2020

Local Binary Pattern based algorithms for the discrimination and detection of crops and weeds with similar morphologies

Vi Nguyen Thanh Le
Edith Cowan University

Edith Cowan University

Copyright Warning

School of Science

AUSTRALIA ECU EDITH COWAN UNIVERSITY

Local Binary Pattern based algorithms for the discrimination and detection of crops and weeds with similar morphologies

This thesis is presented in fulfilment of the requirements for the degree of

Doctor of Philosophy

Vi Nguyen Thanh Le

Supervisors

Prof. Kamal Alameh
Dr. Selam Ahderom

Electron Science Research Institute
School of Science
Edith Cowan University

2020

Declaration

I certify that this thesis does not, to the best of my knowledge and belief:

i. incorporate without acknowledgment any material previously submitted for a degree or diploma in any institution of higher education;

ii. contain any material previously published or written by another person except where due reference is made in the text of this thesis; or

iii. contain any defamatory material;

I also grant permission for the Library at Edith Cowan University to make duplicate copies of my thesis as required.

Vi Nguyen Thanh Le

March 26th, 2020

Acknowledgment

I would like to express my sincerest gratitude to the following individuals and organizations that helped me during my PhD journey. Firstly, I would like to thank Professor Kamal Alameh. His guidance, mentorship, and not to mention the many hours of time committed to proof-reading draft manuscripts proved to be invaluable to my educational experience. Secondly, my academic advisors, Dr. Selam Ahderom and Dr. Beniamin Apopei, whose support and advice greatly facilitated this project. Their wide knowledge has taught me many valuable lessons and helped me to complete this work. They have provided me with a massive source of information throughout my journey. I would like to thank you from the bottom of my heart. Also, I would like to thank Mr. Paul Roach for his administrative assistance.

I would like to acknowledge the financial support for this research project funded by Grains Research Development Corporation (GRDC) and Photonic Detection Systems Pty Ltd. I would also like to thank School of Science and the Electron Science Research Institute for the provision of scholarship funding, without which this project could not have been completed.

Finally, I express my deep gratitude to my parents for their love, encouragement, and giving me all the support I need to finish this thesis.

List of Abbreviations

AMF	Adaptive Median Filter
ANN	Artificial Neural Network
AR	Average Recall
CNN	Convolutional Neural Network
CPU	Central Processing Unit
CUDA	Computer Unified Device Architecture
DCNN	Deep Convolutional Neural Network
DL	Deep Learning
ExGR	Excess Green minus Excess Red
Faster R-CNN	Faster Regions with Convolutional Neural Network features
Fast-RCNN	Fast Regions with Convolutional Neural Network features
FCL	Fully Connected Layers
FCN	Fully Convolutional Network
FLBP	Filtered Local Binary Patterns
FLBPbCM	Filtered LBP-based Contour Mask
FN	False Negative
FP	False Positive
GPU	Graphics Processing Unit
HOG	Histogram of Oriented Gradients
IoU	Intersection over Union
k-FLBPCM	Filtered Local Binary Patterns with Contour Masks and coefficient k
KNN	K-Nearest Neighbour
LBP	Local Binary Pattern
mAP	Mean Average Precision
ML	Machine Learning
NDVI	Normalized Difference Vegetation Index
R, G, B	Red, Green and Blue chromatic values
RCNN	Regions with Convolutional Neural Network features
ReLU	Rectified Linear Unit

ROI	Region of Interest
RPN	Region Proposal Network
SGD	Stochastic Gradient Descent
SIFT	Scale-Invariant Feature Transform
SPPNet	Spatial Pyramid Pooling Network
SSD	Single Shot Detector
SURF	Speeded-Up Robust Features
SVM	Support Vector Machine
TN	True Negative
TP	True Positive
YOLO	You Only Look Once

List of Publications

1. V. N. T. Le, B. Apopei, and K. Alameh, "Effective plant discrimination based on the combination of local binary pattern operators and multiclass support vector machine methods," *Information processing in agriculture*, vol. 6, no. 1, pp. 116-131, 2019.

2. V. N. T. Le, S. Ahderom, B. Apopei, and K. Alameh, "A novel method for detecting morphologically similar crops and weeds based on the combination of contour masks and filtered Local Binary Pattern operators," *GigaScience*, vol. 9, no. 3, 2020, doi: 10.1093/gigascience/giaa017.

3. V. N. T. Le, S. Ahderom, and K. Alameh, "Performances of the LBP Based Algorithm over CNN Models for Detecting Crops and Weeds with Similar Morphologies," *Sensors*, vol. 20, no. 8, p. 2193, 2020.

4. V. N. T. Le, S. Ahderom, and K. Alameh. (2020). Detecting weeds from crops under complex field environments based on Faster RCNN models. Precision Agriculture Journal (Under revision).

Table of Contents

Abstract

In cultivated agricultural fields, weeds are unwanted species that compete with the crop plants for nutrients, water, sunlight and soil, thus constraining their growth. Applying new real-time weed detection and spraying technologies to agriculture would enhance current farming practices, leading to higher crop yields and lower production costs. Various weed detection methods have been developed for Site-Specific Weed Management (SSWM) aimed at maximising the crop yield through efficient control of weeds. Blanket application of herbicide chemicals is currently the most popular weed eradication practice in weed management and weed invasion. However, the excessive use of herbicides has a detrimental impact on the human health, economy and environment. Before weeds are resistant to herbicides and respond better to weed control strategies, it is necessary to control them in the fallow, pre-sowing, early post-emergent and in pasture phases. Moreover, the development of herbicide resistance in weeds is the driving force for inventing precision and automation weed treatments. Various weed detection techniques have been developed to identify weed species in crop fields, aimed at improving the crop quality, reducing herbicide and water usage and minimising environmental impacts.

