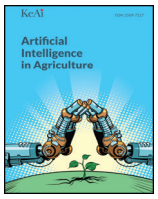
Artificial Intelligence in Agriculture 10 (2023) 13–25

Contents lists available at ScienceDirect

Artificial Intelligence in Agriculture

journal homepage: http://www.keaipublishing.com/en/journals/artificial-intelligence-in-agriculture/



Machine learning for weed–plant discrimination in agriculture 5.0: An in-depth review

Filbert H. Juwono [a,*], W.K. Wong [b], Seema Verma [c], Neha Shekhawat [c], Basil Andy Lease [b], Catur Apriono [d,*]

[a] Department of Electrical and Electronic Engineering, Xi'an Jiaotong-Liverpool University, Suzhou 215123, China
[b] Dpartment of Electrical and Computer Engineering, Curtin University Malaysia, Miri 98009, Malaysia
[c] School of Physical Sciences, Banasthali Vidyapith, Rajasthan 304022, India
[d] Department of Electrical Engineering, Universitas Indonesia, Depok 16424, Indonesia

ARTICLE INFO

Article history:
Received 12 March 2023
Received in revised form 25 July 2023
Accepted 11 September 2023
Available online 19 September 2023

Keywords:
Agriculture 5.0
Machine learning
Unmanned aerial vehicle
Weed-plant discrimination

ABSTRACT

Agriculture 5.0 is an emerging concept where sensors, big data, Internet-of-Things (IoT), robots, and Artificial Intelligence (AI) are used for agricultural purposes. Different from Agriculture 4.0, robots and AI become the focus of the implementation in Agriculture 5.0. One of the applications of Agriculture 5.0 is weed management where robots are used to discriminate weeds from the crops or plants so that proper action can be performed to remove the weeds. This paper discusses an in-depth review of Machine Learning (ML) techniques used for discriminating weeds from crops or plants. We specifically present a detailed explanation of five steps required in using ML algorithms to distinguish between weeds and plants.

© 2023 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co., Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Contents

1. Introduction . . . 14
   1.1. Automated weed control: A technological progress . . . 14
   1.2. Contributions and structure . . . 15
2. Data acquisition . . . 15
3. Image pre-processing: Background removal . . . 15
   3.1. Excess Green Index (ExGI) . . . 16
   3.2. Ratio Vegetation Index (RVI) and Vegetation Index Number (VIN) . . . 16
   3.3. Normalized Difference Vegetation Index (NDVI) . . . 16
      3.3.1. Otsu threshold . . . 17
      3.3.2. Modified Otsu threshold . . . 17
   3.4. Transformed Vegetation Index (TVI), corrected TVI (CTVI), and Thiam's TVI (TTVI) . . . 17
   3.5. Perpendicular Vegetation Index (PVI) . . . 17
4. Feature-based recognition . . . 18
   4.1. Biological morphology features . . . 18
   4.2. Spectral features . . . 18
   4.3. Texture features . . . 19
5. Classification algorithms . . . 20
   5.1. ML-based algorithms . . . 20
   5.2. Ensemble learning . . . 20
   5.3. DL-based algorithms . . . 21
6. Publicly available datasets . . . 22
7. Challenges and opportunities . . . 23

* Corresponding authors.
E-mail addresses: Filbert.Juwono@xjtlu.edu.cn (F.H. Juwono), weikitt.w@curtin.edu.my (W.K. Wong), seemaverma3@yahoo.com (S. Verma), shekhawatneha37@gmail.com (N. Shekhawat), basil.lease@curtin.edu.my (B.A. Lease), catur@eng.ui.ac.id (C. Apriono).

https://doi.org/10.1016/j.aiia.2023.09.002
2589-7217/© 2023 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co., Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

8. Conclusion . . . 23
Declaration of Competing Interest . . . 23
Acknowledgements . . . 23
References . . . 23

1. Introduction

Recently, Information and Communication Technology (ICT) has been immersed in agricultural practises. Agriculture 4.0, which is also known as data-driven agriculture or smart farming or digital farming, uses telematics, data management, and precision agriculture to increase the quality of the crops (Saiz-Rubio and Rovira-Más, 2020). The National Research Council (1997) defines precision agriculture as *the application of modern information technologies to provide, process and analyze multisource data of high spatial and temporal resolution for decision making and operations in the management of crop production*. It is clear that precision agriculture focuses on applying inputs (e.g., fertilizers, pesticides) when necessary in the specific quantities. This will optimize the resources and make the plants receive what they really need. The researchers in this area focus on creating systems that can make operational and strategical decisions using the data supplied. Traditionally, farmers go around the farms to check the conditions of the crops and make decisions based on the acquired experience. This conventional method is no more feasible as the size of fields has grown immensely and traditional methods demand more working resources.

Agriculture 5.0 has evolved as ICT has advanced. Different from Agriculture 4.0, Agriculture 5.0 employs precision agriculture and unmanned and autonomous equipment (Saiz-Rubio and Rovira-Más, 2020). In other words, robots and Artificial Intelligence (AI) play important roles in this field. Technologically, computer vision, Machine Learning (ML), and Unmanned Aerial Vehicles (UAVs), or commonly known as drones, have been utilized to build Agriculture 5.0 system.

Moreover, computer vision has recently played a vital role in enabling the digital world to interact with the physical world. In general, the technology exploits cameras coupled with computer, rather than the human eyes, to identify, track, and estimate targets for image processing. It plays a critical role in technology by powering object detection, classification, and tracking of objects in all domains. Recently, smart agriculture tasks (Neupane and Baysal-Gurel, 2021) that include plant ailment detection, weed detection, crop yield prediction, identification of plant species, are achieved through computer vision technology.

UAVs are recently taking up considerably space in research domain so as to make the industry less labor-intensive. Applications of computer vision to drone technology makes it be able to interpret and interact with surroundings, including buildings, trees, and diverse terrain. The use of data from UAVs for farming applications has been gaining an enormous popularity in both research and industry. Compared to human-based monitoring, UAV is considered to be cheaper and more convenient with shorter inspection time needed. The optimal control of the UAV in the presence of obstacles is commendable, which is a critical factor while working with two crop rows.

The initial step in Agriculture 5.0 is data collection, followed by autonomous decision making and its implementation. The data collected from physical sources such as sensors and cameras contains valuable information, which is directly relevant to the process of decision making. The connection between the data and the decision stage involves filtering routines and AI algorithms for getting only the right data and helping the farmer make correct decisions. The last step of physical execution of the decision, also referred as actuation, is carried out by advanced equipment coupled with computer control units.

The exponential growth of robotic applications in smart farming shows its efficiency in handling workforce and cost of production. The advanced level of precision agriculture deals with applying the above principles using unmanned operated equipment and systems that involve autonomous decisions. The use of autonomous and semi-autonomous robots in the field of agriculture has immensely increased in last decades. They possess arms that can detect weeds and can spray chemicals on the affected areas specifically, thus reducing the overall cost.

1.1. Automated weed control: A technological progress

In general, weeds hinder the movement of irrigation water, disrupt the process of applying pesticides, and serve as a breeding ground for pathogens. Thus, weed detection and control play a vital role in overall crop productivity and farming expenses. Effective weed control plays a crucial role in agriculture by minimizing crop yields, dramatically escalating production costs, negatively impacting crops, and significantly compromising product quality. Note that weeds compete with productive crops for essential resources such as water, nutrients, and light, thereby diminishing overall agricultural output.

Traditional weed management methods include chemical means i.e., uniform application of herbicides throughout the field using mechanical means or physical elimination. The former method leads to the overuse of chemicals as the weed spatial density varies across the field which ends up in environmental apprehension and formation of herbicide-resistant weeds. The latter results in labor intensive farming methods in turn causing a steep increase in production cost. The above issues can be solved by applying concept of Site-Specific Weed Management (SSWM) (Lati et al., 2021). SSWM involves detection of patches of weeds across the field and carrying out spot spraying or mechanical removal ways which can be broadly classified into two categories: 1) prescription maps-based methods (De Castro et al., 2018) where the areas of weed emergence are detected to focus the application of weedicide, and 2) real time monitoring method (Xu et al., 2018) that simultaneously detects and controls weeds by spraying weedicide on the spot.

Existing techniques for image processing provide extremely promising outcomes under optimal imaging conditions, even though they can be quite challenging under real time conditions. The overlap of plant leaves and weeds at different stages of growth makes them synonymous to each other. In addition, the leaves are often occluded, discolored or damaged by undesired materials which alter the morphological as well as spectral characters of the leaves. As the lighting conditions change during different times in a day, the shadows formed by the plant canopy and the solar inclination can directly impact the colour of vegetation. Furthermore, different plant growth stages also result in variations of the morphological, textural, and spectral attributes to leaves.

This paper focuses on the discrimination between crop and weed using ML approaches. Note that traditional image processing techniques like Hough transform can also be used to discriminate crop and weed (Bah et al., 2017). However, these approaches are not of our interest. Instead, we commonly use image processing techniques to extract the features like color, shape, etc. to be fed to the ML algorithms. The efficiency of these methods mostly depends on the manually designed features and have high dependence on image acquisition methods, preprocessing methods, and the standard of features extracted. Earlier studies in this category primarily used color co-occurrence matrix-based texture analysis for images (Chang et al., 2012). With

advancement in technology, optical sensor coupled with image processing algorithms has been implemented for variable rate weedicide applications. The major challenge in classification of weeds from crops is caused by their similar spectral signature.

Image processing-based weed control is usually performed through several steps. After the images are required, they may not be suitable for further processing. For example, it is required to discriminate the vegetation from the soil. The steps will be further discussed in this paper. Note that this paper can be seen as a complementary to a few related works, such as (Wang et al., 2019; Hasan et al., 2021).

