Applications of Artificial Intelligence Based Technologies in Weed and Pest Detection

^{1,2}Nidhi Gupta, ²Bharat Gupta, ³Kalpdrum Passi and ⁴Chakresh Kumar Jain

¹Department of Information Technology, Ajay Kumar Garg Engineering College, Ghaziabad, India ²Department of Computer Science and Information Technology, Jaypee Institute of Information Technology, Noida, India ³School of Engineering and Computer Science, Laurentian University, Sudbury, Canada ⁴Department of Biotechnology, Jaypee Institute of Information Technology, Noida, India

Article history Received: 15-05-2022 Revised: 31-05-2022 Accepted: 03-06-2022

Corresponding Author: Chakresh Kumar Jain Department of Biotechnology, Jaypee Institute of Information Technology, A-10, Sector-62 NOIDA, 201307, India Email: ckj522@yahoo.com Abstract: Unprecedented population growth and climate change has burdened food security and scarcity worldwide, where the agriculture sector can significantly contribute to accomplishing the demands and contribute to the economic growth of a country. Artificial Intelligence (AI) has revolutionized the agricultural domain. Pest and weed detection is significant to yielding good quality crops. The AI-based tools and technologies such as drones and robots bring advancement in crop production by performing the early detection of weeds and pests. The tools utilize image processing and machine learning algorithms to capture, analyze and detect the presence of weeds and pests in plants. The research work carried out provides a comprehensive survey for the application of artificial intelligence for both weed and pest detection. It presents widely used techniques, their evaluation parameters, and publicly available datasets which provide the current status of work for the researchers working in the domain of weed and pest detection.

Keywords: Weed Detection, Pest Detection, Machine Learning, Deep Learning, Image Processing

Introduction

Agriculture makes a considerable contribution to a country's economic sector. The world's population is rapidly increasing, resulting in the increased need for food and employment. By 2050, the world's population is predicted to exceed 9 billion people, necessitating a 70% increase in agricultural and food output to meet demand, posing a severe challenge for the agri-food business as detailed by Rockström et al. (2017), Ben Ayed and Hanana (2021). The farmer's traditional practices are not sufficient to meet the increasing demand of the growing population. As a result, new automated procedures are developed. Agriculture has transformed as a result of artificial intelligence. Artificial Intelligence (AI) is the ability to simulate human intelligence using computer systems, robotics, and digital equipment. This technology has protected crop yields against a variety of factors such as climate change, population increase, presence of pests and weeds, etc.

Agriculture has been raised to a new level due to AIbased equipment and tools. The technology brings advancement in crop production which enabled real-time monitoring, harvesting, processing, and marketing. The development of agricultural drones and robots has contributed tremendously to maintaining crop quality and enhanced productivity through automated irrigation, detection of damaged crops, and yield detection. Weeds and Pests in the plants are among the major factors which hinder plant growth.

Farmers faced problems in controlling the weed due to high resistance to herbicides. The sprayers used for pest and weed management apply uniformly over the entire field. However, the weeds are not uniformly distributed, they are generally patchy. The weed control and management ensure the high crop yield with limited use of herbicides. Therefore, weed detection is significant to produce good quality crops and meet the needs of a growing population. The pests prevent the normal growth of the plants. It damages a significant portion of the plant and affects the process of its development from seed to seedling growth.

Machine learning classification is used to automate the process of rapid detection and recognition of pests and weeds from images and videos of the crops. Weed detection is a challenge as crops and weeds are similar in color and size.

Boulent *et al.* (2019) stated that digital image processing is used to process and manipulate the crop's



images to detect the presence of pests and weeds. The identification of pests is an object detection problem and is based on what, where, and how. The 'what' corresponds to the pest's category label, which provides information on the location of the pests and how corresponds to the image segmentation to detect the pests.

Earlier surveys in the literature are performed either on weed or pest detection. To the best of our knowledge, this is the first review that provides the significant work performed in both weed and pest detection domains.

Different Stages of Classification

The three stages for weed and pests identification and categorization are shown in Fig. 1.

The three stages for weed and pests identification and categorization are shown in Fig 1. The different stages are as follows:

- 1. Input stage
- 2. Processing stage
- 3. Output stage

Input Stage

The input stage consists of a dataset made up of images of weeds or pests taken on a farm. The dataset may contain images of different types of plants having weeds or pests. The different categories of weeds and pests are used for the training dataset. To provide effective training, the input image datasets may be of different sizes, soil environments, diverse visual characteristics, and different growth stages.

The image data is generally collected through Sensors and camera-mounted UAVs. These images vary i.e., RGB, thermal, multi-spectral, hyper-spectral, 3D, and chlorophyll fluorescence, and are captured on type of analysis. Islam *et al.* (2021) have selected RGB images which were captured from cameras mounted in a Phantom 3 Advanced drone and equipped with the 1/2.3" CMOS sensor for early weed detection prediction in chili farm

Wu *et al.* (2021) stated that there are several imagebased features such as texture, shape, spectral, color, etc., which play important role in the classification and help in the early detection of weed. However, their feature values vary due to the natural field conditions i.e., high weed density, overlapping, or obscured weeds and crops.

Texture features are related to spatial distribution among pixels and are directly influenced by the shape and texture of leaves. Many researchers have used this feature information for weed recognition and classification. Bakhshipour *et al.* (2017) have demonstrated the usage of around 52 wavelet texture features in a sugar beet crop and discussed the significance of effective classification in weed recognition. Similarly, Ishak *et al.* (2019) identified the new vector feature on directional texture based on the combination of Gabor Wavelet (GW) and Gradient Field Distribution (GFD) methods.