In this thesis, Local Binary Pattern (LBP)-based algorithms are developed and tested experimentally, which are based on extracting dominant plant features from camera images to precisely detecting weeds from crops in real time. Based on the efficient computation and robustness of the first LBP method, an improved LBP-based method is developed based on using three different LBP operators for plant feature extraction in conjunction with a Support Vector Machine (SVM) method for multiclass plant classification. A 24,000-image dataset, collected using a testing facility under simulated field conditions (Testbed system), is used for algorithm training, validation and testing. The dataset, which is published online under the name "bccr-segset", consists of four subclasses: background, Canola (Brassica napus), Corn (Zea mays), and Wild radish (Raphanus raphanistrum). In addition, the dataset comprises plant images collected at four crop growth stages, for each subclass. The computer-controlled Testbed is designed to rapidly label plant images and generate the "bccr-segset" dataset. Experimental results show that the classification accuracy of the improved LBP-based algorithm is 91.85%, for the four classes.

Due to the similarity of the morphologies of the canola (crop) and wild radish (weed) leaves, the conventional LBP-based method has limited ability to discriminate broadleaf crops from weeds. To overcome this limitation and complex field conditions (illumination variation, poses, viewpoints, and occlusions), a novel LBP-based method (denoted k-FLBPCM) is developed to enhance the classification accuracy of crops and weeds with similar morphologies. Our contributions include (i) the use of opening and closing morphological operators in pre-processing of plant images, (ii) the development of the k-FLBPCM method by combining two methods, namely, the filtered local binary pattern (LBP) method and the contour-based masking method with a coefficient k, and (iii) the optimal use of SVM with the radial basis function (RBF) kernel to precisely identify broadleaf plants based on

their distinctive features. The high performance of this k-FLBPCM method is demonstrated by experimentally attaining up to 98.63% classification accuracy at four different growth stages for all classes of the “bccr-segset” dataset.

To evaluate performance of the k-FLBPCM algorithm in real-time, a comparison analysis between our novel method (k-FLBPCM) and deep convolutional neural networks (DCNNs) is conducted on morphologically similar crops and weeds. Various DCNN models, namely VGG-16, VGG-19, ResNet50 and InceptionV3, are optimised, by fine-tuning their hyper-parameters, and tested. Based on the experimental results on the “bccr-segset” dataset collected from the laboratory and the “fieldtrip_can_weeds” dataset collected from the field under practical environments, the classification accuracies of the DCNN models and the k-FLBPCM method are almost similar. Another experiment is conducted by training the algorithms with plant images obtained at mature stages and testing them at early stages. In this case, the new k-FLBPCM method outperformed the state-of-the-art CNN models in identifying small leaf shapes of canola-radish (crop-weed) at early growth stages, with an order of magnitude lower error rates in comparison with DCNN models. Furthermore, the execution time of the k-FLBPCM method during the training and test phases was faster than the DCNN counterparts, with an identification time difference of approximately 0.224ms per image for the laboratory dataset and 0.346ms per image for the field dataset. These results demonstrate the ability of the k-FLBPCM method to rapidly detect weeds from crops of similar appearance in real time with less data, and generalize to different size plants better than the CNN-based methods.

Chapter 1 – Introduction, Literature Review, Methods and Aims

1.1 Motivation

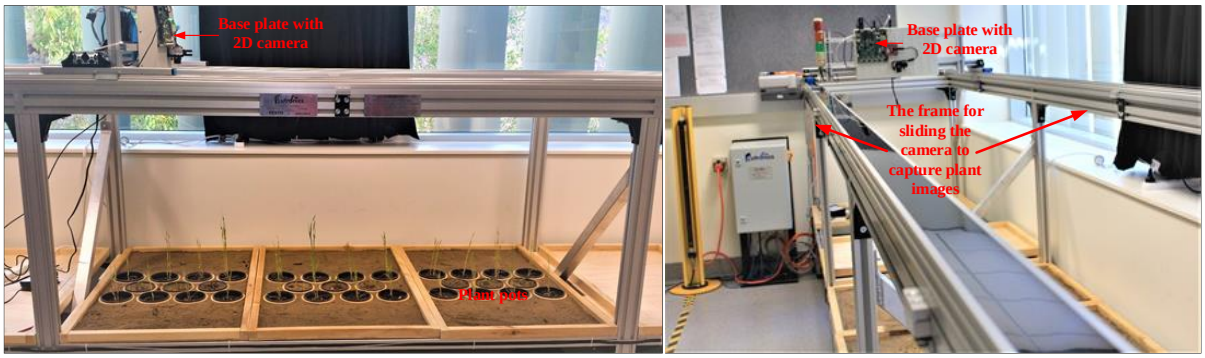
Weed infestation has always been a critical issue that limits the productivity and quality of crops and significantly reduces the farmer's profitability [1]. The most popular method for treating weed infestations is the use of chemical herbicides. However, the excessive use of herbicides gives rise to detrimental problems on the human health and environment, because herbicide waste and residues can be absorbed into foods, groundwater and soil [2-4]. In addition, the frequent use of herbicides increases the herbicide resistance of weeds [5]. Since the overuse of herbicides in agriculture affects the farmer's profitability [6], effective weed detection and spraying is crucial for farmers, since it potentially enable 30-75% savings in herbicide [7], in addition to significant water savings (since herbicide is typically mixed with water before spraying).

Australian grain crops can be grown in two seasons, namely, summer and winter. For instance, wheat, barley and canola are normally planted in winter. Sorghum and sunflowers are often grown in summer. In this project, wheat and barley are particularly investigated, because of their high productivity and importance for the Australian agricultural industry [8]. In Western Australia (WA), wheat is considered the major grain crop, accounting for 70% of annual grain production and bringing A$ 2-3 billion for the economy of this State every year. WA makes up approximately 50% of the total wheat production in Australia. 95% of WA's wheat is exported to Asia and the Middle East [8]. After wheat, barley is the second largest cereal crop in WA and accounts for 25% of the state's total cereal production and generating more than A$ 0.65 billion in barley grain and malt export earnings each year. According to Grains Research and Development Corporation (GRDC), barley crops play a vital role in the Northern grains region, because of their characteristics such as adaptability, tolerance of short seasons, less frost sensitivity at flowering. Particularly, barley is less dependent than wheat on the timing of seasonal breaks and soil moisture profiles [9]. Next, canola is well known for producing one of the world's healthiest vegetable oils with low saturated fat and is considered as an environmentally friendly biofuel. The production of canola (market value of AU$2.2 billion) has increased significantly up to four million tonnes in Australia in the period of 2012 and 2013 [10]. With more than two million tonnes of canola seed exported by Australia every year, Australia has become the world's second largest exporter of canola. However, while canola can be easily grown, farmers need to make more efforts to manage and monitor this crop in comparison with other cereal crops [11].