1.2. Contributions and structure

As previously discussed, discriminating weeds from plants is necessary in weed management, regardless the method that is chosen, i.e., using chemical or non-chemical. In this paper, we focus on the discussion of weed control applications, their recent advances, and also challenges. The steps to do weed-crop discrimination can be simply explained as follows. Initially, the images are captured using robots or UAVs equipped with cameras. The images must be free from any background interference (e.g., soil). Trained experts can be employed to label and distinguish between the weeds and the crops in the images. If any challenges arise during image acquisition, publicly available datasets with annotations can be utilized as an alternative. Afterward, the classification task begins by selecting the features, which can be performed either manually or automatically. ML algorithms use the selected features for learning the patterns, while Deep Learning (DL), which is a subset of ML, can be utilized to perform both automatic feature selection and classification. This concept is depicted in Fig. 1.

The review of the state-of-the-art technology for weed-crop discrimination will be discussed in the following structure. The paper starts by examining the motivation behind weed-crop recognition task in Section 1. Moreover, various data acquisition methods and preprocessing strategies aimed at distinguishing between vegetation and the background (background removal) are presented in Section 2 and Section 3, respectively. Subsequently, feature-based weed-crop recognition techniques are presented in Section 4. Next, ML and DL approaches for classification tasks are elaborated in Section 5. The publicly available datasets are discussed briefly in Section 6. The challenges and opportunities can be found in Section 7. Finally, the paper concludes the discussion in Section 8.

2. Data acquisition

Mobile robots (Xu et al., 2019) or UAVs (Tripicchio et al., 2015) are preferred to be used to capture images of the plantation. The robots and UAVs should be equipped with either multispectral or hyperspectral cameras. A multispectral camera can capture several bands, usually R (red), G (green), B (blue), and NIR (Near-Infrared). In Sankaran et al. (2013), the authors used two types of cameras, each camera had three different bands. The first camera had G, B, and NIR bands while the second one had R, G, and NIR bands.

On the other hand, a hyperspectral sensor may provide hundreds or thousands of bands. A MicroHyperspec sensor which captures images using 325 bands in the visible band and NIR band was used in Lu et al. (2019). In the paper, it is also said that hyperspectral images were in limited supply. In terms of the vegetation properties estimation, the authors concluded that hyperspectral images produced a comparable performance compared with the multispectral images.

Hyperspectral sensors can be classified into point-scan, area-scan, and line-scan categories based on the scanning method they use to obtain 3D-hyperspectral cubes (x, y, λ) (Patrício and Rieder, 2018; Qin et al., 2013). Point-scanning cameras are designed to capture a singular point along two-dimensional area. Spectrum of a single pixel is obtained using this method. An area-scanning camera captures a 2-D grayscale image (x, y) of a single band. Finally, a line-scanning camera, which can be considered as the extension of the point-scanning type, generates a 2-D image (y, λ). The difference can be seen in Fig. 2.

The quality of captured images is influenced by various environmental factors, including lighting conditions (such as day time, night time, and shadows) and humidity levels (wet vs dry crops) (Jakubczyk et al., 2023). In Manea and Calin (2015), it has been clearly proved that illumination conditions plays an important role for determining the quality of the images. High light intensity leads to a loss of information, while low light intensity introduces dark current noise. Furthermore, the quality of images captured using line-scan method is affected by the attitude and position changes of the UAV (Xue et al., 2021).

3. Image pre-processing: Background removal

Image enhancement is required to improve image quality so that it is suitable for feature extraction and classification tasks. The manually taken images sometimes contain disturbances, such as shadows, stones, water, and soil (Gée et al., 2008; Lease et al., 2020). The images are typically multispectral images which may contain the following four channels: Red - R (630–690 nm), Green - G (510–580 nm), Blue - B (450–510 nm), Mid-Infrared - MIR (300–500 nm), and Near-Infrared - NIR (700–1100 nm). Vegetation Indices (VIs) are usually used to separate the vegetation from the disturbances. Note that each component has different reflectance to a specific band. For example, soil may have low reflectance to the red spectrum. Using this principle, we can differentiate vegetation from the disturbances.

According to the bands, VIs can be categorized into VIs without infrared channels (RGB) and with infrared channels (RGB + NIR/MIR). According to the method of obtaining the mathematical formula, VIs can be categorized into first and second generation (Bannari et al., 1995). First generation VIs are obtained using empirical method without

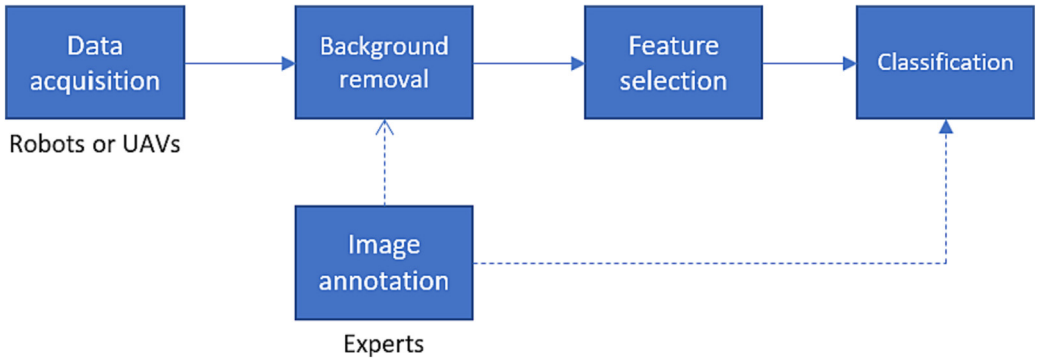


Fig. 1. Concept of weed detection.

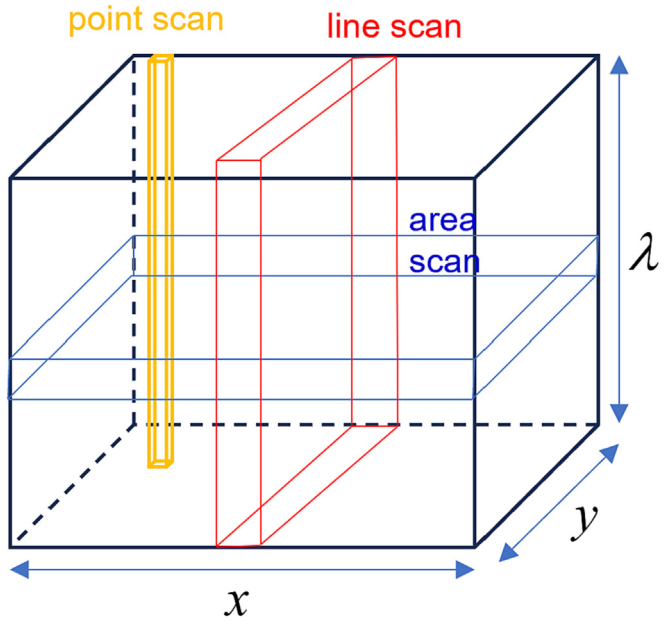


Fig. 2. Hyperpectral sensor scanning methods.

considering atmospheric effects, soil brightness, and soil color. Meanwhile, second generation VIs are obtained through mathematical and physical reasoning, logical experiment, and simulation. Furthermore, according to the method, VIs can be grouped into (Mróz and Sobieraj, 2004):

1. Slope-based: This method uses the ratio of NIR and R. This method includes Ratio Vegetation Index (RVI), Vegetation Index Number (VIN), and Normalized Difference Vegetation Index (NDVI). They may have some variants that have been proposed by various researchers.
2. Distance-based: This method uses soil line as the reference and then measures the perpendicular distance of each pixel point to this soil line. This method includes Perpendicular Vegetation Index 1 and 2 (PVI_1 and PVI_2, respectively), Difference Vegetation Index (DVI), and Soil Adjusted Vegetation Index (SAVI) and its variants.
3. Orthogonal transformation: This group uses orthogonal transformation to create new uncorrelated bands.
4. Red Edge Inflection Point (REIP): This method observes the red edge, which is a fast shift in reflectance found in green plant spectrum at the transition between visible and near-infrared wavelengths, using Gaussian, polynomial, and Lagrangian models.

We notice from the above discussion, there are many well-known VIs in practice. Thorough review papers of VIs can be found in Xue and Su (2017), Bannari et al. (1995), and Mróz and Sobieraj (2004). Here, we revisit some VIs as shown in Fig. 3.

3.1. Excess Green Index (ExGI)

The *ExGI* is defined by Gée et al. (2008) and Woebbecke et al. (1995)

$$ExGI = 2g - r - b, \tag{1}$$

where

$$g = \frac{G_n}{R_n + G_n + B_n}, \tag{2}$$

$$r = \frac{R_n}{R_n + G_n + B_n}, \tag{3}$$

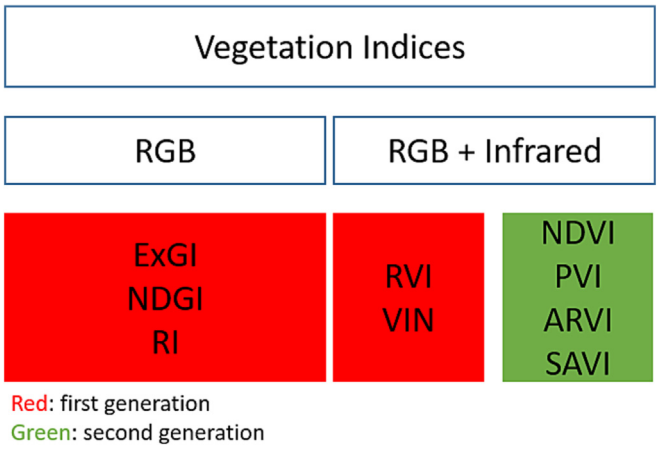


Fig. 3. VIs discussed in this paper.

$$b = \frac{B_n}{R_n + G_n + B_n}, \tag{4}$$

R_n, G_n, and B_n are the normalized RGB coordinates ranging between 0 and 1. The normalized coordinates are given by

$$G_n = \frac{G}{G_m}, \tag{5}$$

$$R_n = \frac{R}{R_m}, \tag{6}$$

$$B_n = \frac{B}{B_m}, \tag{7}$$

where R, G, and B are the true coordinates and R_m, G_m, and B_m are the maximum true coordinates. Note that $R_m = G_m = B_m = 255$ for 24-bit color images. Fig. 4 shows the images before and after ExGI preprocessing.

