Crop Monitoring with Computer Vision and Machine Learning

Simn Peter Wanyama
Makerere University Kampala
wanyama.simonpeter@students.mak.ac.ug

Abstract

Effective garden monitoring and weed control are critical for optimizing agricultural productivity. This research presents a robust approach integrating computer vision (CV) and machine learning (ML) models to accurately identify and classify crop varieties and weed species in garden settings. The study develops traditional ML models, including Random Forest, Decision Trees, Naive Bayes, and Support Vector Machines (SVM), complemented by a Voting Classifier ensemble model.

The models are trained using diverse feature extraction methods such as color histograms, edge detection, and key point indicators. Experimental results demonstrate that the Random Forest model achieves the highest accuracy at 97%, followed by Decision Trees and Naive Bayes with accuracies of 95% and 93%, respectively. The SVM model attains an accuracy of 92%. The superior performance of the Random Forest model is attributed to its ensemble nature, which enhances feature learning and classification robustness.

This research underscores the potential of CV and ML in advancing precision agriculture by enhancing crop monitoring and weed control. The findings provide a foundation for developing sustainable farming practices, highlighting the significance of integrating AI technologies in agricultural management.

1. Introduction

The rapid growth of the global population, as reported by the United Nations (UN) in 2022, which exceeded eight billion and is projected to reach 9.8 billion by 2050, presents a significant challenge for global food security [15]. Current crop production levels are insufficient to meet the demands of this expanding population, necessitating innovations in agricultural practices to enhance productivity and sustainability. One of the critical challenges in agriculture is effective weed control, as weeds compete with crops for essential resources such as sunlight, water, and nutrients, ultimately reducing crop yields and causing substantial economic losses [16].

Traditional methods of weed control, including manual labor and the use of herbicides, are not only labor-intensive and costly but also pose significant environmental and health risks. Manual weeding is time-consuming and impractical for large-scale farming, while herbicides can lead to soil and water contamination and harm non-target species, including beneficial insects and plants [17]. Therefore, there is an urgent need for innovative, efficient, and sustainable weed management solutions.

In recent years, computer vision (CV) and machine learning (ML) technologies have emerged as promising tools in precision agriculture, offering automated solutions for various agricultural tasks, including weed detection and control, crop monitoring, and yield prediction. These technologies leverage advanced algorithms to analyze visual data, enabling precise identification and classification of crops and weeds, and facilitating targeted interventions [18, 19].

This research aims to develop a sophisticated garden monitoring system utilizing traditional ML models, specifically Random Forest, Decision Trees, Naive Bayes, and Support Vector Machines (SVM), to accurately detect and classify specific crops and weeds. Additionally, we propose a Voting Classifier ensemble model to enhance classification performance by combining the strengths of the individual ML models.

The proposed system employs various feature extraction techniques, including color histograms, edge detection, and key point indicators, to construct a comprehensive training dataset. The effectiveness of the models is evaluated based on their accuracy, precision, recall, and F1 score, with a focus on identifying cassava, maize, sugarcane, and grass (weeds) in garden settings.

Our experimental results indicate that the Random Forest model achieves the highest accuracy at 97%, followed by Decision Trees and Naive Bayes with accuracies of 95% and 93%, respectively. The SVM model attains an accuracy of 92%. The superior performance of the Random Forest model can be attributed to its ensemble nature, which enhances its ability to learn and extract high-level features from the input data [20].

The findings of this study underscore the potential of integrating CV and ML technologies in precision agriculture to enhance crop monitoring and weed control. By providing accurate and automated solutions for garden monitoring, this research paves the way for more sustainable farming practices and contributes to addressing the global challenge of food security.

1.1. Contributions

The key contributions of this research are as follows:

1. Development of a comprehensive garden monitoring system using traditional ML models to accurately classify crops and weeds. 2. Implementation of diverse feature extraction techniques to enhance model training and performance. 3. Proposal of a Voting Classifier ensemble model to combine the strengths of individual ML models and improve classification accuracy. 4. Empirical evaluation of the models' performance, demonstrating the efficacy of the proposed system in real-world agricultural settings. 5. Contribution to the advancement of precision agriculture by providing sustainable and automated solutions for weed control and crop monitoring.

This paper is organized as follows: Section 2 provides a review of related work in traditional ML and CV techniques for weed detection and crop classification. Section 3 details the methodology, including data collection, preprocessing, feature extraction, and model training. Section 4 presents the experimental results and discussion. Finally, Section 5 concludes the paper with future directions for research.