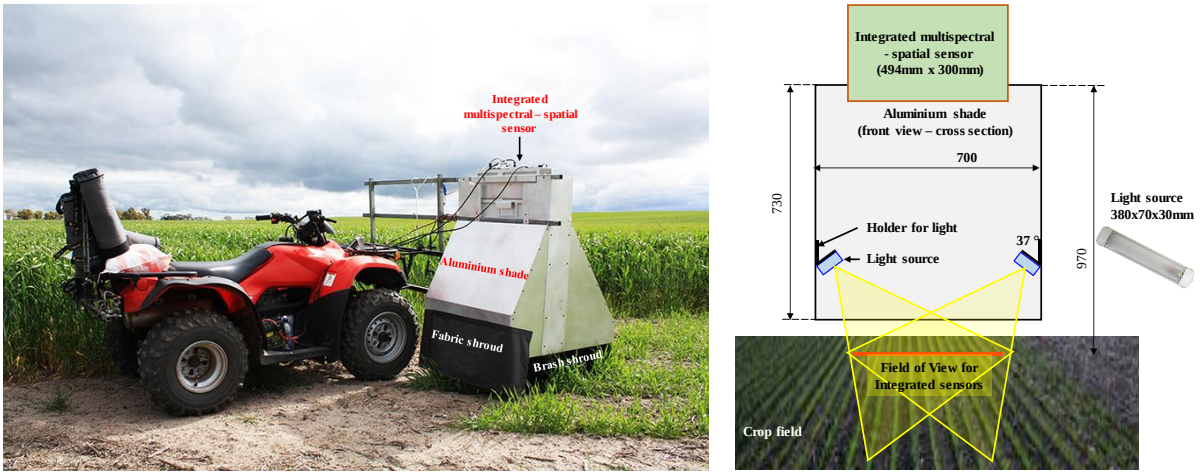
Wild radish (Raphanus raphanistrum) is one of the most competitive and invasive broadleaf weeds throughout Australian cereal-growing regions. Wild radish adapts to different environments and spreads rapidly in patches of varying size, and its herbicide resistance has increased for a wide range of herbicide groups [12, 13]. Several experiments have been conducted in New South Wales to investigate the

detrimental impacts of wild radish on the quality and productivity of canola [14]. The challenging problem of controlling wild radish in canola crops has arisen, especially because of the high spatial correlation (appearance) between canola and wild radish species. Thus, spraying herbicides on only targeted weeds in early growth stage, i.e. before weeds become too widespread and out-of-control, is an effective weed management approach.

To mitigate the effect of weeds on crop yields, precise weed detection and effective management can play an indispensable part of an integrated weed management process. The combination of Machine learning (ML) algorithms and digital image processing enhances the capability of discriminating and detecting weeds under various weather conditions, lighting conditions, leaf overlap, occlusion, and different growth stages, in order to reduce the need for herbicides and major loss on crop yield [15]. Machine vision techniques, in particular, have been widely used in agriculture to discriminate crops and weeds, and their accuracy has been improving at a fast pace [16-18]. However, there still exists some limitations of plant datasets to solve the real agricultural issues and precisely detecting weeds (e.g., wild radish) that look like crops (e.g., canola). In addition, textural and morphological properties of plant leaves are changed at different growth stages. These challenges are the motivation for this research project. In this thesis, we create datasets of plant images by utilizing a testbed system developed at Electron Science Research Institute, Edith Cowan University as shown in Figure 1, to automatically capture plant spatial information, as well as develop and optimize advanced real-time algorithms for improving the accurate identification and detection of crops and weeds with similar morphology at different growth stages. A separate dataset was also created using an integrated weed sensing system of multispectral and spatial sensors, and collected from a commercial farm at Cunderdin, Western Australia as noted in Figure 1.



A Testbed system used to capture plant images in the laboratory



An integrated weed sensing system used to capture plant images under complex field environments

Figure 1. Different viewpoints and structures of systems developed to collect data from the laboratory and practical field environments and built at Electron Science Research Institute, Edith Cowan University.

1.2 Literature Review

Research into the application of digital image processing to the automated detection and discrimination of crop and weed in agriculture is exceptionally diverse, and requires a broad knowledge on image acquisition, segmentation, feature extraction and classification techniques. The variation in crop and weed species used experimentally, along with the different non-standardized approaches adopted by numerous researchers in the capture and processing of (visual) data, makes it extremely difficult to classify and compare the extensive, published research work in this area. This literature review is a selective, comparative overview highlighting pre-processing, segmentation, feature extraction and classification techniques in the context of real-time performance.

1.2.1 Pre-processing

Image pre-processing is the initial step of a weed detection procedure, which focuses on enhancing the visual appearance of original plant images by overcoming the problems of poor contrast and noise. Poor

contrast can be overcome by resampling the captured images and suppressing the problems of shading and background sunlight through the adjustment of the camera settings [19]. Some pre-processing techniques can be applied to remove tiny unexpected objects such as noise removal, low-pass, high-pass, band-pass spatial filtering, mean filtering, median filtering and so on. The ability to reduce the illumination issues using homomorphic filtering was successfully demonstrated on images captured with different environmental conditions [20]. Particularly, homomorphic filtering is used to simultaneously normalize the brightness across an image, increase contrast and reduce illumination variations [21]. Finally, colour conversion and histogram equalisation have been used to detect diseases of plant leaves in initial growth cycles [22].