3.2. Ratio Vegetation Index (RVI) and Vegetation Index Number (VIN)

RVI and VIN are the two first VIs (Bannari et al., 1995). They use the reflectance of the red, *R* and near-infrared, *NIR*. The RVI is mathematically given by

$$RVI = \frac{R}{NIR}, \tag{8}$$

while *VIN* is expressed by

$$VIN = \frac{1}{RVI} = \frac{NIR}{R}. \tag{9}$$

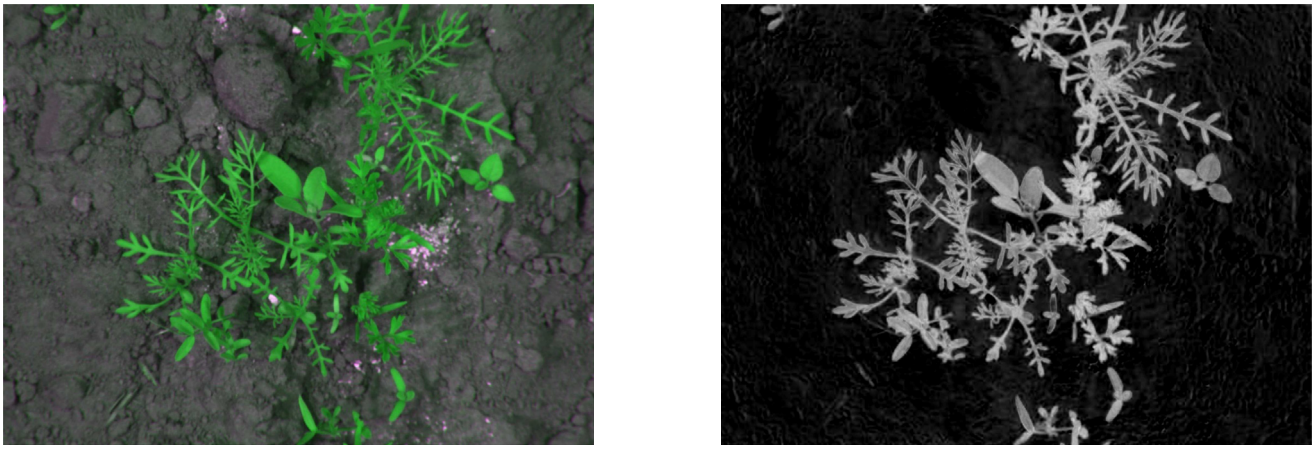
3.3. Normalized Difference Vegetation Index (NDVI)

NDVI is the most popular VI used in the literature (Bosilj et al., 2018). NDVI measures the spectral reflectance difference between near-infrared and red spectrum. Mathematically, NDVI is expressed by

$$NVDI = \frac{NIR - R}{NIR + R}, \tag{10}$$

or using the *VIN* term

$$NVDI = \frac{VIN - 1}{VIN + 1}, \tag{11}$$



(a) Original image (b) ExGI image

Fig. 4. Before and after ExGI Preprocessing.

where *NIR* is the near-infrared spectrum reflectance and *R* is the red spectrum reflectance. It is worth mentioning that NIR can be replaced by mid-infrared (MIR) (Bannari et al., 1995). This process indirectly converts the image into grey scale with most, but not all, background pixels removed. Subsequently, the image is converted into black and white as a binary mask using Otsu threshold (Otsu, 1979), that will be described briefly below. The binary mask is then applied to the gray-scale image to remove the background.

3.3.1. Otsu threshold

Let us assume that a grayscale image has L pixel levels. We would like to separate the image into two classes, background (class C_0) and object (class C_1). Class C_0 is a group of pixels with level $[1,\cdots,T]$ and C_1 is a group of pixels with level $[T+1,\cdots,L]$, where T is the threshold. The probabilities for class occurrence are given by

$$\omega_0 = \omega(T) = Pr(C_0) = \sum_{i=1}^{T} p_i, \tag{12}$$

and

$$\omega_1 = 1 - \omega(T) = Pr(C_1) = \sum_{i=T+1}^{L} p_i. \tag{13}$$

The class mean levels are given by

$$\mu_0 = \sum_{i=1}^{T} ip_i/\omega_0 = \mu(T)/\omega(T), \tag{14}$$

and

$$\mu_1 = \sum_{i=T+1}^{L} ip_i/\omega_1 = (\mu_T - \mu(T))/(1 - \omega(T)), \tag{15}$$

where $\mu(T) = \sum_{i=1}^{T} p_i$ and $\mu_T = \mu(L) = \sum_{i=1}^{L} ip_i = \omega_0\mu_0 + \omega_1\mu_1$.

The optimal threshold, T^* is the threshold which satisfies

$$T^* = \underset{1\le T<L}{\arg\max}\, \sigma_B^2(T), \tag{16}$$

where $\sigma_B^2(T)$ is the between-class variance given by

$$\sigma_B^2(T) = \omega_0\omega_1(\mu_1 - \mu_0)^2, \tag{17}$$

or

$$\sigma_B^2(T) = \frac{[\mu_T\omega(T) - \mu(T)]^2}{\omega(T)[1 - \omega(T)]}. \tag{18}$$

3.3.2. Modified Otsu threshold

Otsu threshold is not the only means to separate the background and the objects. For example, a modified Otsu threshold was proposed in Somasundaram and Genish (2012). Instead of using the class mean values, the standard deviation values are used in Eq. (17). The modified between-class variance is then given by

$$\sigma_B^2(T) = \omega_0\omega_1(S_B - S_O)^2, \tag{19}$$

where S_B and S_O are standard deviation values for the background pixels and the object pixels, respectively.

3.4. Transformed Vegetation Index (TVI), corrected TVI (CTVI), and Thiam's TVI (TTVI)

TVI, CTVI, and TTVI are three VIs that are derived from NDVI (Mróz and Sobieraj, 2004). The TVI is given by

$$TVI = \sqrt{NDVI + 0.5}, \tag{20}$$

for $NDVI \ge -0.5$. To solve the problem of $NDVI < -0.5$, CTVI was proposed. CTVI is given by

$$CTVI = \frac{NDVI + 0.5}{|NDVI + 0.5|}\sqrt{|NDVI + 0.5|}, \tag{21}$$

where $\sqrt{|NDVI + 0.5|}$ is the TTVI.

3.5. Perpendicular Vegetation Index (PVI)

The original PVI takes into account the difference between bare soil reflectance and vegetation reflectance. In particular, it measures the "greenness" level of the vegetation which includes the bare soil and then subtract the bare soil reflectance to obtain the vegetation index. The formula uses two bands, i.e. red channel and NIR channel. The formula is given by Bannari et al. (1995)

$$PVI = \sqrt{(R_s - R_v)^2 + (NIR_s - NIR_v)^2}, \quad (22)$$

where subscripts s and v refer to soil and vegetation, respectively.

An improved PVI formula is given by Bannari et al. (1995)

$$PVI = \frac{NIR - aR - b}{\sqrt{a^2 + 1}}, \quad (23)$$

where a is the slope of the bare soil line and b is the y-axis intercept of the bare soil line.

4. Feature-based recognition

As previously stated in Section 1.2, ML algorithms require features for their operation. In particular, features extracted from the image, such as color, texture, and shape, are fed to the ML algorithms. Conventionally, these methods require manual design of features (engineered features) and heavily depend on image acquisition methods, preprocessing techniques, and the accuracy of feature extraction. For engineered features, researchers have used biological morphology, spectral, or texture features of the field images for the classification of weed and crops. Note that traditional ML methods typically require a smaller sample size and a shorter training time when compared to their DL counterparts due to less computational requirements and complexity. However, the type of features used, the selection, and tuning of those features play a critical role in determining the ability of the classifier to discriminate between crops and weeds.

4.1. Biological morphology features

The physical characteristics of plants are referred to as morphology features. It is obvious that each plant has unique physical traits, such as leaf shape and size. For example, the leaves of *Pluchea indica* have the form of *abovatus*, i.e., round shape like eggs (Susetyarini et al., 2020). Therefore, shape abstraction in regard to plant geometry and structure would give sufficient information on the plant's identity. The approach is mainly based on shape features from leaf edge patterns which includes leaf curvature or lobe. Often, multiple features which contribute to the recognition of the plant are used, such as area, dimensionless ratios, length, perimeter, moments and width of either the leaf or the plant are used. In general, there are many morphological features that can be used: geometry, shape, size, and color of plant organs (e.g., root, stem, leaf and flower) (Speck and Speck, 2021).

Søgaard (2005) used active shape models to classify weed species. The database of the shape model consisted of 19 major weed species in Danish fields. The database holds the information of leaf and plant features. The features included the shape and structure definition of the plants. The database was formed from seedling growth to two true leaves. Color images were taken from individual seedings. A plastic diffuser was used to prevent image overexposure. The images were then categorized using template matching, where each species was discriminated against based on template deformation against the optimal fit. Fig. 5 shows the model applied on weed images.

Ahmed et al. (2012) used the feature vector of shape, color, and moment. An ML model was validated with 224 images consisting of chilli, pigweed, marsh herb, lamb's quarters, cogon grass, and bur cucumber. The authors noted that the method was dependent on a large number of images to ensure a more robust model. They also noted that plant holes and image noise affected the model.

The biological morphology approach to various machine vision methods mentioned previously achieved high recognition rates under optimal conditions. This shows the theoretical feasibility of using such methods in environments where plants are in controlled, damage-free, and unoccluded environment.

4.2. Spectral features

Previous studies by other researchers have employed the reflection of light from plants to identify their types. This feature operates based on the color or spectral reflectance exhibited by the leaves. Remarkably, this technique can effectively identify the plant type even when there is partial occlusion of the plant. The unique cellular characteristics of a leaf lead to distinct responses when exposed to specific wavelengths of electromagnetic waves. For example, there is a significant difference in NIR reflectance between dicotyledonous and monocotyledonous plants (Kim and Reid, 2006). Moreover, other than NIR, some authors have used visible waveband spectrum and water absorption waveband spectrum as features (Kim and Reid, 2006; Hatfield and Pinter, 1993).