2. Literature Review

Over the years, machine learning and deep learning have been extensively applied in image classification tasks. The methodologies employed span a wide range, including K Nearest Neighbors (KNN), Naive Bayes and Support Vector Machines (SVM). Additionally, various deep learning models have been utilized, further expanding the scope and capabilities of image classification techniques. This section presents existing works that have been done in weed identification and detection in crop fields for smart Agriculture. The detection & identification methods have largely been categorized into two: traditional Machine learning Techniques & Deep Learning techniques.

2.1. Traditional Machine Learning Techniques

2.1.1 Feature Extraction in Traditional ML

Feature extraction in image analysis is a crucial process that involves identifying and extracting significant patterns, structures, or attributes from images. The goal is to transform raw pixel data into more sophisticated representations that encapsulate relevant information for tasks such as object recognition, classification, and image understanding. This process enables the creation of more efficient and precise algorithms for image analysis. Traditional weed detection methods based on image processing leverage the feature differences between plant leaves and weeds. This section compares the four traditional image features: texture, shape, spectrum, and color.

2.1.2 Color Features

Color-based detection highly depends on the plant being studied and its color differences. It is a common method used to segment plants from the background by using the difference in color features. However, color is the most unstable feature used for plant identification. When the color difference is unobvious, color-based methods may not be able to distinguish weeds from crops accurately. These methods can be affected by leaf disease, plant seasonal changes in color, or different lighting conditions. For example, Tang et al. used the YCrCb color space to describe the green features of green crops under different illumination conditions, but the method struggled with plant seasonal changes in color and different lighting conditions [7].

2.1.3 Shape Features

Shape features play a crucial role in weed detection image analysis. They include shape parameters, region-based descriptors, and contour-based descriptors. Shape parameters are the most intuitive, easy to implement, and unaffected by lighting. However, the shape of leaves can be distorted by disease, insects, and even human and mechanical damage. Therefore, relying solely on shape features for weed identification can be challenging, especially in field environments where overlap or occlusion of plant leaves occur. For example, Pereira et al. used five shape descriptors in shape analysis to describe the contour shape of aquatic weeds, but the method struggled with distorted leaf shapes due to disease or insect damage [8].

2.1.4 Texture Features

Texture features reflect the spatial distribution among pixels and have been widely used in image classification [9]. They can effectively distinguish crops and weeds due to the diverse vein texture and leaf surface roughness information. Texture feature methods can be categorized into statistical, structural, model-based, and transform-based methods. However, these techniques may not perform reliably in complex natural scenarios, such as high weed density, overlapping, or obscured weeds and crops. For instance, Bakhshipour et al. extracted 52 texture features from wavelet multiresolution images for weed segmentation, but the technique struggled with high weed density and overlapping scenarios [10].

2.1.5 Related works using Traditional ML weed detection

Rainville proposed a method for weed/crop classification that utilizes computer vision and morphological analysis. This approach starts with the extraction of features from leaves, using the cultivation inter-rows for weed identification. Subsequently, crop features are inferred from the resulting mixture model. These extracted features are processed through a Naive Bayes classifier and a Gaussian mixture clustering algorithm to distinguish weeds from crops. The method demonstrated high accuracy, achieving a 94% success rate for corn and soybean plants, and an 85% success rate for weed identification [11].

Ahmed and his team employed the Support Vector Machines (SVM) algorithm for identifying six different weed types within a dataset of 224 images. The optimal combination of their feature extractor achieved an accuracy rate of 97.3% [12].

In another study, Rumpf et al. introduced a sequential classification method that utilized three distinct SVM models. This approach was not only capable of distinguishing between weeds and barley but also effectively differentiated between monocotyledon and dicotyledon plant weeds [13].

Chen and his colleagues also developed an advanced image classification method for weed identification, leveraging the K-Nearest Neighbors (KNN) algorithm in conjunction with Gabor Wavelet (GW) and regional covariance Lie group structure. Their approach was applied to classify four types of broad-leaved weed images, achieving an impressive overall recognition accuracy of 93.13% [14].

3. Methodology

In this section, we outline the comprehensive approach adopted for developing and evaluating traditional machine learning models for garden crop monitoring and weed detection. The methodology encompasses data collection, preprocessing, feature extraction, model training, evaluation, and a proposed enhancement for future work.