1.2.2 Image segmentation

Image segmentation refers to the process of partitioning an image (pixels) into multiple segments or regions. Particularly, with regard to weed detection, this process is based on the segmentation of the different pixels in images into plant areas (crops and weeds) and background areas (non-green species, i.e., soil and residues). Removing the background areas of the images enables better plant feature extraction and classification. The advantages and disadvantages of several segmentation techniques are shown in Table 1.

Table 1. Benefits and drawbacks of several segmentation techniques [23]

Methods	**Advantages**	**Disadvantages**
Otsu threshold, 1979	- Enables automatic operation - Can be used widely	- Producing under-segmentation, i.e. some green pixels are not identified in some cases - Slower than the mean intensity method
Normalised Difference Index [24], 1992	- Easy computation - Robustness to lighting, except for extreme conditions	- When the light intensity is too high or too low, this method results in a low performance. - High false positive rate
Excess Green Index (ExG), 1995	- Computational simplicity - To be used widely - Low sensitivity to lighting conditions and background errors - Ability to effectively adapt to the outdoor environment	- When the light intensity is either weak or strong, the performance is low.

Excess Red Index (ExR), 1998	- Easy computation - Reliance on red component only, but green pixels are still extracted - Soil texture segmentation	- When the light intensity is either weak or strong, the performance is low. - Lower accuracy than ExG.
Colour index of vegetation extraction (CIVE), 2003	- Low running time - Effective adaptation to the outdoor environment (except for shadow)	- When the light intensity is either weak or strong, the performance is poor.
Excess Green minus Excess Red (ExGR), 2004	- Extracting green by ExG and eliminating background noise by ExR - Good adaptability in outdoor environments - Does not require complicated thresholding	- When the light intensity is either weak or strong, the performance is low. - Segmentation of shadow pixels as plants (over-segmentation)
Fuzzy Clustering, 2004	- Extracting the plant region of interest from ExG and ExR images - Identification of green plants from the background, such as soil and residue	- When plant pixel coverage is less than 10% in the image, there is not enough colour information to cluster them
Normalised Green-Red Difference Index (NGRDI), 2005	- Reduction of the differences in the exposure time selected by the digital camera - Ability to discriminate between green plants and soil, and normalise illumination variations between different images	- When the light intensity is either weak or strong, the performance is low. - Limited application
Homogeneity threshold, 2006 and 2007	- Ability to recognise small objects - Considering local information, thus this method is quite useful to address light changes	- Time-consumption calculations as it requires several steps
Mean-shift algorithm with Back Propagation Neural Network, 2009	- Classifying plant and non-plant region → Good segmentation	- Long computation time - Low segmentation rate for green areas with shadows

	performance under illumination changes	
Mean-shift algorithm with Fisher Linear Discriminant, 2009	- Separating green from non-green vegetation → Good segmentation rate for green areas with shadows	- Long computation time
Affinity Propagation-Hue Intensity, 2013	- Separating the pixels of crop and background under light conditions and complicated environmental conditions → Robustness and insensitivity to challenging variation of outdoor environmental conditions	- Misclassification of the highlighted regions in leaves
Morphology Modelling, 2013	- Distinguishing the crop and background pixels under complex illumination conditions - Powerful in illumination variation in the field	- Limited improvement, even if different element sizes are used in the training stage
Decision Tree based Segmentation Model, 2013	- Segmenting vegetation from the background - Ability to handling illumination issues, such as shadow and regions with specular reflection - Does not need to optimise the threshold level for each image	- Reliance on training data

1.2.3 Feature extraction

Feature extraction plays a significant role in object detection and recognition. In addition, feature extraction can be considered as the most common and convenient means of data representation for classification issues. It involves the extraction of the most relevant features of an image and the labelling for image classification in the next stage of image processing [25]. Image features are typically divided into three types, namely, colour, shape and texture.

1.2.3.1 Shape features

Shape features can be regarded as one of the pivotal clues that enable the detection and recognition of objects. Shape feature extraction techniques are typically based on contour and region identification. Contour identification methods are based on calculating the shape features only from the shape

boundary, whereas the region identification methods extract the shape features from the entire region. According to a survey on shape feature extraction approaches reported by Yang et al. [26], the visual features of the images, which represent the content of the images, can be considered as shape descriptors, and the more effective the shape descriptors are, the more similar shapes can be found from the image database to recognise objects. However, applying shape features for the detection of crops and weeds could result in a high false positive rate, especially in a field with harsh environmental conditions and plant canopies. This is because shape feature extraction techniques are typically efficient when individual components (i.e. seedling, leaf or plant) are identified [27], although recent studies have proposed methods based on analysing colour and texture features for the classification of weeds and crops in real-world scenarios.

1.2.3.2 Colour features

According to a review on image feature extraction and representation techniques by Tian [28], the generation of a colour histogram is regarded as the most common method for extracting the colour features of images. However, colour feature extraction methods have limitations, including (i) sensitivity to noise, rotation and scale of images and (ii) very high computation time.

1.2.3.3 Texture features

Texture features can be extracted from a group of pixels as opposed to colour features (typically a pixel property). Numerous texture feature extraction techniques have been proposed, which are based on spatial texture feature extraction and spectral texture feature extraction methods. Spatial texture feature extraction methods are based on calculating the pixel statistics or finding the local pixel structures in the original image region, whereas spectral texture feature extraction methods involve transforming an image into the frequency domain and then computing features from the transformed image [28].

Texture methods based on analysing local spatial variations of colour and intensity levels in neighbour pixels have long been considered challenging in the pattern recognition and computer vision field [29]. Typically, in order to enhance the accuracy of shape-based and colour-based plant detection, the texture information needs to be analysed as feature vectors extracted from the patterns of plants. Many detrimental effects of the outdoor conditions must be considered carefully, including variant illumination, different viewpoints, environmental issues and shading, and, for real-world applications, effective texture operators capable of accurately interpret image contents must be developed.