Woebbecke et al., 1995 proposed the use of color features in standard color slide images to discriminate weeds from the soil. Light reflected from the leaves of the plants were captured and used for classification based on the reflectance wavelength in red, NIR, green, and blue. Their system classified leaves from soybeans, giant foxtail, ivy leaf, morning-glory, and velvetleaf. Reflectances from the other edges of the leaves were used to compute mean, variance, and skewness which were used to determine the type of plant. They experimented with the classifier by either including or excluding leaf orientation in

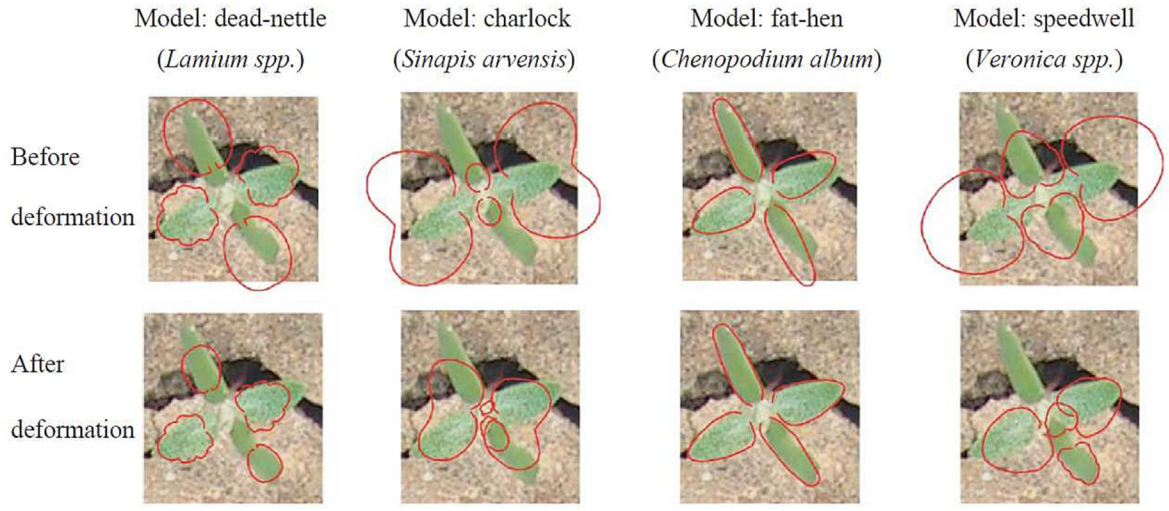


Fig. 5. Example of active shape models by Søgaard (2005).

the classifier. It was found that by excluding leaf orientation features, the classifier managed to obtain a higher classification rate than when including leaf orientation. The researchers also concluded that color images could distinguish between plants and the background but were not sufficient to determine the plant species. It was noted that this technique was affected by leaf orientation as the reflected wavelength changes with the orientation of the leaf.

Another example of using spectral feature for classifying carrots, cabbage, and weeds can be found in Hemming and Rath (2001). The classifier was based on weighted fuzzy classification. It was shown that color features could discriminate plants from the soil, and the classifier showed an improvement in terms of classification accuracy when color features were used in addition to morphology features. Their method resulted in good classification rates depending on the growth and density of the weeds. They noted that this technique is susceptible to occluded leaves.

Spectral features have also been used in Feyaerts and Van Gool (2001). In particular, blue-NIR spectral was observed to discriminate sugar beet plants from weeds. It cannot be denied that crop and weed exhibit comparable spectral signatures. Therefore, Object-Based Image Analysis (OBIA) techniques can improve classification results (De Castro et al., 2018). OBIA works by segmenting images into clusters of neighboring pixels with similar spectral values, referred to as "objects."

4.3. Texture features

Texture features are defined as a collection of metrics used to quantify the apparent texture of an image. The texture provides information on the spatial arrangements of pixels in a region or the entire image. The metrics represent properties such as coarseness, smoothness, and regularity (Bakhshipour et al., 2017).

A texture may be considered as clusters of similarities in an image. There are four types of texture, namely statistical features, structural features, model-based features, and transformation-based features. Statistical features involve calculating the local features of each point in the image, obtaining a collection of statistical data from the distribution of local features, and evaluating the spatial distribution of grey values. Statistical moments are computed through the image or the region's intensity histogram, and the extracted information are sets of statistics that provide properties such as skewness, flatness, and contrast. The Grey Level Co-occurrence Matrix (GLCM) is a common approach for obtaining statistical characteristics (Haralick et al., 1973).

The structural texture is primarily described as a collection of well-defined texture components. The structural texture is defined by the characteristics and placement rules of texture components. Structural textures are rarely utilized in agricultural applications since they can only represent highly regular textures. The texture image is represented as a linear combination of a probability model or a collection of basis functions for model-based textures, and the model coefficients are utilized as texture features (Rallabandi and Sett, 2005).

Model-based texture characteristics, such as structural texture, are rarely utilized in the literature for plant recognition. The converted features are retrieved by converting the image into a new space where the coordinate system has a near interpretation of texture characteristics. For transformation-based features, curve and wavelet transformation are two commonly utilized transformation techniques (López-Granados, 2011). These transformations are typically carried out before the feature extraction process. The Fourier transform method is rarely utilized because of the lack of spatial location. Several wavelet families, including Haar and Gabor, have been utilized to detect specific novel characteristics that are difficult to describe in the spatial domain. Gabor filters improve spatial positioning approaches and have been utilized in a wide range of weed and crop recognition applications.

The authors in Chaki et al. (2015) used the Gabor filter to convolve a grey image of plant leaves with empirically established parameters to create a collection of complex signals, resulting in real and imaginary portions. GLCM was calculated based on the virtual signal following the Gabor filter, followed by a collection of GLCM-based features. In another study, Kumar and Prema (2016) used rotation-invariant wavelet features for weed identification. For acquiring these characteristics, the radon transform approach was used. It substantially lowered the number of characteristics and identified images in a variety of directions rapidly. The impact of rotation on the input and output images might be shown through Radon transform. The wavelet function was then used to split the Radon transform output into distinct sub-bands, and the textural characteristics, such as energy and uniformity of each sub-band, were calculated. Curvelet transform decomposes the image at multiple sizes and angles, making feature extraction easier.

The authors in Bakhshipour et al. (2017) executed a single-level Haar discrete wavelet decomposition on a grey image to get four sub-images representing approximation images, vertical details, horizontal details, and diagonal details. The GLCM texture was then extracted from the four sub-images. Local Binary Pattern (LBP) is a single channel and rotational invariant texture encoder first proposed by He and Wang (1990) describing the spatial structure of the local image texture. The effectiveness of LBP was later proven and described by Ojala et al. (2002). LBP is often used as a visual texture descriptor for many classification problems. Some of the common uses of LBP are in applications of face detection (Jin et al., 2004), facial recognition (Liu et al., 2016), visual inspection (Tajeripour et al., 2008), motion analysis (Zhao and Pietikainen, 2006), and texture analysis (Mäenpää and Pietikäinen, 2005). Literature has shown that LBP is a powerful texture visual descriptor in many applications; thus its use can be further explored in smart agriculture applications, such as weed classification tasks.

The authors in Ahmed et al. (2011) proposed an automated weed classification with LBP-based template matching. The study evaluated the feasibility of using LBP-based texture patterns to classify weed images into broadleaf and grass categories. This is to note that the classification is of the entire image rather than pixel-level classification (i.e., classify the whole image rather than pixel-level labelling). The images used were colored images comprised of 100 broadleafs and 100 grass weeds samples. LBP features were extracted from the images and were trained with two classifiers. The authors concluded that the rotation invariant LBP had strong discriminative abilities, given the correct parameter setting was chosen for LBP.

In the follow-up work, the authors in Ahmed et al. (2014) proposed using local pattern-based texture descriptors, which includes LBP for the classification of broadleaf and grass weeds. Again, the classification was of the entire image rather than pixel-level classification. The researchers experimented with three local texture operators with varying parameter configurations. From their findings, rotation invariant local patterns could provide stable performance in the presence of orientation variations, illumination variations, image noise, and lower memory and computational cost compared with wavelet methods. The desirable characteristics saved considerable computational power as no preprocessing was needed for illumination normalisation and data augmentation in terms of orientation. It showed that that LBP was able to achieve robust performance in uncontrolled natural environment.

The authors in Le et al. (2019) studied on combining multiple LBP features with a classifier to classify canola, corn maize, and radish. The images used were in a lab controlled environment, where only individual plant species were captured at a given time without including other species. Similar to the previous, the classification was of the entire image rather than pixel-level classification. The authors suggested that in plant classification, the accuracy metric itself was not a sufficient indicator to determine its acceptable performance for plant classification. They suggested that the Precision, Recall, and F1-score metrics were suitable to validate the performance of the model for plant classification. From their findings, it was noticed that when visualising the texture feature distribution, corn leaves' texture features were distinct compared to leaves of canola and radish. Their results showed that combining LBP features could achieve accuracy greater than 91% for corn, canola,

and radish plants and background. The authors concluded that combining multiple robust LBP features greatly improved the accuracy and F1 scores in the validation sets.

The authors in Samarajeewa et al. (2018) examined the identification of flower patches in the aerial images using two techniques, namely LBP classification and CIELAB color space threshold classification. In the paper, LBP set pixel intensity thresholds based on the surrounding environment, allowing high-intensity values to be displayed, effectively distinguishing the plants from the background. The CIELAB threshold algorithm determined a threshold value for each RGB channel based on the histogram. Then, in both approaches, erosion was used to eliminate any remaining noise.

Existing research work has shown that there are only few research papers discussing LBP applications in weed classification tasks. Most of the papers utilize LBP to classify entire images or image patches rather than pixel-level classification (Ahmed et al., 2014, 2011; Le et al., 2019; Samarajeewa et al., 2018). Despite the advantages and desirable properties of LBP described in the literature for entire image weed classification, LBP has yet to be applied in pixel-level weed classification. The main challenge of pixel-level classification is the computational cost required when computing multiple LBP feature sets over each image window that contain plant pixels. Each image window represents a plant pixel label (i.e., crop pixel or weed pixel). The high computational cost hinders the advancement of LBP for pixel-level classification. With the current advancement of Application Specific Integrated Circuit (ASIC)-based real-time chip-based LBP extractor, the implementation of a real-time pixel-level weed classification becomes easier. Table 1 shows a summary of the three features.