Torres-Sánchez *et al.* (2015) demonstrated the use of the Object-Based Image Analysis (OBIA) technique with optimal parameters identification through an automatic thresholding algorithm for better classification and detection of weeds from herbaceous plant/crop on the Unmanned Aerial Vehicle (UAV) drone images for large-scale use.

Nursuriati *et al.* (2015) identified that the multi-feature (shape, color, spectral, texture, statistical and structural features, etc.) decision-making fusion recognition method provides better stability and higher accuracy as compared to single-feature recognition approach limiting to feature selection issue. Similarly, Lin *et al.* (2017) described the combination of various features such as texture, spectral, shape, etc., for discrimination of weed from corn crop.

Processing Stage

The processing stage comprises of two different steps:

- a) Image processing
- b) Classification model



Fig. 1: Different Stages of weed or pests detection

a) Image processing: In this, the input image dataset is pre-processed to extract the portion of an image segment to identify the regions of interest. The Region of Interest (RoI) is the image segment that represents the weeds or pests for its detection. The weeds or the pests in the images are required to be isolated from the background (soil). Thereafter, each class of weed in the image is labeled. The classification network uses the labeled images as input. The labeling of images is viewed as its features which are used to perform the classification. The various features used for classification depend on the different classes of weed or pests the image contains. However, the general features of images that are used to perform weed detection are the color of the weed or pests to its background, shape, texture, wavelet transform, etc. In many cases, the fusion of multiple features is also used to achieve acceptable accuracies

Feature Optimization

Feature Selection is a crucial and necessary technique that allows the model to execute fast, eliminate redundancy, eliminate noisy data, reduce overfitting, enhance precision and increase generalization on testing data. The traditional feature selection techniques have been used for classification tasks for decades, however, they are not efficient to reduce the dimensionality, resulting in inefficient predictive models. The new paradigm, uses emerging technologies, such as metaheuristics-based algorithms and hyper-heuristics optimization methods Abiodun et al. (2021). Similarly, Misaghi and Yaghobi (2019) have demonstrated the use of the nature-inspired and chaos theory-based Invasive Weed Optimization algorithm (IWO) for parameters optimization such as standard deviation and logistic chaotic mapping features. In the related work by Saleem et al. (2022) where authors demonstrated the use and impact of image resizers and optimization of the weights of the best-acquired models by initialization techniques, batch normalization, and Deep learning algorithms which consequently led to improved and effective weed detection management system.

These technologies can be applied in weed and pests detection for improving classification accuracy and tackling complex optimization issues in less time

b) Classification model: Classification is a predictive modeling approach that categorizes a set of unseen data into classes. The classification model uses various classification algorithms to train images to classify outputs into several classes. The algorithms apply neural networks, deep learning, etc. The classifier is developed using various algorithms such as K-Nearest Neighbor (KNN), Support Vector Machine (SVM), neural networks, deep learning, Adaboost, etc.

Output Stage

The output stage produces the classifier network as the output of the processing stage. The classifier network produced can detect the required infestation in the plants.

Materials and Methods

Image Segmentation and Feature Extraction

Image segmentation is a technique for breaking down a digital image into subgroups called Image segments. It reduces the image's complexity, making it easier to handle or analyze. Feature extraction refers to the process of defining a set of image characteristics, that will most efficiently represent the information needed for analysis and classification. The image segmentation and feature extraction are applied to the images of the crops to distinguish between soil background and crops. It is performed due to the color difference between the two.

The early works on weed detection were limited to the particular type of crops. They differentiate between the two based on the color and shape of the weeds. Lee et al. (1999) detected the weeds in the tomato crops. Aitkenhead et al. (2003) performed weed detection on carrot crops only. Karger and Shirzadifar (2013) focused on weed detection in corn crops. It achieved 90% accuracy due to the wider size of corn leaves in comparison to weeds. Aravind et al. (2015) identify the weed from the ragi plantation. The region of spraying herbicide is identified using a threshold-based approach. McCarthy et al. (2010) surveyed smart spraying for weeds using machine vision and concluded that the approaches can distinguish between soil and vegetation only due to their color difference. The traditional weed and pest detection techniques faced various challenges such as the small difference between damaged areas and the background, issues with the contrast of the crop images, and noise in the images, which makes identification of objects difficult.

Classification Model

Classification is a predictive modeling approach that categorizes a set of unseen data into classes. In weed detection, the classification approaches include traditional machine learning and deep learning algorithms.

Traditional Algorithms for Weed Detection

The efficiency and effectiveness of the traditional methods are primarily based upon the suitable feature selection and feature extraction techniques apart from the classifier methods. There are several lists of features such as shapes, color size, textures of weed are used for easy discrimination.

Tellaeche *et al.* (2011) and Behmann *et al.* (2014) stated that the classification of crops and weeds is largely performed using SVMs and Artificial Neural Networks

(ANNs). SVM has advantage in solving nonlinear and high-dimensional pattern recognition, as well as in small-sample size problems. Bakhshipour *et al.* (2017) identified that ANNs have a strong learning capability and can give high accuracy with unseen data. Besides these algorithms, other approaches have also been used for the classification of crops and weeds such as K-Nearest Neighbor (KNN) Kazmi *et al.* (2015), random forest Lottes *et al.* (2017), Naive Bayesian algorithm, and Bayesian classifier De Rainville *et al.* (2012) and AdaBoost.

Chen *et al.* (2011) classified four types of broadleaved weed images using an enhanced KNN combined with GW and regional covariance Lie group structure.