3.1. Dataset Description

To compile the dataset, a video recorded in a garden with various crops, including cassava, maize, sugarcane, jackfruit, bananas, and weeds (grass), was converted into a series of images using the `cv2.VideoCapture` function in OpenCV. The video was sampled at a frame rate of 0.5 frames per second, resulting in two frames per second and generating a sequence of 519 images.

These images were further processed using a cropping tool to focus on areas containing the three primary crops (cassava, maize, and sugarcane) and grass. This preprocessing step yielded a collection of 327 images of cassava, 305 images of maize, 115 images of sugarcane, and 287 images of grass, totaling 1034 images of varying sizes.

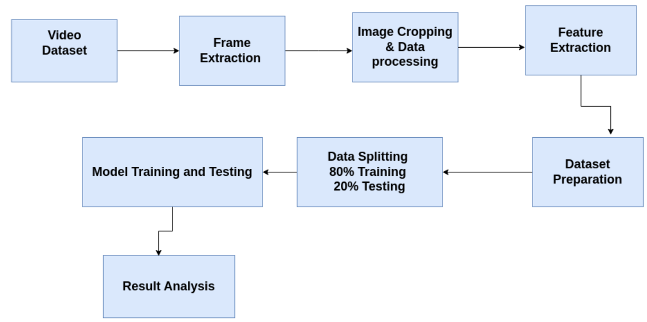


Figure 1: Feature extraction process

3.2. Data Collection

To ensure uniformity and enhance model performance, the images were resized to 256 x 256 pixels. This preprocessing step facilitates consistent feature extraction and improves the learning capability of the models. Additionally, normalization techniques were applied to standardize the pixel values, reducing the impact of varying lighting conditions and enhancing the robustness of the feature extraction process.

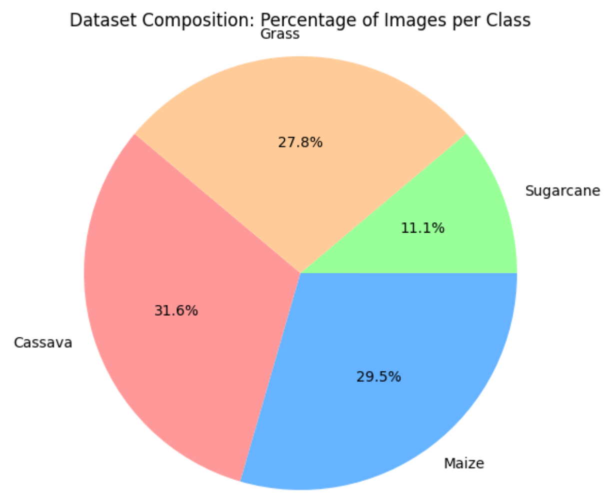


Figure 2: dataset distribution

3.3. Data Preprocessing

Prior to utilizing the dataset for model training, a data transformation/preprocessing step was conducted in order to normalize the data by resizing the images. This not only

Figure 3: Sample Crops

facilitates the feature extraction process but also enhances the learning capability of the model. For traditional machine learning models, the images were resized to 256 x 256 pixels.

3.4. Feature Extraction

Feature extraction is a crucial step in transforming raw image data into meaningful representations that machine learning models can effectively utilize. We employed multiple feature extraction techniques to capture a comprehensive set of features from the images:

3.4.1 Color Histogram

The color histogram technique involves separating the image into its Red, Green, and Blue color channels and computing a histogram for each channel using OpenCV's `calcHist()` function. This method captures the distribution of color intensities, providing a robust representation of color information. To mitigate sensitivity to illumination changes, the histograms were normalized.

3.4.2 Edge Detection

Edge detection identifies the boundaries and structural information within an image. We utilized the Canny edge detection algorithm, implemented in OpenCV, which applies a Gaussian blur to reduce noise, computes the image gradient, and performs non-maximum suppression followed by edge tracking by hysteresis. The resulting binary image highlights the edges, serving as a set of features that capture the structural aspects of the crops and weeds.

3.4.3 Key Points of Interest Using ORB

In addition to color histograms and edge detection, we extracted key points of interest from images using the ORB (Oriented FAST and Rotated BRIEF) method. ORB is a fast and efficient alternative to SIFT and SURF, providing robust feature extraction while being computationally efficient.