5. Classification algorithms

5.1. ML-based algorithms

ML-based classifiers are coupled with various features to increase the accuracy of the system for precision agriculture. These traditional methods, once again, require to be designed on hand crafted features and are highly depended on the image quality, pre-processing, and learning algorithms. Advantages of ML-based techniques are advocated on grounds of small sample sizes required to train the system and low requirements on graphics processing units, thus, making agricultural machinery and equipment inexpensive. Traditional ML methods could give efficient performance in research, given the soil conditions and illumination remain the same throughout. These conditions make the real-time applications difficult to be implemented.

In this subsection, we elaborate several ML classification algorithms that are commonly used for crop/weed discrimination task, such as *k*-Nearest Neighbor (KNN) (Ahmad et al., 2011), Artificial Neural Networks (ANNs) (Jeon et al., 2011), decision trees (Goel et al., 2003), and Support Vector Machine (SVM) (Shahbudin et al., 2017; Le et al., 2019). KNN belongs to a non-parametric algorithm, i.e., it does not have the knowledge of the data distribution (lazy learner). It works by calculating the distance (similarity) of a new data sample with its *k* neighbours. Following that, the new data sample will be classified to the class with the shortest distance. Note that *k* is usually an odd number to avoid a tie.

ANN is an ML algorithm that is designed to mimic the principle of human brain. An ANN consists of layers of linked artificial neurons. The fundamental ANN structure is known as a perceptron. A perceptron works by receiving input signals, weighting and summing them. Subsequently, it passes the sum to a nonlinear activation function to produce the output. ANN is trained to adjust the weights that minimize the difference between the target output and the resulted output. The backpropagation algorithm is usually used to optimize the weights.

SVM is considered as a linear classifier. Suppose that we have two classes, SVM creates a hyperplane that optimally separates the two classes. For multiclass classification problem, we can extend the concept by using Error-Correcting Output Coding (ECOC) with either One vs. All (OVA), One vs. One (OVO), Dense Random (DR), or Sparse Random (SR) configuration (Wong et al., 2021). If the data samples are not distributed linearly, a kernel approach, such as the Radial Basis Function (RBF), can be employed to increase the dimension.

A decision tree is supervised ML algorithm with tree-like structure that can be used for both classification and regression. A decision tree consists of a root node, splitting nodes, branch/sub-tree nodes, decision nodes, and terminal nodes. Several algorithms have been used for building the decision tree. In general, entropy measure is used to find the appropriate feature to split the data at each node of the tree.

5.2. Ensemble learning

Ensemble learning methods employ social learning behaviours like those observed in human social interactions, such as working in a group, receiving peer feedback, and voting before deciding. Working in a group allows each member to have various perspectives and representations of the data, which aids in getting a more reliable prediction (Kuncheva, 2014). As a result, ensemble learning approaches are intended to increase the prediction performance of a particular statistical learner or model by combining basic base learners. An interesting characteristic of ensemble learning is that it encourages diversity among simple base learners (Bishop, 2006). Individual learners create independent errors; therefore, diversity would minimize the error generated by the learners. There are several techniques for building an ensemble model. However, it has been discovered that utilizing various learning sets increases variety within the ensemble (Torres Sospedra et al., 2011).

While the notion of combining classifiers is not new, research into various combination techniques, features, classifiers, and innovative applications is still ongoing. The authors in Fumera and Roli (2005) offered a theoretical and experimental study of linear combiners for multiclassifier models. Their study focused on basic and weighted averaging at the categorization decision level. The research revealed that the output of linear combinations was affected by the accuracy and correlation of the individual classifiers. Weighted averaging outperformed basic averaging, according to the data. Weight optimization for classifiers was still an unresolved issue. The authors suggested that future studies might involve the use of metaheuristic optimization techniques.

Table 1
Summary of feature types.

	Morphology	Spectral	Textural
Characteristics	Physical characteristics	Reflection of light	Spatial arrangements of pixels
Examples	Leaf shape, stem color	NIR, infrared	LBP, GLCM
Pros	Rotation and scale invariant Useful in segmentation tasks Visible to human	Rich information Discriminative power	Discriminative power Robust to noise
Cons	Noise sensitivity Limited contextual information	High dimesionality Limited in spatial information More complex	High computational No spatial information

The authors in Kittler et al. (1998) presented a theoretical structure for classifier combinations that is universal. While the authors focused on ensembles with unique feature sets, in their comparable work (Kittler, 1998), both unique and common characteristics were investigated. It was created as a mathematical framework using the sum and product rules. The rules serve as the foundation for others, such as maximum, minimum, median, and majority voting. Moreover, the authors ran extensive tests on the handwritten digits dataset. The sum rule outperformed other rules despite their most stringent assumptions. It demonstrates that the prediction error have a considerably less impact on the sum rule; hence, the theoretical paradigm established is consistent with their previous experimental results.

In the study of Torres-Sospedra and Nebot (2014), the authors utilized an ensemble of NN classifiers combined with noisy statistical characteristics of HSL (hue, saturation, brightness) color images in orange groves to classify weeds. Their implementation included boosting techniques and was tested on 130 images of orange trees with weeds, resulting in a good classification accuracy. The authors discovered that the ensemble is appropriate for weed detection with noisy characteristics. The authors in Ahmad et al. (2018) utilized AdaBoost with Nave Bayes, a boosted ensemble technique combined with statistical visual characteristics, to categorize wide and grass weed images in another study. Their technique correctly identified 250 images with a classification accuracy ranging from 94.72% to 98.40%. The authors discovered that utilising ensemble techniques improved overall accuracy by 4.7% compared to using simply an individual classifier in difficult classification tasks such as weed classification.

Combining different perspectives, forecasts, or models to minimize uncertainty is not a novel concept. Around 200 years ago, Laplace officially stated that he had demonstrated that the correct combination of two probabilistic techniques would produce better results than compared to only one (Laplace, 1820). Numerous methods have been developed with the ensemble technique in combining various models, including aggregation, combination, committee, and fusion (Drucker et al., 1994; Lam and Suen, 1995; Kittler et al., 1998; Polikar et al., 2008).

5.3. DL-based algorithms

DL algorithms use a large number of images to train and validate for recognition and classification. With further progress in computational power as well as the availability of data, DL algorithms have gained popularity due to their capability to extract multidimensional and multiscale spatially essential information of crops and weeds using Convolutional Neural Networks (CNNs) (Dyrmann et al., 2016). The basis of CNN is a stack of convolutional layers which basically perform filtering process (Albawi et al., 2017). Each convolutional layer accepts input data, transform as required and transfers them to the next layer. The convolutional operation finally helps in simplifying the data for better processing and feature extraction.

In specific tasks like image classification and object detection and recognition, DL algorithms possess many advantages over conventional learning approaches. The disadvantages of the conventional human intervened feature extraction methods are usually overcome by its capability of enhanced data expression by automated feature extraction. Recently, a lot of research articles on DL-based weed detection and classification have been published.

As implicitly mentioned in the previous section, there are two types of weed-crop discrimiation approaches that are investigated: pixel-level (or pixel-wise) and image-wise classifications. This subsection mainly focuses on pixel-wise classification. In the pixel-wise classification, each pixel in an image is segmented and marked as weed or crop pixel. Pixel-wise approaches offer the advantage of being able to estimate the yield of each individual pixel. In applications that specifically require this level of information, using a pixel-wise implementation would be advantageous.

The authors in Di Cicco et al. (2017) used a pixel-wise classification CNN model SegNet. The SegNet model uses an encoder-decoder network structure and the fully connected layer is a softmax layer which enables pixel classification. The author used a synthetic dataset generated by procedural generative model with real world plant textures in a virtual environment to train the classification model. The method was able to reduce efforts on manual pixel annotation of real world images by a human. The synthetic dataset was used to train SegNet for crop-weed pixel classification and was able to achieve per-class average accuracy of 91.3%.

The authors in McCool et al. (2017) proposed the use of leveraging Deep CNN (DCNN) to learn lightweight models for pixel level crop-weed classification problems. The authors proposed three stages to train the classification model. The first stage used a pre-trained DCNN model which provides cutting edge classification performance but at a high computational cost. In mobile automated field robotics application, this model was not suitable with low frame rate of 0.12 fps. Thus, the second stage used the pre-trained DCNN model as a teacher model to train a lightweight DCNN student model. This method was known as model compression which results in lower accuracy, but much lower complexity. The third stage involved multiple lightweight model combined as a mixture of models which resulted in higher classification performance. The classification rate of the combination of lightweight models was 88.9%. The research found that while cutting edge DCNN model was able to achieve 93.9% classification accuracy, it was only able to process at 0.12 fps. The mixture of combined lightweight DCNN was able to compute at between 1.07 and 1.83 fps which is faster with less parameters.

The authors in Haug and Ostermann (2015) proposed a plant classification model without the use of segmentation. Often, in the agriculture field, crop and weeds grow too close together and overlapping is a common occurrence. Thus, the authors proposed the use of a Random Forest classifier with statistical features to classify pixels rather than segments. The classification of pixels eliminated the challenges which were associated with segmentation methods such as overlapping and complex background. The classifier was used to estimate the spatial pixel location based on features of overlapping neighborhood pixels. The model was further spatially smoothed with Markov Random Field. A sample of the prediction from this model is shown in Fig. 6. The model was able to achieve an average classification accuracy of 85.9% in the CWFID benchmark dataset on binary classification.

One of the most prominent DL approach that is specific for segmentation is U-Net. U-Net has commonly been applied for medical images segmentation as demonstrated in Arun et al. (2020). The authors in Subeesh et al. (2022) applied various pre-trained models for distinguishing bell peppers and weeds. The overall accuracy of the investigated models ranged from 94.5% to 97.7% with InceptionV3 outperformed the other models with 97.7% accuracy, 98.5% precision, and 97.8% recall at 30-epoch and 16-batch size. Furthermore, the false-positive for this Inception3 model was 1.4% and the false-negative was 0.9%. The approaches detected the entire vegetation object into either weed or crop class and they are not based on segmentation.