Neural Networks and Deep Learning for Weed Detection

In the modern era of increasing agricultural yields, the concept of smart farming has been realized with greater significance. The increasing usage of Machine learning-based algorithms for the classification of images captured through the installed camera on moving objects such as tractors, drones, Unmanned Aerial Vehicles (UAV), etc., are supportive of weed detection as described by Rakhmatuiln et al. (2021). Although real-time practicing of the Machine learningbased image classification provides a solution to image detection, identifying the actual weed in the real field is still a challenge due to the variation in light images, size of weed, angle of light and color-based weed detection. Image Pre-processing is one of the important aspects in the field of machine vision. Panqueba and Medina (2016) studied the use of neural networks and deep learning for the detection of weeds. Much of the research uses image segmentation and a neural network to detect weeds in the crops.

Partel *et al.* (2019) focus on the development of a smart sprayer for weeds detection among vegetation. It helps in spraying only on the affected areas. The approach uses the deep learning neural network to detect the target and non-target plants for effective weed management.

Badhan *et al.* (2021) propose a real-time weed detection approach for onion and cucumber crops. It captures video of the crop field and identifies the frames for weed detection. It applies Convolutional Neural Networks (CNN) and Residual neural network-50 (ResNet50) to train the machine learning classification model on onion and cucumber crops. The results reveal that ResNet-50 showed higher accuracy in comparison to CNN with an accuracy of 90% for Onion crops and 84.6% for cucumber crops.

Islam *et al.* (2021) describe the application of various machine learning-based classifiers such as Random Forest (RF), Support Vector Machine (SVM), and K-Nearest Neighbors (KNN), to detect weeds using UAV images from a chili crop field.

Subeesh *et al.* (2022) demonstrated the use of deep learning-based techniques (Alex Net, Google net, InceptionV3, Exception) for the weed identification from RGB images of bell pepper fields with the varied accuracy of 94.5 to 97.7% of different models where InceptionV3 model exhibited the higher performance with a 97.7% accuracy. This study facilitates further integration of both the herbicide applications and weed management system with more preciseness and accuracy.

Narassiguin *et al.* (2016) applied ensemble learning by training several learners to solve the same problem. It incorporates the boosting algorithm which combines several weaker boosting algorithms to create a more powerful machine learning classifier. Adaboost. M1 and Logit Boost are two of the most used boosting algorithms. Freund and Schapire (1997) proposed Adaboost. M1. It is the generalization of the Adaboost algorithm mainly used for more than two class problems. Friedman *et al.* (2000) proposed Logit Boost as an extension of Adaboost that combines a combination of the boosting method and logistic regression for classification.

Image Datasets for Weed Classification

The classification of crops requires training data. Table 1 presents some of the publicly available datasets that can be trained using deep learning.

Pest Detection

Automatic pest detection is significant for the estimation of the damage caused by pests and taking preventive measures. Many researches focuses on remote monitoring of crops to find the infestations. It is being performed using Aerial images captured using Unmanned Aerial Vehicles (UAVs) as proposed by Vanegas *et al.* (2018). However, it required a high resolution of images to detect small pests. Earlier studies on pest detection applied acoustical analysis for the sound emitted by pests. Digital imaging is extensively used for the manipulation of captured images and their analysis.

The problem of pest detection aims to find the presence or absence of pests in the images. It is a binary classification problem that identifies the presence of an object from all the elements. Detection gives an estimate of the degree of spread of pests in an image. Classification of the pests is performed after their detection. The classification problem is the identification of the kind of pests present in the image. It is a multiclass detection problem that results in the most probable class of pests. The classification is divided into three different steps as follows.

Image Acquisition

Martineau *et al.* (2017) proposed that the type of image captured is one of the significant factors for the identification of techniques for pest classification. It includes the different poses of the images captured, the

shape, orientation, and size are considerable factors in the classification techniques of pests. Table 1 lists some of the publicly available datasets for crops that can be trained using deep learning.

Feature Extraction

The research on feature extraction is carried out in the extraction of the desired area of the image. It includes isolating the pests from their background or foreground image. The extraction is performed using the image segmentation technique. Various studies have been carried out for segmentation tasks. The segmentation can be performed as follows:

- a) Supervised segmentation: It aims to learn to remove the background. The classifier is trained using a set of images with negative or positive background images as stated by Xie *et al.* (2015)
- b) Threshold segmentation: In this, an intensity value is set. The images are split and are grouped into objects or backgrounds according to their intensity value. The thresholding can also be performed using clustering. It formed two or more clusters that define the different regions of the images. Various clustering methods have been implemented for the thresholding such as k-means

by Fina *et al.* (2013), ISODATA by Mayo and Watson (2007), and mean shift by Zhu and Zhang (2011), etc.

- c) User segmentation: Sometimes the user performed the segmentation in which it selects the region for extraction
- d) Edge detection segmentation: It performs the segmentation to get the edge of the image using Sobel filters or order statistics filters as stated by Leow *et al.* (2015)

Classification Techniques

The various classification techniques for detection of different kinds of pests are shown in Table 2.

Image Datasets for Pest Classification

Some of the datasets used for pests and plant diseases are shown in Table 3. The dataset can be used to classify the type of pests.

Evaluation Parameters for Weed and Pest detection

The weed detection algorithms are evaluated and compared based on parameters such as their accuracy, precision, recall, etc. The various parameters of machine learning algorithms on which the weed detection is performed are presented in Table 4.