1.2.4 Image feature descriptors

Finding useful image features plays a crucial role in recognising objects. Consequently, practical image feature descriptors must take into consideration the following attributes [30]:

- Repeatability: That is, the detection must be independent of changes in the imaging conditions, such as conditions of illumination, parameters of the camera and positions of

the camera relative to the scene [31]. The repeatability rate is defined as the number of repeated points between two images with respect to the total number of detected points. The same features extracted from two images must show the same object, despite geometric and photometric transformations.

- Distinctiveness/Informativeness: Features can be distinguished and matched by observing a lot of variation in the intensity patterns underlying the detected features.
- Locality: This property helps to reduce the probability of occlusion and object deformations.
- Quantity: It is necessary to reflect the content of the image to enhance the image representation by the number of detected features.
- Accuracy: Features must be accurately localised in images.
- Efficiency: The time to detect features needs to be fast, in order to apply in real-time applications.

Knowing the criteria for useful image features enables the development of efficient methods for extracting useful features from images. Numerous studies describing the role of image feature descriptors in object recognition and discrimination have been published. The more useful features are found, the more accurately objects are detected and identified. Subsequently, the most popular and effective descriptors are discussed and described in detail.

1.2.4.1 Scale-Invariant Feature Transform (SIFT)

A local feature description algorithm-SIFT has been proposed by David Lowe [32]. The procedure of the SIFT algorithm consists of four basic steps:

- Detection of Scale-Space Extrema
- Accuracy Keypoint Localisation
- Orientation Assignment
- Keypoint Descriptor

Scale invariance plays an indispensable role in the success of the SIFT method. To obtain scale invariance, SIFT firstly applies a Difference-of-Gaussian (DoG) function to identify potential points of interest, which are invariant to scale and orientation. Subsequently, Taylor series is used to remove unstable feature points, such as low contrast, poor localised and unstable edge points, in order to improve and select good keypoints. After having a set of good points, a window region around each point is chosen to compute the gradient magnitude and direction of every neighbourhood. Next, the gradient orientation of points within the region generates an orientation histogram. The highest orientation values in the histogram are located and regarded as dominant directions of local gradients. Finally, the gradients and the direction around the keypoint are sampled. By comparing each keypoint

extracted from the test image and the set of keypoints from the training image, the best feature points for matching leaves can be found [32]. The SIFT algorithm shows good performance in scale invariance, rotation invariance and illumination changes.

Despite the excellent features of the SIFT algorithm, it typically has several drawbacks. For large scale images, the SIFT algorithm requires a very long time to calculate the descriptors, and the construction of its feature vector is complex. In addition, the SIFT algorithm is based on the use of grey-scale images only, hence, it cannot be applied for analysing colour images. Furthermore, for images with affine transformations, the accuracy of the SIFT algorithm is typically low [33]. As a result, extensive research has been carried out focusing on developing advanced methods based on the SIFT method, e.g., Principal Component Analysis combined with SIFT (PCA-SIFT) [34], global information integrated into SIFT (GSIFT) [35], colour invariance integrated into SIFT (CSIFT) [36], and affine transformation solved by using ASIFT [37]. An example of applying the original SIFT method is illustrated in Figure 2. Specifically, we captured a plant image under the field environment including barley crops and wild radish weeds as shown in Figure 2 (a). By converting the image to grayscale a radish leaf can be identified, sliced into sections, and scaled up and rotated 90° anticlockwise as shown in Figure 2 (b). Figure 2 (c) shows the wild radish leaf sliced, mapped at a different position and rotated 180° anticlockwise. Figure 2 (d) shows a barley leaf sliced, scaled up and rotated 90° clockwise. These leaves can be detected by using the SIFT algorithm despite their scales and viewpoint changes.

