International Journal of Applied and Advanced Multidisciplinary Research (IJAAMR)
Vol. 1, No. 4, 2023: 327-332

Machine Learning Based Weed Detection System

Dr. Meenakshi Thalor[1]*, Prathamesh Gajbhiye[2]
Department of Information Technology, AISSMS Institute Of Institute Of Information Technology, Pune, Maharashtra, India.

Corresponding Author: Prathamesh prathamesh280129@gmail.com

ARTICLEINFO

Keywords: Weed Detection, Crop Cultivation, Machine Learning Algorithms, YOLOv5 Neural Network, Classifier Accuracy.

Received : 22, October
Revised : 20, November
Accepted: 27, December

ABSTRACT

This abstract underscores the importance of weed detection in crop cultivation to prevent plant diseases and minimize crop losses. To address these challenges and promote eco-friendly practices, the authors propose a weed detection program employing K-Nearest Neighbors, Random Forest, Decision Tree algorithms, and the YOLOv5 neural network. The abstract also provides a concise overview of existing research in weed identification using machine learning and deep learning. The authors developed a YOLOv5-based weed detection system and evaluated the performance of the algorithm, showing traditional classifiers achieve accuracies of 83.3%, 87.5%, and 80%, while the neural network scores range from 0.82 to 0.92 for each class. The study demonstrates the effectiveness of this approach in classifying low-resolution weed images.

DOI: https://doi.org/10.59890/ijaamr.v1i4.568
https://journal.multitechpublisher.com/index.php/ijaamr

INTRODUCTION

The agricultural sector is a crucial part of our country's economy, contributing 35-40% of the annual income to the state budget and employing 15% of the total workforce. With the rise of agricultural robotization, controlling weeds and monitoring crop diseases has become a pressing issue. Detecting and preventing diseases and minimizing crop losses during cultivation stages are vital, as traditional methods are costly, labor-intensive, and can expose workers to harmful chemicals. Therefore, the primary focus in the agricultural industry is the development of pest control systems that can detect and eliminate weeds.

Currently, the most effective pest control method is the widespread use of herbicides. However, these methods often do not consider variations in weed growth, leading to unnecessary chemical treatment of crops, which can harm the environment. Previously, technologies could only distinguish between the presence or absence of plants and could not differentiate between weeds and agricultural crops. Modern technologies now enable more precise herbicide application, targeting only the areas with weeds to preserve crops and protect the environment. Implementing intelligent weed detection systems will also help conserve herbicides and pesticides, which are essential for combating plant diseases, various weeds, and disease vectors in industrial agriculture. Utilizing autonomous robots and automated systems in agriculture has the potential to significantly reduce human efforts required for various agricultural tasks. To address these challenges, new classification systems have been introduced to identify agricultural crops and distinguish them from unwanted harmful vegetation.

Agriculture, as the second-largest sector of material production, involves over a billion people in the cultivation of grains, vegetables, and fruits. The introduction of automation and robotization in this industry holds great potential for enhancing efficiency and can even replace heavy agricultural machinery. Implementing systems capable of distinguishing weeds from agricultural crops can lead to substantial savings in chemicals by targeting weed-infested areas accurately. Additionally, processes such as harvesting and sorting by crop quality and size often result in increased expenses, driving up the overall cost of crops. Therefore, it is essential to conduct scientific research in the field of weed detection and discrimination to enable swift and precise operations in agricultural fields.

The outcomes of this research pave the way for the deployment of robot systems, which promise to boost productivity, reduce material waste, and enhance crop quality through the principles of precision farming.

LITERATURE REVIEW

In the scientific paper, the authors employed 3D Otsu's method for segmentation to differentiate between agricultural crops and weeds. Classification was then achieved by compressing three-dimensional image vectors through Principal Component Analysis (PCA). By combining these methods, the authors devised a real-time weed detection program.

In the authors conducted a comprehensive review of computer vision methods for locating weeds and crops. This study assessed the pros and cons of algorithms based on deep learning and examined their utility in weed detection. Notably, AgroAVNET (a hybrid of AlexNet and VGGNET), Graph Weeds Net, MTD (multiscale triangle descriptor), and LBP-HF (local binary pattern histogram Fourier) methods exhibited high accuracy in weed detection compared to other algorithms. This paper presented a thorough comparative analysis of various architectures and algorithms using the same weed dataset. A similar study, employed deep learning in an image processing framework to classify various crops and weeds, achieving an impressive 95% accuracy with a convolutional neural network and max pooling layers. This approach notably reduced the frequency of misclassifying weeds.

In a model for detecting weeds in agricultural crop fields based on deep learning was proposed. The authors utilized a convolutional neural network for extracting informative features and performed data augmentation to create additional training images. They employed Inception V3 as a feature extractor and used U-Net for weed classification. The system achieved a detection accuracy of 90% when tested on 158 images. However, it's worth noting that the authors did not compare their system with works by other researchers, and it could benefit from testing on larger datasets under varying lighting conditions. The research outlined in holds significant relevance for the development of weed seed detection systems, as precise classification of weed seeds plays a pivotal role in controlling crop pests. During this investigation, six different deep convolutional neural network models were compared to identify the most effective method for weed species detection. The results indicated that the GoogLeNet architecture exhibited notably high accuracy, while the SqueezeNet architecture proved superior in terms of seed and background detection speed. However, there is potential for further improvement in identification accuracy by considering scenarios involving the fusion of various seed types and their close proximity.