In Farooq et al. (2019), remote sensing vegetation patches were evaluated and classified as weed or crops. Because of high spectral similarity between weeds and crops, patch-based classification approaches were used in the paper. Furthermore, the methods of CNN and Histogram of Oriented Gradients (HoG) were assessed and compared. In Brilhador et al. (2019), the authors investigated the impact of individual data augmentation transformations on the pixel-level classification of crops and weeds when using a DL model. The research focused on implementation of the image augmentation approach to increase training image sets similar to the Generative Adversarial Networks (GAN) approaches in generating synthetic images. In Wang et al. (2020), an encoder-decoder DL network was investigated for pixel-wise semantic crop and weed segmentation. To optimize the network's input, different input representations, including different color space transformations

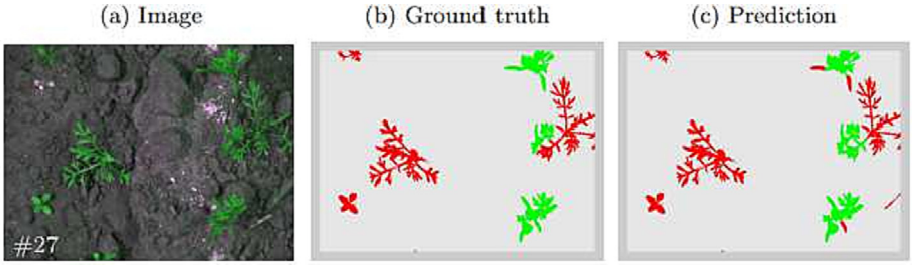


Fig. 6. Example pixel based classification without segmentation by Haug and Ostermann (2015).

and color indices, were compared. In Sa et al. (2018), the authors proposed a method for dense semantic weed categorization using multispectral Micro Aerial Vehicle (MAV) images. In particular, the authors employed SegNet, a recently designed encoder-decoder cascaded CNN, to infer dense semantic classes in sugar beet and weed datasets.

The authors in Jiang et al. (2020) proposed a CNN-based Graph Convolutional Network (GCN). Based on semi-supervised learning, the GCN graph exploited labelled and unlabeled features to enrich the model, and testing samples obtained label information from labelled weed data via propagating through the network. GCN-ResNet-101 obtained 97.80%, 99.37%, 98.93%, and 96.51% recognition accuracies on four separate weed datasets, outperforming state-of-the-art networks (AlexNet, VGG16 and ResNet-101). In Dyrmann et al. (2016), the authors used a CNN to classify seedlings of 22 plant species. Plants were accurately classified 86.2% of the time. In dos Santos Ferreira et al. (2017), a DL method was proposed for soybean weed detection. In particular, CNN was used on SLIC superpixels-segmented images and UAV photos formed a picture database. This experiment obtained 97% weed detection accuracy using CNN.

In Bah et al. (2018), UAV images were collected and an attempt was made to classify into crop/weed pixels. Instead of manual annotation of pixels which is time consuming, the authors proposed an unsupervised approach. The procedure contains multiple steps. First, crop rows were detected to identify inter-row weeds. Inter-row weeds formed the training dataset in the second phase which was to distinguish from the weed. CNNs on this dataset were used to build a model which was able to detect the crop and the weeds in the images. The results obtained were comparable to those of traditional supervised training data labeling, with differences in accuracy of 1.5% in the spinach field and 6% in the bean field when compared with supervised approach (hand annotation).

In Chavan and Nandedkar (2018), AlexNet and VGGNET were combined in AgroAVNET for crop weed classification. Incremental learning used a system to learn new weeds and crops more successfully. Performance was compared to AlexNet, VGGNET, and other approaches. The proposed approach was evaluated using seedling dataset. The implementation yielded 98.21% (7 species detection) and 93.64% (12 species detection). In Hu et al. (2020), the authors proposed a graph-based DL architecture that recognizes weeds from RGB photographs of range lands, called Graph Weeds Net (GWN). GWN collects regional patterns and creates multi-scale graph representations for weed classification. GWN also suggests key regions for robotic in-field actions. The architecture achieved state-of-the-art performance with top-1 accuracy (98.1%).

Another category of DL algorithms is Recurrent Neural Networks (RNNs) (Hochreiter and Schmidhuber, 1997; Verma and Dubey, 2021). The feedback loop available in RNNs allows them to perform as an expert forecasting machine rather than recognition by CNNs. The networks try to make an influence of previous inputs in current inputs and outputs, which helps in increasing the overall performance. They provide a pixel-level classification for all the multispectral images available. The generator in the network works in creating photo-realistic images, aiming for extra training data.

Table. 2 summarizes some significant research exploration involving DL. The literature review has revealed that the majority of implementations have utilized CNN as the foundational approach for either segmentation-(or pixel-wise) or annotation (or image-wise)-based classification.

6. Publicly available datasets

Acquiring crop/weed images using UAV is not an easy task. Fortunately, there exist various public domain datasets that enable researchers to explore implementation of ML algorithm. We also divide the datasets into two types: segmentation and annotation. Due to the morphology of the plants, pixel-wise classification is more complex as they are mostly in green spectrum. Most of these datasets provide a logical mask to indicate if the referred pixel is weed or crop category.

Table 2
DL implementation on weed crop recognition.

Authors	Approaches	Application	Achievement (Accuracy)	Segmentation/Annotation
Subeesh et al., 2022	Pre-trained DL models	Bell pepper vs. Weed	94.5%–97.7%	Annotation
Arun et al., 2020	U-net	Carrot vs weed (CWFID)	95.34%	Segmenttion
Farooq et al., 2019	CNN	Remote sensing vegetation patches	94.72%	Segmentation
Brilhador et al., 2019	CNN/data augmentation	Carrot vs weed (CWFID)	83.44% (average dice similarity)	Segmentation
Wang et al., 2020	Coder/encoder	Sugar beet vs weed (sugar beet dataset)	96.12%	Segmentation
Sa et al., 2018	Coder/encoder CNN	Sugar beet vs weed (sugar beet dataset)	~0.8 for F1/ 0.78 for AUC	Segmentation
Jiang et al., 2020	GCN	Multiple public dataset	96.51%–99.37%	Segmentation
Dyrmann et al., 2016	CNN	22 Species of seedling	86.2%	Annotation
dos Santos Ferreira et al., 2017	CNN	Weed vs soybean	97%	Segmentation
Bah et al., 2018	Unsupervised method/CNN	Spinach/bean vs weed	94%/88.73%	Segmentation
Chavan and Nandedkar, 2018	Hybrid ResNet/AlexNet(Pre-trained CNN)	Dataset of seedlings from various species.	98.21% (7 species) 93.64% (12 species)	Segmentation
Hu et al., 2020	Graph deep learning	DeepWeeds dataset	98.21%	Segmentation

Table 3
Details on publicly available dataset.

Name	Author	No of samples	Application	Segmentation/Annotation
CWFID	Haug and Ostermann, 2015	60	Carrot-weed	Segmentation
Cowpea dataset	LR et al., 2021	150	Southern pea-weed	Segmentation
Carrot-Weed dataset	Lameski et al., 2017	39	Carrot-weed	Segmentation
Crop-weed discrimination dataset	Bosilj et al., 2020	20	Onion-weed	Segmentation
Paddy crops-weeds digital image dataset	Kamath and Balachnadra, 2021	300	Paddy-weed	Segmentation
Rice seedling and weed image segmentation	Ma et al., 2019	224	Rice seedling-weed	Segmentation
Food crops and weed images	Sudars et al., 2020	1118	6 food crops-8 weed species	Annotation
Weed species classification	Skovsen et al., 2019	261	Vegetation biomass	Segmentation
Plant seedling classification	Giselsson et al., 2017	961	12 species at several growth stages	Annotation
DeepWeeds	Olsen et al., 2019	17509	8 species of weeds native to Australia in situ with neighboring flora	Annotation

For pixel-based implementation, dataset normally consist of several High Definition (HD) images in the order of 20–100 as the images can be divided into smaller sections for evaluations/training. For the annotation type, the object plants are segmented using green index. Subsequently, the entire image can be categorized as either weed or crop. This category normally carries more images as most implementation applies some forms of DL approach such as CNN. The summary of the datasets is shown in Table 3.

7. Challenges and opportunities

The problem of identifying background from vegetation (crop or weed) is much more straightforward. However, this still presents a challenge. Agricultural scenes often have cluttered and textured backgrounds, which can interfere with the accurate segmentation of weed and crop regions. Shadows, soil variations, and other artifacts can obscure plant boundaries. Environmental conditions, such as soil type, humidity, and temperature, can affect plant appearances. Models trained in one environment may not generalize well to different conditions.

The key challenge in crop-weed classification for annotation-based techniques is the construction of the classification model and the optimization of the model parameters. The classifier model, like any other image classification issue, is created for specific applications, and its parameters must be fine-tuned and optimized. Classifier optimization necessitates the use of several algorithms in order to attain a high classification rate while minimizing false positives and data overfitting.

Pixel-wise segmentation for weed–crop discrimination is also a challenging task. The most apparent is the high expense of resources and computation. Each pixel must be identified to either a crop or a weed pixel, which is a difficult task. This also has an impact on the annotation task for each pixel, which may result in incorrect annotation. It is also worth noting that the majority of agriculture segmentation takes use of UAV datasets. It is evident that weed pixels would be substantially smaller than crop pixels. This is known as a class imbalance situation in machine learning. Models may become biassed in favour of the dominant class (crops), resulting in poor weed segmentation. Finally, we see that some weeds, in terms of color, texture, and shape, strongly resemble certain crops. Therefore, distinguishing between these visually similar classes becomes challenging, especially when the plants are in early growth stages.