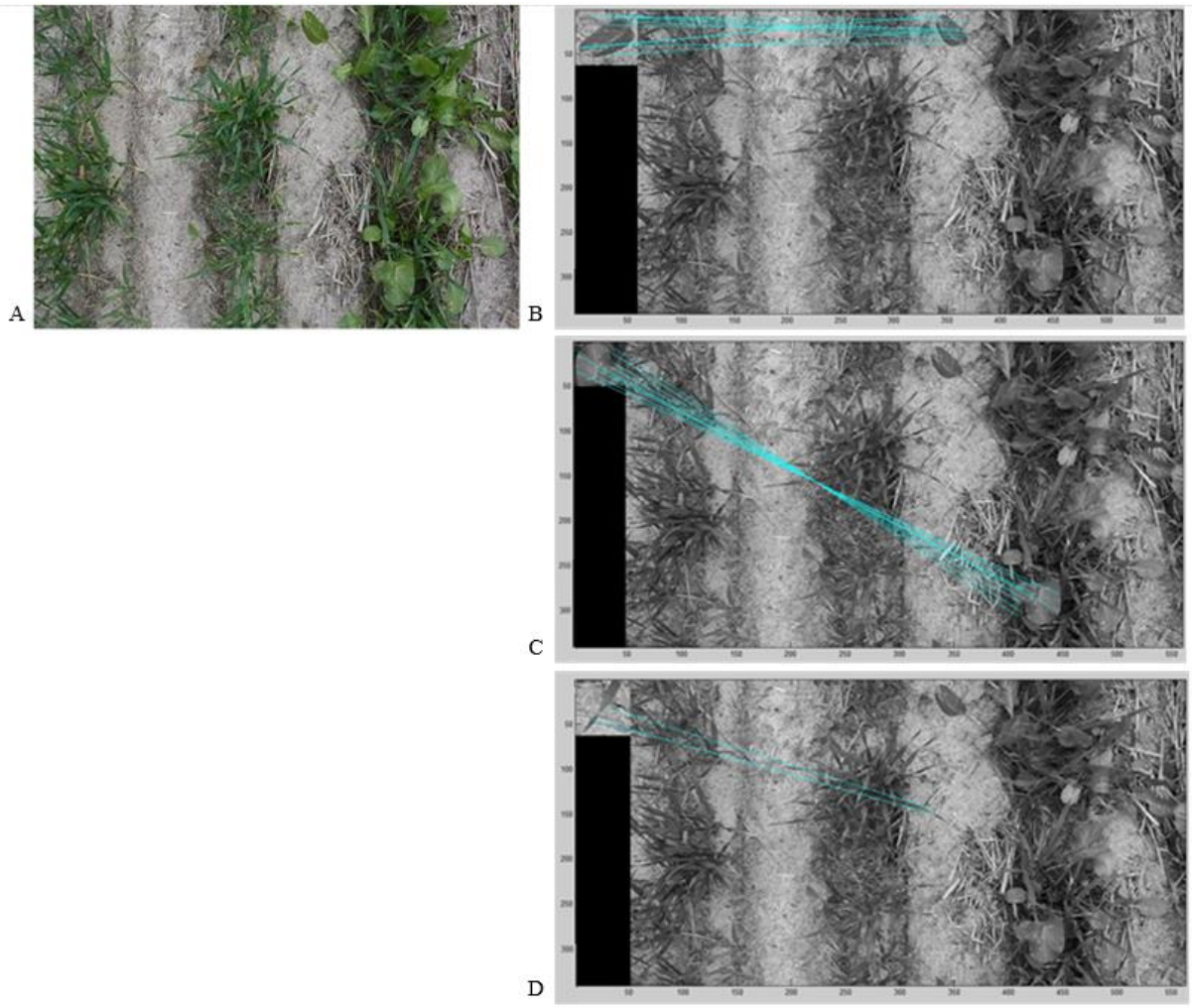


Figure 2. An example of matching barley and wild radish leaves using the SIFT algorithm, despite their scales and viewpoint changes. (A) An original plant image; (B), (C) and (D): Leaves can be detected by using the SIFT algorithm despite their scales and rotation.

1.2.4.2 Speeded-Up Robust Features (SURF)

The principle of the SURF method is similar to that of the SIFT method. However, SURF applies different methods to detect keypoint location and generate descriptors. Besides, the SURF method is proposed to address the time-consuming problem of the SIFT algorithm, which is caused by its computational complexity.

The SURF algorithm has proven to be fast and more robust when applying a quick Hessian matrix to detect interest points. Concurrently, the procedure of constructing the Gaussian pyramid in SIFT is replaced by using an integral image algorithm. In the description stage, a square region is applied around the detected interest points. For example, the SURF algorithm divides a 20×20-pixel region into 4×4 sub-regions. After that, a Haar wavelet response for each sub-region is computed, represented by a 4-

dimensional vector, each keypoint is represented by a 64-dimensional feature describing all sub-regions. To reduce illumination effects, a unit vector is generated by normalising the feature descriptor [38].

It is worth noting that the important advantage of the SURF algorithm is that its processing speed is approximately 3-4 times faster than that of the SIFT algorithm. However, the rotational invariance performance of the SURF algorithm is low, especially when 2D or 3D objects are compared and on affine invariance [39]. Figure 3 illustrates the detection of green features (300 strongest feature points) on the original plant image by using the SURF algorithm.

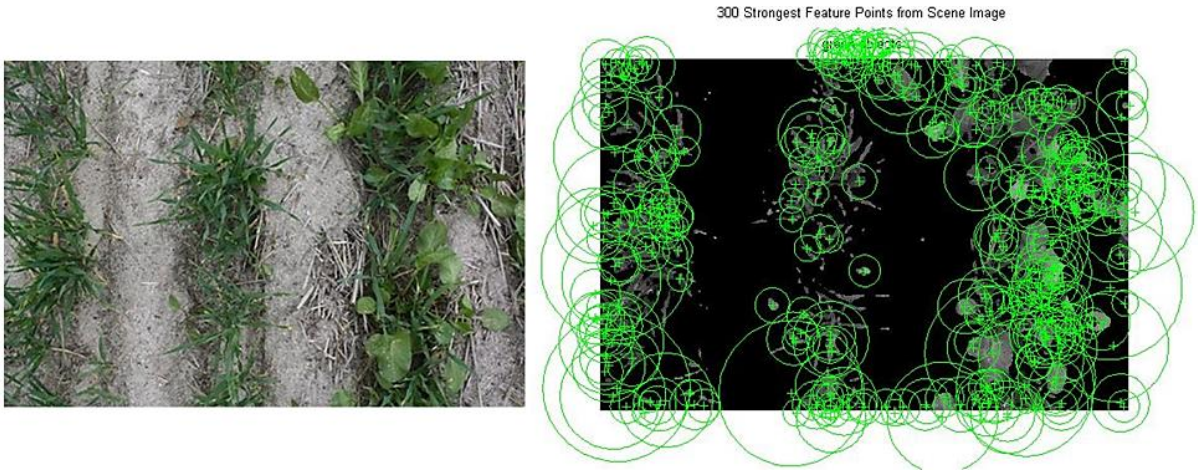


Figure 3. An example of detected features from a grey-scale plant image using the SURF algorithm.

1.2.4.3 Local Binary Patterns (LBP)

The third descriptor investigated in this review is Local Binary Pattern (LBP). This algorithm was introduced to the public in 1996 [40], and since then, it has primarily been developed to detect dominant features in images. The LBP has been regarded as one of the powerful tools for extracting good features from texture-based image analysis and classifying objects based on local image texture properties. LBP operators typically enable powerful discrimination performance in many applications, such as face recognition [41], facial expression analysis [42], and weed detection and classification [43, 44]. Moreover, the LBP method also has a computational simplicity that enables higher processing speeds to be attained for plant detection. Consequently, LBP texture operator has become a popular approach in various applications. The advantages of this algorithm include i) computation efficiency and ii) robustness to different lighting conditions, scaling, rotation, viewpoint variation, and distorted objects [45].