Table 1. Some of t	the Publicly	z available	datasets for	crops that	can he	trained	using dee	n learning
Table 1: Some of t	the Fublicity	available	ualasets 101	crops mat	can be	uameu	using uee	p learning

Reference	Type of crops	Number of images
Chebrolu et al. (2017)	Weed in sugar beet crops	12340
Lameski et al. (2017)	Carrot with weed	39
dos Santos Ferreira et al. (2017)	Soyabean and weed dataset	400
kovsen et al. (2019)	Red and white clover	39600 (real and synthetic)
Bosilj <i>et al.</i> (2019)	Onions with weed	20
Jiang <i>et al.</i> (2020)	Weed with Corn and lettuce	6800
Sudars <i>et al.</i> (2020)	Food crops and weed	1118

Table 2: Various techniques for pest classification

Table 2. Various teeninques for pest classificatio	11	
Classification technique	Kind of pests	References
Template matching	Whiteflies in crops	Wang <i>et al.</i> (2013)
Template matching	Red palm weevil	Al-Saqer (2011)
Artificial neural network	Beet armyworms	Asefpour and Massah (2013)
Support vector machine	beetles in beet and potato crops	Roldán-Serrato et al. (2018)
Support vector machine	Rice planthoppers	Yao, (2016)
K-means clustering	Thrips in strawberry flowers	Ebrahimi et al. (2017)
K-means clustering	Whiteflies in crops	Wang et al. (2018)

Table 3: Some of the publicly available datasets for pests and plant diseases

Reference	Dataset	Number of images
Hughes and Salathé (2015)	Plant village a database of plant diseases	50,000
	around 14 crop varieties with 26 diseases	
Wu et al. (2019)	IP 102: Insect pest recognition database	75,000 related to 102 pests species
Shah et al. (2016;	Rice leaf diseases data set: Three classes of diseases:	
Prajapati et al., 2018)	Bacterial leaf blight, brown spot and leaf smut	40 images for each class
Thapa <i>et al.</i> (2020)	apple leaf disease	3651 images
Wang et al. (2021)	Agri Pest: A dataset of wild pest images	49700 images containing 14 species of pests
Huang and Chuang (2020)	tomato pest images	8 classes

S. no Metric Description 1. Classification accuracy It is the percentage of correctly identified crops among the input 2. True Positive (TP) It is the number of records that correctly predict a positive class. It indicates the presence of weed or pests class when it is not a crop 3. True Negative (TN) It is the number of records that correctly predicts a negative class. It detects non-weed or no-pests class when it is crop actually 4. False Positive (FP) It is the number of records when a weed or pest is incorrectly detected when it is a crop 5. False Negative (FN) It is the number of records when the crop is incorrectly detected when it is a weed or pest	Table 4: Ev	aluation parameters	
1.Classification accuracy True Positive (TP)It is the percentage of correctly identified crops among the input It is the number of records that correctly predict a positive class. It indicates the presence of weed or pests class when it is not a crop3.True Negative (TN)It is the number of records that correctly predicts a negative class. It detects non-weed or no-pests class when it is crop actually4.False Positive (FP)It is the number of records when a weed or pest is incorrectly detected when it is a crop5.False Negative (FN)It is the number of records when the crop is incorrectly detected when it is a weed or p6.Precision $\frac{TP}{TP + FP}$ 7.Recall $\frac{TP}{TP + FP}$ 8.Confusion matrixIt describes the performance of an algorithm in terms of the number of TP, TN, FP, ar	S. no	Metric	Description
2.True Positive (TP)It is the number of records that correctly predict a positive class. It indicates the presence of weed or pests class when it is not a crop3.True Negative (TN)It is the number of records that correctly predicts a negative class. It detects non-weed or no-pests class when it is crop actually4.False Positive (FP)It is the number of records when a weed or pest is incorrectly detected when it is a crop5.False Negative (FN)It is the number of records when the crop is incorrectly detected when it is a weed or pest6.Precision $\frac{TP}{TP + FP}$ 7.Recall $\frac{TP}{TP + FP}$ 8.Confusion matrixIt describes the performance of an algorithm in terms of the number of TP, TN, FP, ar	1.	Classification accuracy	It is the percentage of correctly identified crops among the input
3.True Negative (TN)It is the number of records that correctly predicts a negative class. It detects non-weed or no-pests class when it is crop actually4.False Positive (FP)It is the number of records when a weed or pest is incorrectly detected when it is a crop It is the number of records when the crop is incorrectly detected when it is a weed or p5.False Negative (FN)It is the number of records when the crop is incorrectly detected when it is a weed or p6.Precision $\frac{TP}{TP + FP}$ 7.Recall $\frac{TP}{TP + FP}$ 8.Confusion matrixIt describes the performance of an algorithm in terms of the number of TP, TN, FP, and TP	2.	True Positive (TP)	It is the number of records that correctly predict a positive class. It indicates the presence of weed or pests class when it is not a crop
4.False Positive (FP)It is the number of records when a weed or pest is incorrectly detected when it is a crophone.5.False Negative (FN)It is the number of records when the crop is incorrectly detected when it is a weed or performance.6.Precision $\frac{TP}{TP + FP}$ 7.Recall $\frac{TP}{TP + FP}$ 8.Confusion matrixIt describes the performance of an algorithm in terms of the number of TP, TN, FP, and TP.	3.	True Negative (TN)	It is the number of records that correctly predicts a negative class. It detects non-weed or no-pests class when it is crop actually
5.False Negative (FN)It is the number of records when the crop is incorrectly detected when it is a weed or p6.Precision $\frac{TP}{TP + FP}$ 7.Recall $\frac{TP}{TP + FP}$ 8.Confusion matrixIt describes the performance of an algorithm in terms of the number of TP, TN, FP, and TP, TN, FP, and TS	4.	False Positive (FP)	It is the number of records when a weed or pest is incorrectly detected when it is a crop
6.Precision $\frac{TP}{TP + FP}$ 7.Recall $\frac{TP}{TP + FP}$ 8.Confusion matrixIt describes the performance of an algorithm in terms of the number of TP, TN, FP, and TP,	5.	False Negative (FN)	It is the number of records when the crop is incorrectly detected when it is a weed or pest
7.Recall $\frac{TP}{TP + FP}$ 8.Confusion matrixIt describes the performance of an algorithm in terms of the number of TP, TN, FP, arEN	6.	Precision	$\frac{TP}{TP + FP}$
8. Confusion matrix It describes the performance of an algorithm in terms of the number of TP, TN, FP, and	7.	Recall	$\frac{TP}{TP + FP}$
FIN	8.	Confusion matrix	It describes the performance of an algorithm in terms of the number of TP, TN, FP, and FN
9. Mean pixel accuracy Percentage of a correctly classified pixel in the image	9.	Mean pixel accuracy	Percentage of a correctly classified pixel in the image
10. K-fold cross-validation k-fold is used to avoid overfitting the model. It splits the dataset into K parts and uses different parts iteratively to perform tests and train. The average accuracy would give the final performance of the weed detection algorithm	10.	K-fold cross-validation	k-fold is used to avoid overfitting the model. It splits the dataset into K parts and uses different parts iteratively to perform tests and train. The average accuracy would give the the final performance of the weed detection algorithm
11 ROC curve Depict performance of model graphically for classification. it is plotted between True	11	ROC curve	Depict performance of model graphically for classification. it is plotted between True
Positive Rate (TPR) and False-Positive Rate (FPR) with varied cutoff			Positive Rate (TPR) and False-Positive Rate (FPR) with varied cutoff
12. Kappa coefficient Evaluate the performance of the classifier based on a random classifier	12.	Kappa coefficient	Evaluate the performance of the classifier based on a random classifier