In general, a few key challenges are listed out in the following:

- The vegetation images obtained for the training of the ML model are often obtained at different growth stages of the vegetation with varying vegetation orientation and size.
- In a natural growing environment, crop and weed tend to overlap each other, thus applying image segmentation methods to separate crops and weeds is difficult.
- When using too many features, clustering or forming of the relationship between observations becomes more difficult and complex for the model or, in other words, the “curse of dimensionality” issue is the challenge.

One essential to encouraging development is the use of edge devices rather than offline or cloud-based alternatives. High quality images can be difficult to create and store. The majority of the ideas presented here are computationally demanding and unsuitable for edge devices. More research work using techniques with reduced computation complexity or low energy consumption, such as FPGA-based approaches, may be included. Because the majority of these systems are linked to the Internet of Things (IoT), edge devices can further optimize data transmission and reduce the number of data packets delivered.

8. Conclusion

Various techniques employing ML have been investigated in this paper. We present an overview of crop and weed recognition or discriminating algorithms. The biggest problem is how the technology is used, followed by the ways used owing to crop and weed morphology. As a result, determining which techniques are preferable on a “apple-to-apple” basis is extremely challenging. In this regard, stakeholders may consider implementing ways that are most comparable in order to benefit from proven outcomes. We expect that this paper may give a complete overview of crop-weed detection in Agriculture 5.0 point of view.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was supported partly by ASEAN-India Collaborative R&D scheme under ASEAN-India S&T Development Fund (AISTDF), File Number: CRD/2020/000248 and partly by Universitas Indonesia's International Indexed Publication (PUTI) Q2 Grant, year 2023, number: NKB-803/UN2.RST/HKP.05.00/2023.

References

Ahmad, I., Siddiqi, M.H., Fatima, I., Lee, S., Lee, Y.-K., 2011. Weed classification based on haar wavelet transform via k-nearest neighbor (k—nn) for real-time automatic sprayer control system. Proceedings of the 5th International Conference on Ubiquitous Information Management and Communication, ICUIMC '11, New York, NY, USA.

Ahmad, J., Muhammad, K., Ahmad, I., Ahmad, W., Smith, M.L., Smith, L.N., Jain, D.K., Wang, H., Mehmood, I., 2018. Visual features based boosted classification of weeds for real-time selective herbicide sprayer systems. Comput. Ind. 98, 23–33.

Ahmed, F., Bari, A.H., Shihavuddin, A., Al-Mamun, H.A., Kwan, P., 2011. A study on local binary pattern for automated weed classification using template matching and support vector machine. 2011 IEEE 12th International Symposium on Computational Intelligence and Informatics (CINTI), pp. 329–334.
Ahmed, F., Al-Mamun, H.A., Bari, A.H., Hossain, E., Kwan, P., 2012. Classification of crops and weeds from digital images: a support vector machine approach. Crop Prot. 40, 98–104.
Ahmed, F., Kabir, M.H., Bhuyan, S., Bari, H., Hossain, E., 2014. Automated weed classification with local pattern-based texture descriptors. Int. Arab J. Inform. Technol. 11 (1), 87–94.
Albawi, S., Mohammed, T.A., Al-Zawi, S., 2017. Understanding of a convolutional neural network. 2017 International Conference on Engineering and Technology (ICET), pp. 1–6.
Arun, R.A., Umamaheswari, S., Jain, A.V., 2020. Reduced u-net architecture for classifying crop and weed using pixel-wise segmentation. 2020 IEEE International Conference for Innovation in Technology (INOCON), pp. 1–6.
Bah, M.D., Hafiane, A., Canals, R., 2017. Weeds detection in uav imagery using slic and the hough transform. 2017 Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA), pp. 1–6.
Bah, M., Hafiane, A., Canals, R., 2018. Deep learning with unsupervised data labeling for weed detection in line crops in uav images. Remote Sensing 10.
Bakhshipour, A., Jafari, A., Nassiri, S.M., Zare, D., 2017. Weed segmentation using texture features extracted from wavelet sub-images. Biosyst. Eng. 157, 1–12.
Bannari, A., Morin, D., Bonn, F., Huete, A.R., 1995. A review of vegetation indices. Remote Sens. Rev. 13 (1–2), 95–120.
Bishop, C.M., 2006. Pattern Recognition and Machine Learning.
Bosilj, P., Duckett, T., Cielniak, G., 2018. Connected attribute morphology for unified vegetation segmentation and classification in precision agriculture. Comput. Indus. 98, 226–240.
Bosilj, P., Aptoula, E., Duckett, T., Cielniak, G., 2020. Transfer learning between crop types for semantic segmentation of crops versus weeds in precision agriculture. J. Field Robot. 37 (1), 7–19.
Brilhador, A., Gutoski, M., Hattori, L.T., de Souza Inácio, A., Lazzaretti, A.E., Lopes, H.S., 2019. Classification of weeds and crops at the pixel-level using convolutional neural networks and data augmentation. 2019 IEEE Latin American Conference on Computational Intelligence (LA-CCI), pp. 1–6.
Chaki, J., Parekh, R., Bhattacharya, S., 2015. Plant leaf recognition using texture and shape features with neural classifiers. Pattern Recogn. Lett. 58, 61–68.
Chang, Y.K., Zaman, Q.U., Schumann, A.W., Percival, D.C., Esau, T.J., Ayalew, G., 2012. Development of color co-occurrence matrix based machine vision algorithms for wild blueberry fields. Smart Agric. Technol 28 (3), 315–323.
Chavan, T.R., Nandedkar, A.V., 2018. Agroavnet for crops and weeds classification: a step forward in automatic farming. Comput. Electron. Agric. 154, 361–372.
De Castro, A.I., Torres-Sánchez, J., Peña, J.M., Jiménez-Brenes, F.M., Csillik, O., López-Granados, F., 2018. An automatic random forest-obia algorithm for early weed mapping between and within crop rows using uav imagery. Remote Sens. (Basel) 10 (2).
Di Cicco, M., Potena, C., Grisetti, G., Pretto, A., 2017. Automatic model based dataset generation for fast and accurate crop and weeds detection. 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 5188–5195.
dos Santos Ferreira, A., Matte Freitas, D., Gonçalves da Silva, G., Pistori, H., Theophilo Folhes, M., 2017. Weed detection in soybean crops using convnets. Comput. Elect. Agric. 143, 314–324.
Drucker, H., Cortes, C., Jackel, L.D., LeCun, Y., Vapnik, V., 1994. Boosting and other ensemble methods. Neural Comput. 6 (6), 1289–1301.
Dyrmann, M., Karstoft, H., Midtiby, H.S., 2016. Plant species classification using deep convolutional neural network. Biosyst. Eng. 151, 72–80.
Farooq, A., Hu, J., Jia, X., 2019. Analysis of spectral bands and spatial resolutions for weed classification via deep convolutional neural network. IEEE Geosci. Remote Sens. Lett. 16 (2), 183–187.
Feyaerts, F., Van Gool, L., 2001. Multi-spectral vision system for weed detection. Pattern Recogn. Lett. 22 (6), 667–674.
Fumera, G., Roli, F., 2005. A theoretical and experimental analysis of linear combiners for multiple classifier systems. IEEE Trans. Pattern Anal. Mach. Intell. 27 (6), 942–956.
Gée, C., Bossu, J., Jones, G., Truchetet, F., 2008. Crop/weed discrimination in perspective agronomic images. Comput. Elect. Agric. 60 (1), 49–59.
Giselsson, T., Jørgensen, R., Jensen, P., Dyrmann, M., Midtiby, H., 2017. A Public Image Database for Benchmark of Plant Seedling Classification Algorithms.
Goel, P., Prasher, S., Patel, R., Landry, J., Bonnell, R., Viau, A., 2003. Classification of hyperspectral data by decision trees and artificial neural networks to identify weed stress and nitrogen status of corn. Comput. Elect. Agric. 39 (2), 67–93.
Haralick, R.M., Shanmugam, K., Dinstein, I.H., 1973. Textural features for image classification. IEEE Trans. Syst. Man Cybern. 6, 610–621.
Hasan, A.S.M.M., Sohel, F., Diepeveen, D., Laga, H., Jones, M.G., 2021. A survey of deep learning techniques for weed detection from images. Comput. Elect. Agric. 184, 106067.
Hatfield, P., Pinter, P., 1993. Remote sensing for crop protection. Crop Prot. 12 (6), 403–413.
Haug, S., Ostermann, J., 2015. A crop/weed field image dataset for the evaluation of computer vision based precision agriculture tasks. Computer Vision - ECCV 2014 Workshops.
He, D.-C., Wang, L., 1990. Texture unit, texture spectrum, and texture analysis. IEEE Trans. Geosci. Remote Sens. 28 (4), 509–512.
Hemming, J., Rath, T., 2001. Pa—precision agriculture: computer-vision-based weed identification under field conditions using controlled lighting. J. Agric. Eng. Res. 78 (3), 233–243.
Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9 (8), 1735–1780.
Hu, K., Coleman, G., Zeng, S., Wang, Z., Walsh, M., 2020. Graph weeds net: a graph-based deep learning method for weed recognition. Comput. Elect. Agric. 174, 105520.
Jakubczyk, K., Siemiatkowska, B., Wieckowski, R., Rapcewicz, J., 2023. Hyperspectral imaging for mobile robot navigation. Sensors 23 (1).
Jeon, H.Y., Tian, L.F., Zhu, H., 2011. Robust crop and weed segmentation under uncontrolled outdoor illumination. Sensors 11 (6), 6270–6283.
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. Comput. Elect. Agric. 174, 105450.
Jin, H., Liu, Q., Lu, H., Tong, X., 2004. Face detection using improved lbp under bayesian framework. Third International Conference on Image and Graphics (ICIG'04), pp. 306–309.
Kamath, R., Balachnadra, M., 2021. Paddy Crop and Weeds Digital Image Dataset.
Kim, Y., Reid, J.F., 2006. Spectral sensing for plant stress assessment - a review. Agric. Biosyst. Eng. 7 (1), 27–41.
Kittler, J., 1998. Combining classifiers: a theoretical framework. Pattern. Anal. Applic. 1 (1), 18–27.
Kittler, J., Hatef, M., Duin, R.P., Matas, J., 1998. On combining classifiers. IEEE Trans. Pattern Anal. Mach. Intell. 20 (3), 226–239.
Kumar, D.A., Prema, P., 2016. A novel wrapping curvelet transformation based angular texture pattern (wctatp) extraction method for weed identification. ICTACT J. Image Video Process. 6 (3).
Kuncheva, L.I., 2014. Combining Pattern Classifiers: Methods and Algorithms. John Wiley & Sons.
Lam, L., Suen, C.Y., 1995. Optimal combinations of pattern classifiers. Pattern Recogn. Lett. 16 (9), 945–954.
Lameski, P., Zdravevski, E., Trajkovik, V., Kulakov, A., 2017. Weed Detection Dataset with RGB Images Taken Under Variable Light Conditions. pp. 112–119.
Laplace, P.S., 1820. Théorie analytique des probabilités. Courcier.
Lati, R.N., Rasmussen, J., Andujar, D., Dorado, J., Berge, T.W., Wellhausen, C., Pflanz, M., Nordmeyer, H., Schirrmann, M., Eizenberg, H., Neve, P., Jørgensen, R.N., Christensen, S., 2021. Site-specific weed management—constraints and opportunities for the weed research community: insights from a workshop. Weed Res. 61 (3), 147–153.
Le, V.N.T., Apopei, B., Alameh, K., 2019. Effective plant discrimination based on the combination of local binary pattern operators and multiclass support vector machine methods. Inform. Process. Agric. 6 (1), 116–131.
Lease, B.A., Wong, W.K., Gopal, L., Chiong, C.W.R., 2020. Pixel-level weed classification using evolutionary selection of local binary pattern in a stochastic optimised ensemble. SN Comput. Sci. 1 (337), 1–13.
Liu, L., Fieguth, P., Zhao, G., Pietikäinen, M., Hu, D., 2016. Extended local binary patterns for face recognition. Inform. Sci. 358, 56–72.
López-Granados, F., 2011. Weed detection for site-specific weed management: mapping and real-time approaches. Weed Res. 51 (1), 1–11.
LR, V.N., Krishnan, A., Krishnan, K.V., P, H, Haritha, Z.A.S.A., 2021. Southern pea/weed field image dataset for semantic segmentation and crop/weed classification using an encoder-decoder network. Proceedings of the International Conference on Systems, Energy & Environment (ICSEE) 2021.
Lu, B., He, Y., Dao, P.D., 2019. Comparing the performance of multispectral and hyperspectral images for estimating vegetation properties. IEEE J. Select. Topics Appl. Earth Observ. Remote Sensing 12 (6), 1784–1797.
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), 1–13.
Mäenpää, T., Pietikäinen, M., 2005. Texture analysis with local binary patterns. Handbook of Pattern Recognition and Computer Vision, pp. 197–216.
Manea, D., Calin, M.A., 2015. Hyperspectral imaging in different light conditions. Imaging Sci. J. 63 (4), 214–219.
McCool, C., Perez, T., Upcroft, B., 2017. Mixtures of lightweight deep convolutional neural networks: applied to agricultural robotics. IEEE Robot. Automat. Lett. 2 (3), 1344–1351.
Mróz, M., Sobieraj, A., 2004. Comparison of several vegetation indices calculated on the basis of a seasonal spot xs time series, and their suitability for land cover and agricultural crop identification. Technical Report. University of Warmia and Mazury in Olsztyn.
Neupane, K., Baysal-Gurel, F., 2021. Automatic identification and monitoring of plant diseases using unmanned aerial vehicles: a review. Remote Sens. (Basel) 13 (19).
Ojala, T., Pietikäinen, M., Mäenpää, T., 2002. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans. Pattern Anal. Mach. Intel. 7, 971–987.
Olsen, A., Konovalov, D.A., Philippa, B., Ridd, P., Wood, J.C., Johns, J., Banks, W., Girgenti, B., Kenny, O., Whinney, J., Calvert, B., Azghadi, M.R., White, R.D., 2019. Deepweeds: a multiclass weed species image dataset for deep learning. Sci. Rep. 9, 2058.
Otsu, N., 1979. A threshold selection method from gray-level histograms. IEEE Trans. Syst. Man Cybern. 9 (1), 62–66.
Patrício, D.I., Rieder, R., 2018. Computer vision and artificial intelligence in precision agriculture for grain crops: a systematic review. Comput. Elect. Agric. 153, 69–81.
Polikar, R., Topalis, A., Parikh, D., Green, D., Frymiare, J., Kounios, J., Clark, C.M., 2008. An ensemble based data fusion approach for early diagnosis of Alzheimer's disease. Inform. Fusion 9 (1), 83–95.
Qin, J., Chao, K., Kim, M.S., Lu, R., Burks, T.F., 2013. Hyperspectral and multispectral imaging for evaluating food safety and quality. J. Food Eng. 118 (2), 157–171.
Rallabandi, V.S., Sett, S., 2005. Unsupervised texture classification and segmentation. WEC. 5, pp. 299–302.