METHODOLOGY

Segmentation

For image segmentation, the Otsu's method is employed . This adaptive algorithm is based on binarization and relies on the maximum variance value, which represents the deviations in average brightness between the background and the selected image, to determine the threshold for segmentation. The process begins by splitting the image into foreground and background based on its grayscale characteristics. The goal is to find the optimal threshold value where the difference between these two parts is maximized. Minimizing the probability of incorrect classification occurs when this difference reaches its maximum. The segmentation of images using Otsu's method is executed as follows:

1. The original image is divided into l=[0, 1, ..., l–1] levels. The number of pixels at a certain level i is denoted by ni,and the total

number of pixels is denoted by the sum of all ni,that is, n=n1+n2+...+nl.

2. The pixels in an image are categorized into classes based on their gray levels, and this categorization can be done both without using a specific threshold and with a threshold. The process involves calculating the probability of each gray level, which is then used to assign these levels to their respective classes.

Classification

Three classical machine learning algorithms, namely K-Nearest Neighbors, Random Forest, and Decision Tree, were employed for weed classification. To ensure the classifier's effectiveness, it was necessary to use a class-balanced training set . This was achieved by implementing a random sampling process that selected an equal number of objects for each class throughout the entire image.

Subsequently, various features were extracted from each object composing the training set. These features included NIR/G (Near-Infrared/Green) values, average red, average green, average NIR, brightness, and the standard deviation of NIR. These features were utilized in distinguishing between weeds, crops, and bare soil using the Random Forest (RF) algorithm. The training process involved the use of 400 decision trees, a value found to be suitable for the RF classifier. Notably, the height of the object was excluded as a classification parameter to prevent any misclassification of large weeds, whether they were located between or within rows.

Evaluation of Algorithm

Various metrics have been developed to evaluate the performance of machine learning algorithms. These metrics are typically calculated using combinations of the confusion matrix, which provides information about the true positives, true negatives, false positives, and false negatives generated by the classifier.

Metrics such as False Positive Rate (FPR), False Negative Rate (FNR), Recall, Precision, Accuracy, and a few others are calculated using different combinations of the values in the confusion matrix. However, these metrics may not always provide an objective assessment of classification results.

On the other hand, metrics like F1 score, Cohen's kappa, and the Matthews correlation coefficient utilize all elements of the confusion matrix and are particularly useful for assessing classifier performance in the presence of unbalanced data. Here are the formulas for these metrics:

1. **False positive Rate:** FPR =FP/(FP+TN)
2. **False Negative Rate:** FNR=FN/(FN+TP)
3. **Recall:** Recall=TP/(TP+FN)
4. **Precision:** Precision=TP/(TP+FP)
5. **Accuracy:**Accuracy=(TP+TN) /(TP+FP+FN+TN)

RESEARCH RESULT AND DISCUSSION

In this study, the development of weed and crop detection systems in modern agriculture was addressed. The YOLOv5 neural network architecture was used for weed identification. A dataset of over 5,000 images of common weeds was collected. 80% of the images were used for training, and 20% for validation. Image segmentation and preprocessing were performed, followed by classification using KNN, Random Forest, and Decision Tree algorithms, each with varying accuracy results.

The YOLOv5 architecture was introduced, which demonstrated significant improvements in size and speed compared to its predecessors. The accuracy of weed detection was evaluated, and YOLOv5 outperformed other algorithms. The proposed architecture incorporated an attention module based on the ECA-Net algorithm to enhance performance. Comparisons with similar work by other researchers showed that this system achieved higher accuracy in weed detection. The dataset's low-resolution images were noted as an advantage for the system.

However, a limitation was identified due to the small number of classes used in training. Future research plans involve expanding the dataset with images of different weed species, growth stages, and increasing the number of classes to improve accuracy. Additionally, the study considered integrating this weed detection system with a tomato recognition system to create a comprehensive robotic complex capable of distinguishing between crop pests and cultivated plants, which aligns with earlier proposed agricultural robot mechanisms. Lastly, the importance of selecting high-performance computers and optimizing parameters for implementing the YOLOv5-based detection system in real-world agricultural applications was emphasized.

CONCLUSION

In conclusion, this study focused on identifying and addressing weed-related challenges in agricultural fields. Localities were surveyed to select common weeds like Bindweed, Amaranthus, Bromus, and Ambrosia. Over 1,000 photos were collected for each weed species, forming a valuable image database for future research. The study applied noise reduction techniques during image segmentation and evaluated various classifiers. Results showed that K-Nearest Neighbors achieved 83.3% accuracy, Random Forest 87.5%, and Decision Tree 80%. Additionally, the YOLOv5 neural network with attention modules exhibited strong performance, with class-specific indicators ranging from 0.82 to 0.92. This research culminated in the development of a weed detection system based on YOLOv5, which proved effective in classifying low-resolution weed images, offering a promising solution for weed management in agriculture.

REFERENCES

Dos Santos Ferreira, A., Matte Freitas, D., Et All. (2017). Weed detection in soybean crops using ConvNets. Computers and Electronics in Agriculture, 143, 314–324.

Ma, X., Deng, X., Qi, L., Jiang, Y., Li, H., Wang, Y., Xing, X. (2019). Fully convolutional network for rice seedling and weed image segmentation at the seedling stage in paddy fields. PLOS ONE, 14 (4), e0215676.

Thuan, D. (2021). Evolution of YOLO algorithm and YOLOv5: The state-of-the-art object detection algorithm. Information Technology Oulu University of Applied Sciences, 61.

Wang, A., Zhang, W., Wei, X. (2019). A review on weed detection using ground-based machine vision and image processing techniques. Computers and Electronics in Agriculture, 158, 226–240.

Wang, Q., Wu, B., Zhu, P., Li, P., Zuo, W., Hu, Q. (2020). ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).

Xu, R., Lin, H., Lu, K., Cao, L., Liu, Y. (2021). A Forest Fire Detection System Based on Ensemble Learning. Forests, 12 (2), 217.