Discussion

The existing weed detection methods are specific to a particular crop instead of actual field images. When the existing approaches are applied to weed detection in a field, the accuracy and the performance become low. Weeds have a diverse range of species, a wide distribution, a diversity of leaf shapes and sizes, and irregular development, which results in a variety of textural aspects. Weeds in the bud stage are typically tiny, have a variety of appearances, and have a high germination density. As a result, obtaining accurate statistics is challenging. The following are the primary factors that influence weed detection performance:

- Weed detection is highly dependent on different stages 1. of development of a plant. In changing seasons or phases of growth and development, most plants vary their leaf morphology, texture, and spectral properties
- 2. Weed detection is highly dependent on the variation in the light received by the plant. The color of the vegetation is affected by the amount of sunlight and the shade of the plant canopy
- The presence of overlapping leaves, dead leaves, and 3. damaged leaves causes problems in the image segmentation task

At present, machine learning and deep learning methods are widely used in various computer vision tasks based on image analysis i.e., food safety, plant diseases, and pests detection, and play an important role in agriculture as proposed by Khan et al. (2021), Khan (2022). Khan et al. (2022).

However, the detection of agricultural pests from images is a hard problem because of small sample size availability, unlabeled dataset, noisy data, and lack of sufficient benchmark datasets like image net classifiers process around 14 million datasets which are needed improved and cleaned dataset for better accuracy and to reduce the spurious prediction. Apart from that, the incidence rate of some pests causing plant disease is very low and the image acquisition cost is also too high leading to a lesser number of useful data set availability impacting the effective and precise classification of plant diseases and pests identification through various deep learning models.

Conclusion and Future Work

The research work carried out provides a comprehensive survey in the domain of weed and pests detection in the crops. The work presents the role of artificial intelligence and machine learning algorithms to automate the detection of weeds and pests among plants. Various publicly available datasets are also presented. High levels of automatic weed and pests detection have been attained using multiple platforms and mechanical equipment based on machine learning methods and deep learning-based approaches. However, the domain of weed and pests detection is still in its early stage. The presence of current machine learning and deep learning algorithms will act as a foundation for achieving higher accuracy of results. The work requires focusing on the generalization of results and robustness. Apart from that the real field conditions of the crops and weed growth are controlled by several other factors i.e., biotic to abiotic, which poses another line of the major challenge to image-based data collection with more preciseness. Though sufficient research efforts were made in past to cope with accurate data collection still it is required to be addressed under the

advanced UAV-based image collection technology for precision agriculture. Based on the challenges discussed, future work can be focused on overlapping leaves, optimization of weed detection in the fields, and increasing the pests samples for training data.

Acknowledgment

We are thankful to the Jaypee Institute of Information Technology for facilitating the necessary support.

Author's Contributions

Nidhi Gupta: Research plan and organized the study, analysis, and written manuscript.

Bharat Gupta: Research plan, analysis, and organized the study and written manuscript.

Kalpdrum Passi: Research plans direction in the study and critical suggestions and contributed to the writing of the manuscript.

Chakresh Kumar Jain: All experiments, coordinated, performed data analysis, and contributed to the writing of the manuscript.

Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript and no ethical issues involved.