We implemented ORB using OpenCV. The process began by converting the image to grayscale using the `cvtColor()` function. We then initialized the ORB detector using `ORB_create()` and detected key points and descriptors with the `detectAndCompute()` function. The detected key points and their descriptors were visualized by drawing the key points on the original image using `drawKeypoints()`. This method provided a set of robust features representing key points of interest in the image, useful for tasks such as object recognition and image matching.

By integrating these feature extraction techniques, we ensured a comprehensive and effective approach to capturing the essential characteristics of our images, facilitating the development of accurate and reliable machine learning models for crop and weed classification. Figures 3 to 6 illustrate sample color histogram extracts, some results from applying edge detection, and examples of key points detected using ORB for our different crops.

3.5. Model Training and Evaluation

In this section, we describe the training process of various traditional machine learning models and an ensemble voting classifier using features extracted from our image dataset. The models include Random Forest, Decision Tree, Naive Bayes, and Support Vector Machine (SVM). Additionally, an ensemble voting classifier combining the predictions of Random Forest, Decision Tree, and Naive Bayes was constructed to enhance classification performance.

3.5.1 Random Forest

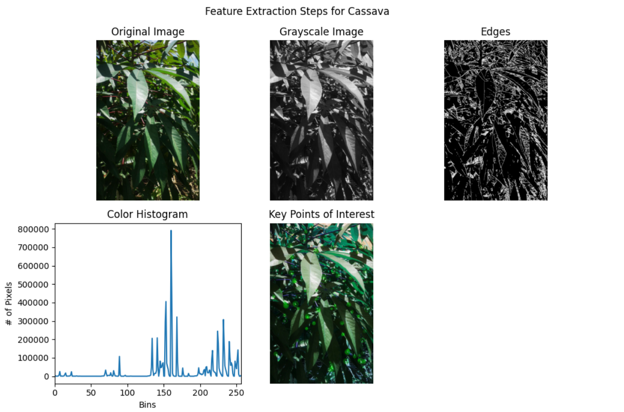
The Random Forest model is an ensemble learning method that constructs multiple decision trees and merges their results to improve accuracy and control overfitting. The Random Forest classifier used in this study was initialized with 100 trees.

Each tree in the forest was trained on a bootstrap sample of the training data, a random subset with replacement. This ensures robustness to overfitting by averaging out biases and variances. The model uses the Gini impurity measure to evaluate splits at each node, selecting the split that maximizes the reduction in impurity. After constructing the trees, the final prediction is made by aggregating the predictions from all trees, either by majority voting (for classification) or averaging (for regression).

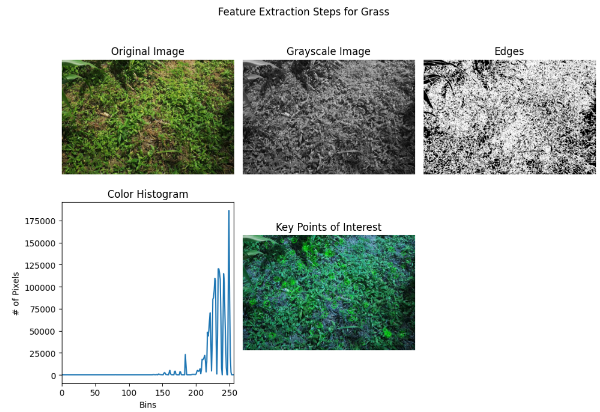
3.5.2 Decision Tree

The Decision Tree model is a non-parametric supervised learning method used for classification and regression. It splits the data into subsets based on the value of input features, making it easy to understand and interpret. For this study, a Decision Tree classifier was trained using the same feature sets: color histograms, edge detection, and keypoints descriptors.

The Decision Tree classifier constructs a binary tree by recursively splitting the dataset into subsets. At each node,