Numerous extended LBP methods have recently been developed to enhance the performance of LBP operators for different applications. Consequently, the original LBP operators have been improved in different aspects, including i) improving its discriminative capability; ii) enhancing its robustness; iii) selecting its neighbourhoods; iv) extending to 3-D data; and v) combining it with other approaches [46].

The key extended LBP approaches that have recently been developed are reviewed in Table 2, which gives a clear understanding of the evolution of LBP techniques.

Table 2. The development and improvement of LBP methods [46]

Purposes	**LBP extension**	**Year**	**Properties**
Enhance the discriminative ability	Improved LBP (Mean LBP)	2004,2005, 2008	- Considers the effects of central pixels and presents complete structure patterns
	Hamming LBP	2007	- Incorporates non-uniform patterns into uniform patterns
	Extended LBP	2007	- Discriminates the same local binary patterns - The drawback of this method is low feature dimensionality
	Completed LBP	2010	- Included both the sign and the magnitude information of the given local region
	Median robust extended LBP	2016 [47]	- Robustness to image noise - Strong discrimination and computational efficiency - Has no realistic and high-level applications (such as object recognition and image matching) yet
Improve the robustness	Local Ternary Patterns	2007	- Developed a new threshold to resistant to noise - No longer strictly invariant to grey-level transformation
	Soft LBP	2007	- Not invariant to monotonic grey scale changes - Causes high computational complexity

Choosing the neighbourhood	Elongated LBP	2007	- Extracts the anisotropic information and lose anisotropic information - Not invariant to rotation
	Multi-Block LBP	2007	- Captures micro-and macro-structure information
	Three/Four Patch LBP	2008	- Encodes patch type of texture information
Extending to 3D	3D LBP	2007,2008	- Extends LBP to 3D volume data
	Volume LBP (LBP-TOP)	2007	- Describes dynamic texture - Causes high dimensionality
Combining with other features	LBP and Gabor wavelet	2005,2006, 2007,2008	- Utilises the benefits of Gabor to improve the results - Disadvantages: increase time, cost and cause high dimensionality
	LBP and SIFT	2006,2009, 2010	- Combines with the advantages of SIFT method - Reduces feature vector length
	LBP Histogram Fourier	2009	- Obtains rotational invariance globally for the whole region

1.2.5 Classification

The last stage of image processing is classification. With regards to the classification of plant images, there are different machine learning methods such as Naive Bayes Classifier, Support Vector Machine (SVM) [48-50], Artificial Neural Networks (ANN) [51], area thresholding [4], Fuzzy measure [52], Nearest Neighbour [53, 54], Decision Trees [55], and Random Forest [56]. For a classifier to attain a good performance, sufficient data needs to be collected and the training performance analysed.

Furthermore, the extracted features represent the information content of the plant images, so the classification accuracy rate relies on the careful selection of the applied approaches.

To have a better understanding of feature extraction and classification methods, their benefits and limitations, a study on several popular techniques in agricultural applications is presented in Appendix 7.2. In addition, several factors affect the weed/crop discrimination process, namely [23]:

- *Lighting conditions*: Poor illumination in cloudy, overcast or sunny days has an impact on plant detection. For example, when leaves are under strong light intensities, captured images from these leaves exhibit specular reflection and the leaves may also turn into another colour. Consequently, since the dominant colour of a leaf is green, it is hard to segment green colour.
- *Shadow*: Plants or other objects can cause shadowing effect on sunny days.
- *Temperature*: Typically, the temperature of daylight is variable, resulting in changes of the colour of leaves.
- *Occlusion*: In crop field, crops might sometimes be partially occluded by weeds or other objects and vice versa.
- *Shape and rotation variation*: Weeds are usually transformed or distorted in reality.
- *Complex background*: Images contain stones, broom grasses, water, etc.

Typically, plants are segmented to remove the soil background before going through the process of feature extraction. However, inevitable illumination variation in outdoor conditions is a key factor in determining the ability to accurately detect weeds from crops. The other challenge impacting on weed detection is related to the computational efficiency of the weed detection algorithm. It is important to note that it is typically difficult for an imaging-based plant identification sensor to analyse crops and weeds with different pose angles. Therefore, the image processing technique must have the ability to detect rotated images in the training data set. In addition, training large datasets is also one of the big challenges. It is crucial for attaining acceptable accuracy levels.

After reviewing the techniques and the obstacles associated with weed detection, the potential performance of two Machine Learning (ML) models including SVM and ANNs has been presented in Figure 4. Therefore, LBP techniques in conjunction with the most popular SVM method and Artificial Neural Networks (ANNs) [57] seems to be the most viable option for weed detection, which will be selected for investigation and development in this thesis. This selection is in agreement with recommendations stating noteworthy combination of these approaches for the classification and detection of crops and weeds [43, 44, 58, 59].

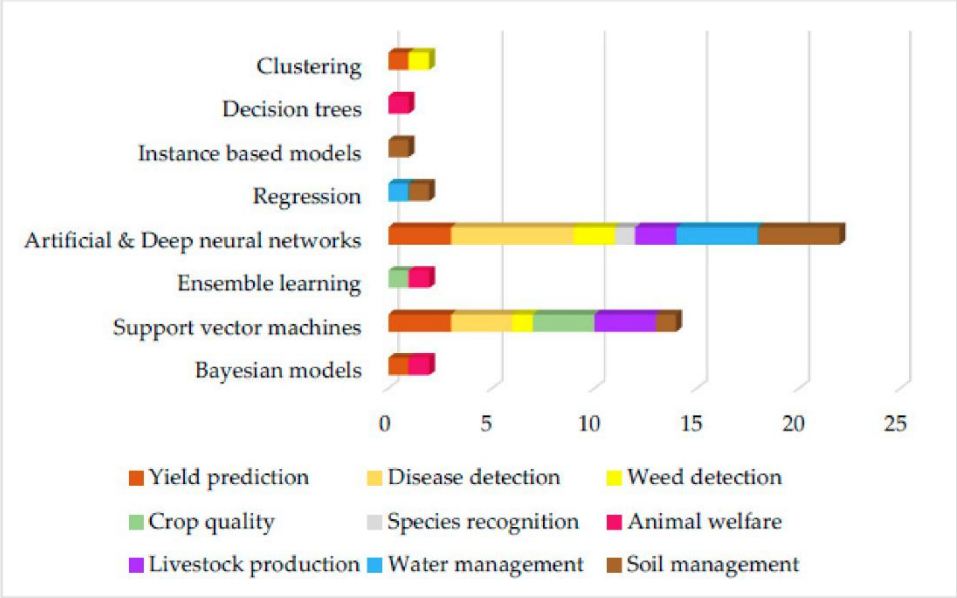


Figure 4. Machine learning models in agricultural areas [57].