References

- Abiodun, E. O., Alabdulatif, A., Abiodun, O. I., Alawida, M., Alabdulatif, A., & Alkhawaldeh, R. S. (2021). A systematic review of emerging feature selection optimization methods for optimal text classification: The present state and prospective opportunities. Neural Computing and Applications, 33(22), 15091-15118. doi.org/10.1007/s00521-021-06406-8
- Aitkenhead, M., Dalgetty, I., Mullins, C., McDonald, A., & Strachan, N. (2003). Weed and crop discrimination using image analysis and artificial intelligence methods. Computers and Electronics in Agriculture, 39(3), 157-171.

doi.org/10.1016/s0168-1699(03)00076-0

- Al-Saqer. (2011). Red palm weevil (Rynchophorus ferrugineous, Olivier) recognition by image processing techniques. American Journal of Agricultural and Biological Sciences, 6(3), 365-376. doi.org/10.3844/ajabssp.2011.365.376
- Aravind, R., Daman, M., & Kariyappa, B. S. (2015, December). Design and development of automatic weed detection and smart herbicide sprayer robot. In 2015 IEEE Recent Advances in Intelligent

Computational Systems (RAICS) (pp. 257-261). IEEE. doi.org/10.1109/RAICS.2015.7488424

Asefpour, V. K., & Massah, J. (2013). Performance evaluation of a machine vision system for insect pests identification of field crops using artificial neural networks. Archives of phytopathology and plant protection, 46(11), 1262-1269.

doi.org/10.1080/03235408.2013.763620

- Badhan, S., Desai, K., Dsilva, M., Sonkusare R. & Weakey, S. (2021). Real-Time Weed Detection using Machine Learning and Stereo-Vision. 6th International Conference for Convergence in Technology (I2CT), pp. 1-5, doi.org/10.1109/I2CT51068.2021.9417989.
- Bakhshipour, A., Jafari, A., Nassiri, S., & Zare, D. (2017). Weed segmentation using texture features extracted from wavelet sub-images. Biosystems Engineering, 157, 1-12.

doi.org/10.1016/j.biosystemseng.2017.02.002

- Behmann, J., Mahlein, A., Rumpf, T., Römer, C., & Plümer, L. (2014). A review of advanced machine learning methods for the detection of biotic stress in precision crop protection. Precision Agriculture, 16(3), 239-260. doi.org/10.1007/s11119-014-9372-7
- Ben Ayed, R., & Hanana, M. (2021). Artificial intelligence to improve the food and agriculture sector. Journal of Food Quality, 2021. doi.org/10.1155/2021/5584754
- Bosilj, P., Aptoula, E., Duckett, T., & Cielniak, G. (2019). Transfer learning between crop types for semantic segmentation of crops versus weeds in precision agriculture. Journal of Field Robotics, 37(1), 7-19. doi.org/10.1002/rob.21869
- Boulent, J., Foucher, S., Théau, J., & St-Charles, P. (2019). Convolutional neural networks for the automatic identification of plant diseases. Frontiers in Plant Science, 10. doi.org/10.3389/fpls.2019.00941
- Chebrolu, N., Lottes, P., Schaefer, A., Winterhalter, W., Burgard, W., & Stachniss, C. (2017). Agricultural robot dataset for the plant classification, localization and mapping of sugar beet fields. The International Journal of Robotics Research, 36(10), 1045-1052. doi.org/10.1177/0278364917720510
- Chen, Y., Lin, P., He, Y., & Xu, Z. (2011). Classification of broadleaf weed images using Gabor wavelets and Lie group structure of region covariance on Riemannian manifolds. Biosystems Engineering, 109(3), 220-227. doi.org/10.1016/j.biosystemseng.2011.04.003
- De Rainville, F., Durand, A., Fortin, F., Tanguy, K., Maldague, X., Panneton, B., & Simard, M. (2012). Bayesian classification and unsupervised learning for

isolating weeds in row crops. Pattern Analysis and Applications, 17(2), 401-414.

doi.org/10.1007/s10044-012-0307-5

- dos Santos Ferreira, A., Freitas, D. M., da Silva, G. G., Pistori, H., & Folhes, M. T. (2017). Weed detection in soybean crops using ConvNets. Computers and Electronics in Agriculture, 143, 314-324. doi.org/10.1016/j.compag.2017.10.027
- Ebrahimi, M. A., Khoshtaghaza, M. H., Minaei, S., & Jamshidi, B. (2017). Vision-based pest detection is based on the SVM classification method. Computers and Electronics in Agriculture, 137, 52-58. doi.org/10.1016/j.compag.2017.03.016
- Fina, F., Birch, P., Young, R., Obu, J., Faithpraise, B., & Chatwin, C. (2013). Automatic plant pest detection and recognition using k-means clustering algorithm and correspondence filters. International Journal of Advanced Biotechnology and Research, 4(2), 189-199.
- Freund, Y., & Schapire, R. E. (1997). A decision-theoretic generalization of online learning and an application to boosting. Journal of computer and system sciences, 55(1), 119-139. doi.org/10.1006/jcss.1997.1504
- Friedman, J., Hastie, T., & Tibshirani, R. (2000). Additive logistic regression: A statistical view of boosting (with discussion and a rejoinder by the authors). The annals of statistics, 28(2), 337-407. doi.org/10.1214/aos/1016218223
- Huang, M. L., & Chuang, T. C. (2020). A database of eight common tomato pest images. Mendeley data. DOI:10.17632/s62zm6djd2.1
- Hughes, D., & Salathé, M. (2015). An open access repository of images on plant health to enable the development of mobile disease diagnostics. arXiv preprint arXiv:1511.08060. https://arxiv.org/abs/1511.08060
- Ishak, A. J., Hussain, A., & Mustafa, M. M. (2009). Weed image classification using Gabor wavelet and gradient field distribution. Computers and Electronics in Agriculture, 66(1), 53-61. doi.org/10.1016/j.compag.2008.12.003
- Islam, N., Rashid, M. M., Wibowo, S., Xu, C. Y., Morshed, A., Wasimi, S. A., & Rahman, S. M. (2021). Early weed detection using image processing and machine learning techniques in an Australian chili farm. Agriculture, 11(5), 387. doi.org/10.3390/agriculture11050387
- Jiang, H., Zhang, C., Qiao, Y., Zhang, Z., Zhang, W., & Song, C. (2020). CNN feature-based graph convolutional network for weed and crop recognition in smart farming. Computers and Electronics in Agriculture, 174, 105450.