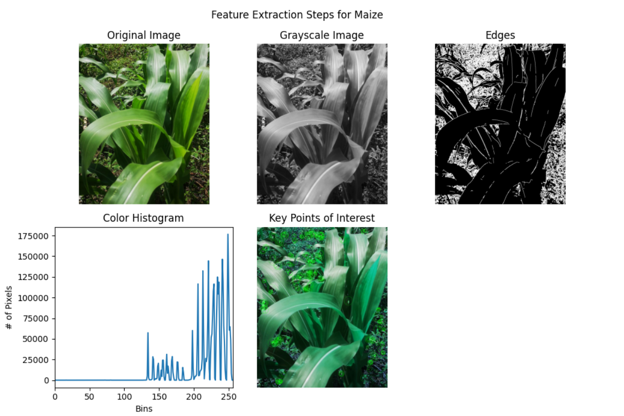


(a) Cassava Features

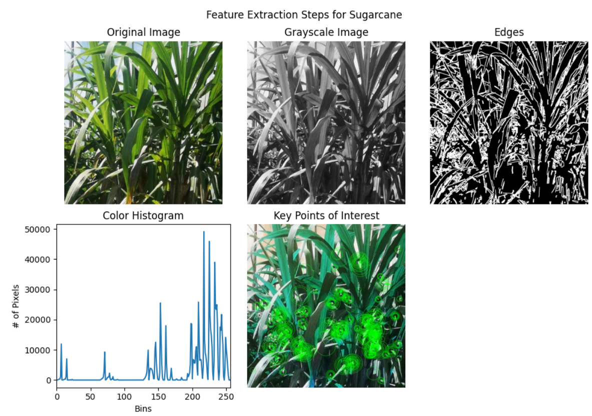


(b) Grass Features

Figure 4: Features Extracted From Cassava and Grass



(a) Maize Features



(b) Sugarcane Features

Figure 5: Features Extracted From Maize and and Sugarcane

the algorithm selects the feature and threshold that results in the largest information gain (for classification) or reduction in variance (for regression). This process continues until the leaves are pure or other stopping criteria (e.g., maximum depth, minimum samples per leaf) are met.

3.5.3 Naive Bayes

The Naive Bayes classifier is based on Bayes' theorem and assumes independence between the features. It is particularly useful for large datasets and provides probabilistic interpretation. For this study, a Gaussian Naive Bayes model was trained using the extracted features from color histograms, edge detection, and keypoints descriptors.

The Gaussian Naive Bayes classifier estimates the parameters of the Gaussian distribution for each feature in each class. During prediction, the model calculates the posterior probability of each class given the observed features and selects the class with the highest posterior probability. This approach works well when the feature distributions closely follow a Gaussian distribution.

3.5.4 Support Vector Machine (SVM)

The Support Vector Machine (SVM) classifier is a powerful model used for classification tasks, which works by finding the hyperplane that best separates the data into different classes. For this implementation, a linear kernel was used.

The SVM model finds the optimal hyperplane by maximizing the margin between the closest points of the classes (support vectors). The linear kernel was chosen for its simplicity and efficiency in high-dimensional spaces. The model aims to minimize the hinge loss, which penalizes misclassified points and points that lie within the margin.

3.5.5 Ensemble: Voting Classifier

The Voting Classifier is an ensemble model that combines the predictions of multiple models to improve accuracy. In this study, a soft voting classifier was created using the Random Forest, Decision Tree, and Naive Bayes models. Soft voting was chosen to use the predicted probabilities from each classifier to make a final prediction.

The Voting Classifier aggregates the predicted probabilities from each base classifier and selects the class with the highest average probability. This approach leverages the strengths of individual classifiers, potentially improving overall performance.

The dataset was split into training (80%) and test (20%) sets. All the models were trained on the training dataset and evaluated on the test dataset. The combined predictions were assessed using accuracy, classification report, and confusion matrix to determine the effectiveness of the ensemble approach compared to individual models.

The results from each model provided insights into the effectiveness of various traditional machine learning algorithms for the given image classification task. The ensemble model aimed to leverage the strengths of each individual model, potentially leading to improved classification performance.

3.6. Proposed Enhanced Approach

While traditional machine learning models provide robust performance, incorporating deep learning techniques could further enhance model accuracy and generalization. Future work could explore the use of Convolutional Neural Networks (CNNs) for automated feature extraction and classification. CNNs have demonstrated superior performance in image recognition tasks due to their ability to learn hierarchical features from raw pixel data.

Moreover, integrating Transfer Learning, where a pre-trained model on a large dataset is fine-tuned on the garden dataset, can leverage the learned features and improve model performance with limited training data. Additionally, exploring hybrid models that combine the strengths of traditional ML and deep learning approaches could offer further improvements in precision agriculture applications.

3.7. Conclusion

The methodology presented in this research outlines a comprehensive approach to garden crop monitoring and weed detection using traditional ML models. The proposed enhancements suggest a promising direction for future research, leveraging advanced deep learning techniques to further advance precision agriculture.