Sa, I., Chen, Z., Popović, M., Khanna, R., Liebisch, F., Nieto, J., Siegwart, R., 2018. Weednet: dense semantic weed classification using multispectral images and mav for smart farming. IEEE Robot. Automat. Lett. 3 (1), 588–595.
Saiz-Rubio, V., Rovira-Más, F., 2020. From smart farming towards agriculture 5.0: a review on crop data management. Agronomy 10 (2).
Samarajeewa, T., Suduwella, C., Jayasuriya, N., Kumarasinghe, P., Gunawardana, K., De Zoysa, K., Keppitiyagama, C., 2018. Identification of lantana camara distribution using convolutional neural networks. 2018 18th International Conference on Advances in ICT for Emerging Regions (ICTer), pp. 221–228.
Sankaran, S., Khot, L.R., Maja, J.M., Ehsani, R., 2013. Comparison of two multiband cameras for use on small uavs in agriculture. 2013 5th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), pp. 1–4.
Shahbudin, S., Zamri, M., Kassim, M., Abdullah, S.A.C., Suliman, S.I., 2017. Weed classification using one class support vector machine. 2017 International Conference on Electrical, Electronics and System Engineering (ICEESE), pp. 7–10.
Skovsen, S., Dyrmann, M., Krogh Mortensen, A., Laursen, M., Gislum, R., Eriksen, J., Farkhani, S., Karstoft, H., Jørgensen, R., 2019. The Grassclover Image Dataset for Semantic and Hierarchical Species Understanding in Agriculture. pp. 2676–2684.
Søgaard, H.T., 2005. Weed classification by active shape models. Biosyst. Eng. 91 (3), 271–281.
Somasundaram, K., Genish, T., 2012. Modified otsu thresholding technique. Commun. Comput. Inform. Sci. 283, 445–448.
Speck, O., Speck, T., 2021. Functional morphology of plants – a key to biomimetic applications. New Phytol. 231 (3), 950–956.
Subeesh, A., Singh, B., Chandel, N.S., Rajwade, Y.A., Rao, K.V.R., Kumar, S.P., Jat, D., 2022. Deep convolutional neural network models for weed detection in polyhouse grown bell peppers. Artif. Intel. Agric. 6, 47–54.
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 Brief 31, 105833.
Susetyarini, E., Wahyono, P., Latifa, R., Nurrohman, E., 2020. The identification of morphological and anatomical structures of pluchea indica. J. Phys. Conf. Ser. 1539 (1), 012001.
Tajeripour, F., Kabir, E., Sheikhi, A., 2008. Fabric defect detection using modified local binary patterns. EURASIP J. Adv. Signal Process. 2008, 60.
Torres Sospedra, J., et al., 2011. Ensembles of artificial neural networks: Analysis and development of design methods. PhD thesisUniversitat Jaume I.
Torres-Sospedra, J., Nebot, P., 2014. Two-stage procedure based on smoothed ensembles of neural networks applied to weed detection in orange groves. Biosyst. Eng. 123, 40–55.
Tripicchio, P., Satler, M., Dabisias, G., Ruffaldi, E., Avizzano, C.A., 2015. Towards smart farming and sustainable agriculture with drones. 2015 International Conference on Intelligent Environments, pp. 140–143.
Verma, T., Dubey, S., 2021. Prediction of diseased rice plant using video processing and LSTM-simple recurrent neural network with comparative study. Multimed. Tools Appl. 80, 29267–29298.
Wang, A., Zhang, W., Wei, X., 2019. A review on weed detection using ground-based machine vision and image processing techniques. Comput. Elect. Agric. 158, 226–240.
Wang, A., Xu, Y., Wei, X., Cui, B., 2020. Semantic segmentation of crop and weed using an encoder-decoder network and image enhancement method under uncontrolled outdoor illumination. IEEE Access 8, 81724–81734.
Woebbecke, D.M., Meyer, G.E., Bargen, K.V., Mortensen, D.A., 1995. Color indices for weed identification under various soil, residue, and lighting conditions. Trans. ASAE 38 (1), 259–269.
Wong, W.K., Juwono, F.H., Apriono, C., 2021. Vision-based malware detection: a transfer learning approach using optimal ecoc-svm configuration. IEEE Access 9, 159262–159270.
Xu, Y., Gao, Z., Khot, L., Meng, X., Zhang, Q., 2018. A real-time weed mapping and precision herbicide spraying system for row crops. Sensors 18 (12).
Xu, X.-k, Li, X.-m, Zhang, R.-h, 2019. Remote configurable image acquisition lifting robot for smart agriculture. 2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC). vol. 1, pp. 1545–1548.
Xue, J., Su, B., 2017. Significant remote sensing vegetation indices: a review of developments and applications. J. Sensors 2017.
Xue, Q., Yang, B., Wang, F., Tian, Z., Bai, H., Li, Q., Cao, D., 2021. Compact, uav-mounted hyperspectral imaging system with automatic geometric distortion rectification. Opt. Express 29 (4), 6092–6112.
Zhao, G., Pietikainen, M., 2006. Local binary pattern descriptors for dynamic texture recognition. 18th International Conference on Pattern Recognition (ICPR'06). vol. 2, pp. 211–214.