doi.org/10.1016/j.compag.2020.105450

- Karger, & Shirzadifar. (2013) Automatic weed detection system and smart herbicide sprayer robot for corn fields. doi.org/10.14419/ijet.v7i3.6.14952
- Kazmi, W., Garcia-Ruiz, F., Nielsen, J., Rasmussen, J., & Andersen, H. (2015). Exploiting affine invariant regions and leaf edge shapes for weed detection. Computers and electronics in agriculture, 118, 290-299. doi.org/10.1016/j.compag.2015.08.023
- Khan, R. (2022). Artificial Intelligence and Machine Learning in Food Industries: A Study. J Food Chem Nanotechnol, 7(3), 60-67. doi.org/10.17756/jfcn.2021-114
- Khan, R., Dhingra, N., & Bhati, N. (2022). Role of Artificial Intelligence in Agriculture: A Comparative Study. In Transforming Management with AI, Big-Data and IoT (pp. 73-83). Springer, Cham. https://link.springer.com/chapter/10.1007/978-3-030-86749-2_4
- Khan, R., Kumar, S., Dhingra, N., & Bhati, N. (2021). The Use of Different image recognition techniques in food safety: A study. Journal of Food Quality, 2021. doi.org/10.1155/2021/7223164
- Lameski, P., Zdravevski, E., Trajkovik, V., & Kulakov, A. (2017). Weed detection dataset with RGB images taken under variable light conditions. ICT Innovations 2017, 112-119. doi.org/10.1007/978-3-319-67597-8_11
- Lee, W. S., Slaughter, D. C. & Giles, D. K. (1999). Robotic Weed Control System for Tomatoes. Precision Agriculture 1, 95-113. doi.org/10.1023/A:1009977903204
- Leow, L. K., Chew, L. L., Chong, V. C., & Dhillon, S. K. (2015). Automated identification of copepods using digital image processing and artificial neural network. BMC bioinformatics, 16(18), 1-12. doi.org/10.1186/1471-2105-16-s18-s4
- Lin, F., Zhang, D., Huang, Y., Wang, X., & Chen, X. (2017). Detection of corn and weed species by the combination of spectral, shape and textural features. Sustainability, 9(8), 1335. doi.org/10.3390/su9081335
- Lottes, P., Hörferlin, M., Sander, S., & Stachniss, C. (2017). Effective vision-based classification for separating sugar beets and weeds for precision farming. Journal of field Robotics, 34(6), 1160-1178. doi.org/10.1002/rob.21675
- Martineau, M., Conte, D., Raveaux, R., Arnault, I., Munier, D., & Venturini, G. (2017). A survey on image-based insect classification. Pattern recognition, 65, 273-284. doi.org/10.1016/j.patcog.2016.12.020
- Mayo, M., & Watson, A. T. (2007). Automatic species identification of live moths. Knowledge-based systems, 20(2), 195-202. doi.org/10.1016/j.knosys.2006.11.012
- McCarthy, C., Rees, S., & Baillie, C. (2010). Machine vision-based weed spot spraying: A review and where

next for sugarcane? In Proceedings of the 32nd Annual Conference of the Australian Society of Sugar Cane Technologists (ASSCT 2010) (Vol. 32, pp. 424-432). Australian Society of Sugar Cane Technologists.

- Misaghi, M., & Yaghoobi, M. (2019). Improved invasive Weed Optimization algorithm (IWO) based on chaos theory for the optimal design of PID controller. Journal of Computational Design and Engineering, 6(3), 284-295. doi.org/10.1016/j.jcde.2019.01.001
- Narassiguin, A., Bibimoune, M., Elghazel, H., & Aussem, A. (2016). An extensive empirical comparison of ensemble learning methods for binary classification. Pattern Analysis and Applications, 19(4), 1093-1128. doi.org/10.1007/s10044-016-0553-z
- Nursuriati, J., Hussin, N., Nordin, S., Awang, K. (2015). Automatic Plant Identification: Is Shape the Key Feature? Procedia Comput. Sci. 2015, 76, 436–442. doi.org/10.1016/j.procs.2015.12.287
- Panqueba, B. N. S., & Medina, C. A. C. (2016). A computer vision application to detect unwanted weeds in early-stage crops. WSEAS transactions on computer research, 4, 41-45. http://www.wseas.us/journal/pdf/computerresearc

h/2016/a105818-093.pdf

Partel, V., Kakarla, S. C., & Ampatzidis, Y. (2019). Development and evaluation of low-cost and smart technology for precision weed management utilizing artificial intelligence. Computers and electronics in agriculture, 157, 339-350.