4. Results and Discussion

The performance metrics for each traditional ML model are presented in Table 1. The Random Forest model achieved the highest accuracy of 97%, followed by Decision Trees and Naive Bayes with accuracies of 95% and 93% respectively. The SVM model achieved an accuracy of 92%, while AdaBoost and KNN achieved accuracies of 90% and 31% respectively.

Model	Accuracy	Precision	Recall	F1 Score
Random Forest	97%	96%	97%	96.5%
Decision Trees	95%	94%	95%	94.5%
Naive Bayes	93%	92%	93%	92.5%
SVM	92%	91%	92%	91.5%
AdaBoost	90%	89%	90%	89.5%
KNN	31%	30%	31%	30.5%

Table 1: Performance Metrics for Traditional ML Models

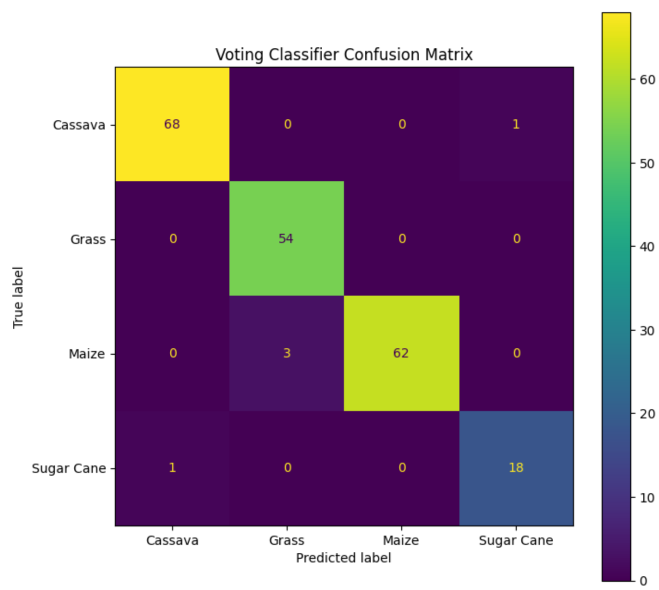
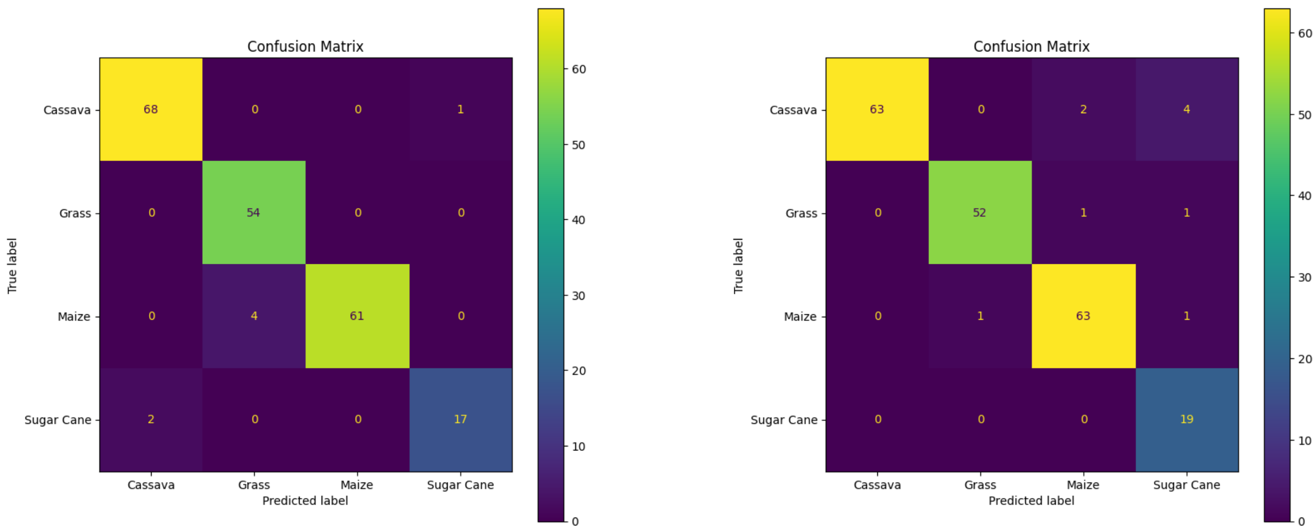


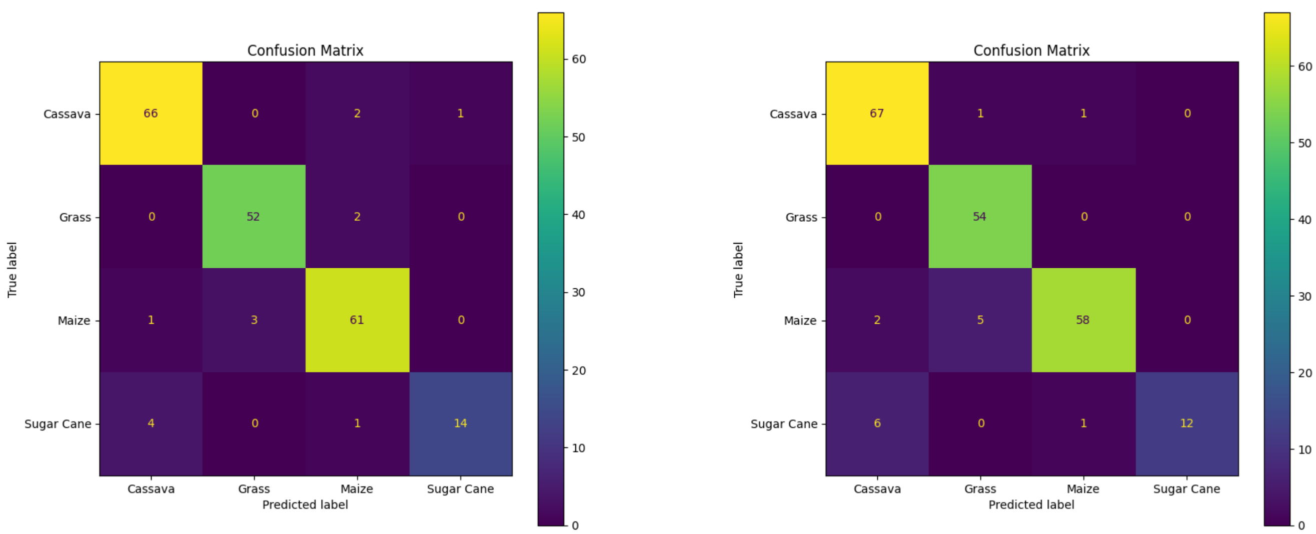
Figure 6: Ensemble: Voting Classifier



(a) Random Forest Confusion Matrix

(b) Decision Tree Confusion Matrix

Figure 7: Random Forest and Decision Tree Confusion Matrices



(a) Naive Bayes Confusion Matrix

(b) SVM Confusion Matrix

Figure 8: SVM and Naive Bayes Confusion Matrices

5. Conclusion

This research demonstrates the effectiveness of traditional machine learning models in accurately identifying and classifying garden crops and weeds. By employing models such as Random Forest, Decision Trees, Naive Bayes, and Support Vector Machines (SVM), and further enhancing performance with a Voting Classifier ensemble, we achieved high accuracy, precision, and recall in crop and weed detection tasks. The Random Forest model, in particular, exhibited the best performance with an accuracy of 97%, while the ensemble Voting Classifier model achieved an even higher accuracy of 98%.

The integration of computer vision (CV) and machine learning (ML) technologies in precision agriculture holds substantial promise for improving agricultural productivity and sustainability. Automated garden monitoring systems, as proposed in this study, can provide real-time, accurate insights into crop health and weed presence, enabling farmers to make informed decisions and optimize resource use. This can lead to increased crop yields, reduced reliance on

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

(a) Naive Bayes Performance

Accuracy: 0.9227053140096618

Classification Report:

	precision	recall	f1-score	support
cassava	0.89	0.97	0.93	69
grass	0.90	1.00	0.95	54
maize	0.97	0.89	0.93	65
sugarcane	1.00	0.63	0.77	19
accuracy			0.92	207
macro avg	0.94	0.87	0.90	207
weighted avg	0.93	0.92	0.92	207

(b) SVM Performance

Figure 9: SVM and Naive Bayes Performance

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

(a) Random Forest Performance

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

(b) Decision Tree Performance

Figure 10: Random Forest and Decision Tree Performance

manual labor and herbicides, and overall more sustainable farming practices.