doi.org/10.1016/j.compag.2018.12.048

- Prajapati, H., Shah, J., & Dabhi, V. (2018). Detection and classification of rice plant diseases. Intelligent decision technologies, 1-18. doi.org/10.3233/idt-180338
- Rakhmatuiln, I., Kamilaris, A., & Andreasen, C. (2021).
 Deep Neural Networks to Detect Weeds from Crops in Agricultural Environments in Real-Time: A Review. Remote sensing, 13(21), 4486. doi.org/10.1007/s13280-016-0793-6
- Roldán-Serrato, K. L., Escalante-Estrada, J., & Rodríguez-González, M. (2018). Automatic pest detection on bean and potato crops by applying neural classifiers. Engineering in agriculture, environment and food, 11(4), 245-255. doi.org/10.1016/j.eaef.2018.08.003
- Rockström, J., Williams, J., Daily, G., Noble, A., Matthews, N., Gordon, L., & Smith, J. (2017). Sustainable intensification of agriculture for human prosperity and global sustainability. Ambio, 46(1), 4-17. doi.org/10.1007/s13280-016-0793-6
- Saleem, M. H., Velayudhan, K. K., Potgieter, J., & Arif, K. (2022). Weed Identification by Single-Stage and Two-Stage Neural Networks: A Study on the Impact of image resizers and weights optimization algorithms. doi.org/10.3389/fpls.2022.850666
- Shah, J. P., Prajapati, H. B., & Dabhi, V. K. (2016, March). A survey on detection and classification of

rice plant diseases. In 2016 IEEE International Conference on Current Trends in Advanced Computing (ICCTAC) (pp. 1-8). IEEE. doi.org/10.1109/ICCTAC.2016.7567333

kovsen, S., Dyrmann, M., Mortensen, A. K., Laursen, M. S., Gislum, R., Eriksen, J., & Jorgensen, R. N. (2019). The grass clover image dataset for semantic and hierarchical species understanding in agriculture. In Proceedings of the IEEE/CVF Conference on Computer Vision and

Recognition

Pattern

doi.org/10.1109/cvprw.2019.00325
Subeesh, A., Bhole, S., Singh, K., Chandel, N. S., Rajwade, Y. A., Rao, K. V. R., & Jat, D. (2022). Deep convolutional neural network models for weed detection in polyhouse-grown bell peppers. Artificial Intelligence in Agriculture, 6, 47-54. doi.org/10.1016/j.aiia.2022.01.002

Workshops

0-0).

(pp.

- Sudars, K., Jasko, J., Namatevs, I., Ozola, L., & Badaukis, N. (2020). Dataset of annotated food crops and weed images for robotic computer vision control. Data in brief, 31, 105833. doi.org/10.1016/j.dib.2020.105833
- Tellaeche, A., Pajares, G., Burgos-Artizzu, X., & Ribeiro, A. (2011). A computer vision approach for weeds identification through Support Vector Machines. Applied Soft Computing, 11(1), 908-915. doi.org/10.1016/j.asoc.2010.01.011
- Thapa, R., Zhang, K., Snavely, N., Belongie, S., & Khan, A. (2020). The Plant Pathology Challenge 2020 data set to classify foliar disease of apples. Applications in plant sciences, 8(9), e11390. doi.org/10.1002/aps3.11390
- Torres-Sánchez, J., López-Granados, F., & Pena, J. M. (2015). An automatic object-based method for optimal thresholding in UAV images: Application for vegetation detection in herbaceous crops. Computers and Electronics in Agriculture, 114, 43-52. doi.org/10.1016/j.compag.2015.03.019
- Vanegas, F., Bratanov, D., Powell, K., Weiss, J., & Gonzalez, F. (2018). A novel methodology for improving plant pest surveillance in vineyards and crops using UAV-based hyperspectral and spatial data. Sensors, 18(1), 260. doi.org/10.3390/s18010260
- Wang, K., Zhang, S., Wang, Z., Liu, Z., & Yang, F. (2013). Mobile smart device-based vegetable disease and insect pest recognition method. Intelligent Automation & Soft Computing, 19(3), 263-273. doi.org/10.1080/10798587.2013.823783
- Wang, R., Liu, L., Xie, C., Yang, P., Li, R., & Zhou, M. (2021). Agri Pest: A large-scale domain-specific benchmark dataset for practical agricultural pest detection in the Wild. Sensors (Basel, Switzerland), 21(5), 1601. doi.org/10.3390/s21051601
- Wang, Z., Wang, K., Liu, Z., Wang, X., & Pan, S. (2018). A cognitive vision method for insect pest image segmentation. IFAC-Papers on Line, 51(17), 85-89. doi.org/10.1016/j.ifacol.2018.08.066

- Wu, X., Zhan, C., Lai, Y. K., Cheng, M. M., & Yang, J. (2019). Ip102: A large-scale benchmark dataset for insect pest recognition. In proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 8787-8796). doi.org/10.1109/CVPR.2019.00899.
- Wu, Z., Chen, Y., Zhao, B., Kang, X., & Ding, Y. (2021). Review of weed detection methods based on computer vision. Sensors, 21(11), 3647. doi.org/10.3390/ s21113647
- Xie, C., Zhang, J., Li, R., Li, J., Hong, P., Xia, J., & Chen, P. (2015). Automatic classification for field crop insects via multiple-task sparse representation and multiplekernel learning. Computers and Electronics in Agriculture, 119, 123-132. doi.org/10.1016/j.compag.2015.10.015
- Yao, Q. (2016). Automatic detection and counting of rice planthoppers in Paddy fields by image processing methods. 2016 International Congress of Entomology.

doi.org/10.1603/ice.2016.110884

Zhu, L., & Zhang, Z. (2011). Insect recognition is based on integrated region matching and dual-tree complex wavelet transform. Journal of Zhejiang University science C, 12(1), 44-53. doi.org/10.1631/jzus.c0910740