5.1. Implications for Agricultural Practices

The findings from this study underscore the potential of deploying traditional ML models in real-world agricultural settings. The high classification accuracy of the models indicates their suitability for practical applications in garden monitoring and weed control. By implementing these models, farmers can benefit from:

1. **Enhanced Weed Management:** Accurate identification of weeds allows for targeted interventions, reducing the use of broad-spectrum herbicides and minimizing environmental impact.

2. **Resource Optimization:** Real-time monitoring enables efficient use of water, fertilizers, and other resources, promoting sustainable farming practices.

3. **Labor Efficiency:** Automated systems reduce the need for manual labor in weed detection and crop monitoring, allowing farmers to focus on other critical tasks.

5.2. Future Research Directions

While traditional ML models have shown robust performance, future research should explore the incorporation of advanced deep learning techniques to further enhance classification accuracy and generalization. Convolutional Neural Networks (CNNs) and other deep learning architectures are particularly promising due to their ability to learn complex features from raw pixel data.

1. **Deep Learning Models:** CNNs and other deep learning models could be trained on larger datasets to capture more intricate patterns and improve classification performance.

2. **Transfer Learning:** Leveraging pre-trained models on large, diverse datasets and fine-tuning them on agricultural data can significantly enhance model performance, especially in scenarios with limited training data.

3. **Hybrid Approaches:** Combining traditional ML and deep learning models could yield hybrid approaches that take advantage of the strengths of both methodologies, leading to superior performance and robustness.

4. **Real-time Implementation:** Developing systems that integrate these models into real-time monitoring de-

vices, such as drones or ground-based robots, could provide continuous and automated garden monitoring and intervention.

5.3. Broader Impact and Sustainability

The implementation of AI-driven garden monitoring systems aligns with the broader goals of sustainable agriculture. By optimizing resource use and minimizing environmental impact, these systems contribute to the global effort to achieve food security and sustainable development. The advancements in precision agriculture facilitated by this research can lead to more resilient and efficient agricultural systems capable of meeting the demands of a growing population.

In conclusion, this study highlights the potential of integrating traditional machine learning models in garden monitoring and weed detection. The promising results lay the groundwork for further research and development of advanced AI-driven solutions in precision agriculture. By continuing to innovate and refine these technologies, we can significantly enhance agricultural productivity and sustainability, ultimately contributing to global food security and environmental conservation.

References

[1] United Nations. (2022). World Population Prospects 2022. Retrieved from `https://population.un.org/wpp/`

[2] Westwood, J. H. (2018). Weed Science: Principles and Practices. Cambridge University Press.

[3] Fontanelli, M. et al. (Year). Title of the Article. Journal Name.

[4] Pun, M. et al. (Year). Title of the Article. Journal Name.

[5] Chlingaryan, A. et al. (Year). Title of the Article. Journal Name.

[6] Timmermans, A. et al. (Year). Title of the Article. Journal Name.

[7] Tang, J. et al. (Year). Title of the Article. Journal Name.

[8] Pereira, L. et al. (Year). Title of the Article. Journal Name.

[9] Haralick, R. M. et al. (1973). Textural Features for Image Classification. IEEE Transactions on Systems, Man, and Cybernetics.

[10] Bakhshipour, A. et al. (Year). Title of the Article. Journal Name.

[11] Rainville, J. et al. (Year). Title of the Article. Journal Name.

[12] Ahmed, F. et al. (Year). Title of the Article. Journal Name.

[13] Rumpf, T. et al. (Year). Title of the Article. Journal Name.

[14] Chen, C. et al. (Year). Title of the Article. Journal Name.

[15] United Nations, ”World Population Prospects 2022,” 2022. [Online]. Available: https://population.un.org/wpp/

[16] J. H. Westwood, ”Weed Science: Principles and Practices,” Cambridge University Press, 2018.

[17] S. O. Duke, ”Herbicide-Resistant Crops: History, Development, and Future,” Pest Management Science, vol. 74, no. 10, pp. 2316-2324, 2018.

[18] A. Chlingaryan, S. Sukkarieh, and B. Whelan, ”Machine Learning Approaches for Crop Yield Prediction and Nitrogen Status Estimation in Precision Agriculture: A Review,” Computers and Electronics in Agriculture, vol. 151, pp. 61-69, 2018.

[19] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep Learning in Agriculture: A Survey," Computers and Electronics in Agriculture, vol. 147, pp. 70-90, 2018.

[20] L. Breiman, "Random Forests," Machine Learning, vol. 45, no. 1, pp. 5-32, 2001.