1.3 Project Aims

The purpose of this research project is to investigate efficient techniques for the detection and classification of morphologically similar crops and weeds in images under complex field environments and occlusions in real time. In recent years, a variety of projects have addressed automated recognition of weeds by using cameras in order to develop weed sensing systems for controlling weeds more efficiently and intelligently. This leads to demands on automatically analysing plant images under uncontrolled field conditions. Crops and weeds with visual similarities at four different growth stages as presented in Figure 5 poses a challenge to precisely classify and detect them under complex field conditions in real time. Consequently, advanced algorithms have been developed to automatically detect weeds which becomes a promising potential in precision weed control as well as in precision agriculture. While a variety of local descriptors have been proposed for plant feature extraction and classification, including Scale Invariant Feature Transform (SIFT) [32], Speeded Up Robust Features (SURF) [38], Histogram of Oriented Gradients (HOG) [60], Local Binary Patterns (LBP), we particularly adopt the LBP technique for several reasons, namely:

1. The LBP technique was first introduced by Ojala et al. in 1996 [40]. It is considered a very flexible and robust method to monotonic grey-level changes, rotation invariance. It is computationally less complex than the SIFT or SURF methods and exhibits high discriminative capability [45, 61].
2. Due to its computational simplicity, the LBP methods have the ability to analyse images in challenging real-time settings [62].

3. The LBP technique, which has been investigated and developed for almost 20 years [63], is also a powerful descriptor for the representation of local features in images and this method.
4. The LBP techniques have been extensively exploited in many applications, such as face image analysis and face recognition [41, 64, 65], face expressions [66, 67], texture classification [45, 68, 69], and motion analysis [70, 71]. However, most developed LBP techniques focused on facial image analysis and face recognition, and too little attention has been devoted towards crop/weed discrimination. In other words, the development and optimisation of LBP methods for discriminating crops and weeds have not been thoroughly investigated [43, 44].

The objective of this work is to develop efficient LBP-based algorithms for real-time automatic crop/weed discrimination and detection by using an embedded target hardware platform. Results will be validated by using a test rig already installed in one of ESRI's laboratories[1] and by a weed sensing system under complex field environments.

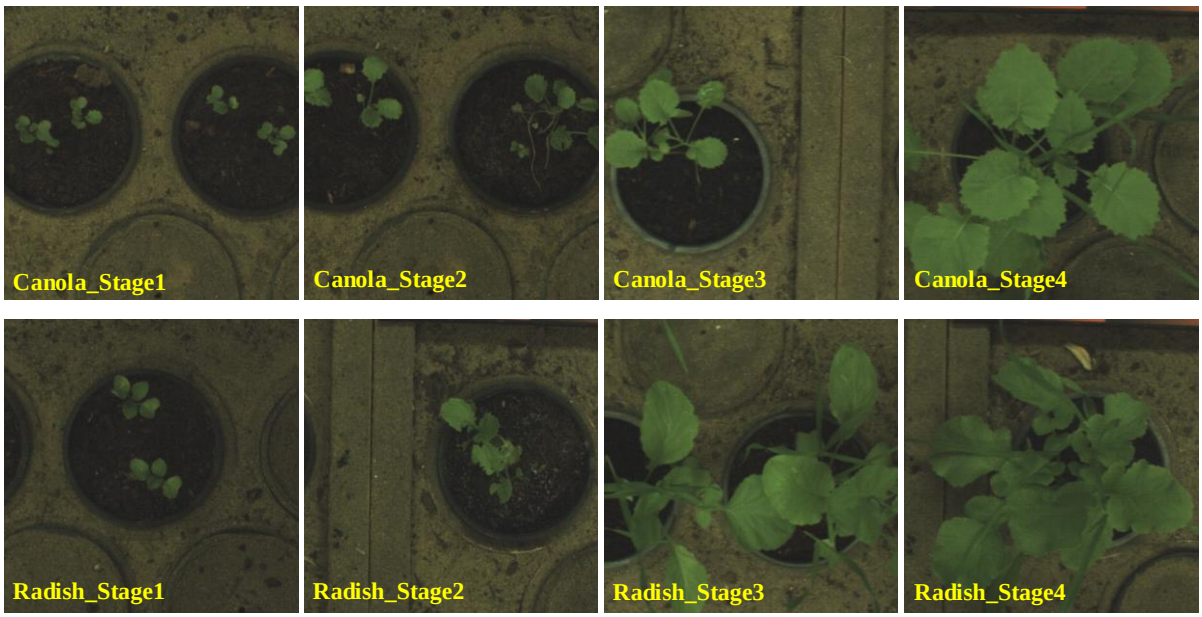


Figure 5. Canola (crop) and wild radish (weed) collected by the Testbed system have many visual similarities at every growth stage.

1.4 Contribution

This thesis is to explore and propose methods for discriminating between crop and weed species. Initially, plant datasets were collected by using the Testbed system with the main component as shown in Figure 1 and Figure 6.

- Bccr-segset dataset (published online) [72]

[1] Electron Science Research Institute: http://www.ecu.edu.au/schools/science/research-activity/electron-science-research-institute/overview

- Can-rad dataset (published online) [73]
- Mixed-plants dataset (published online) [74]

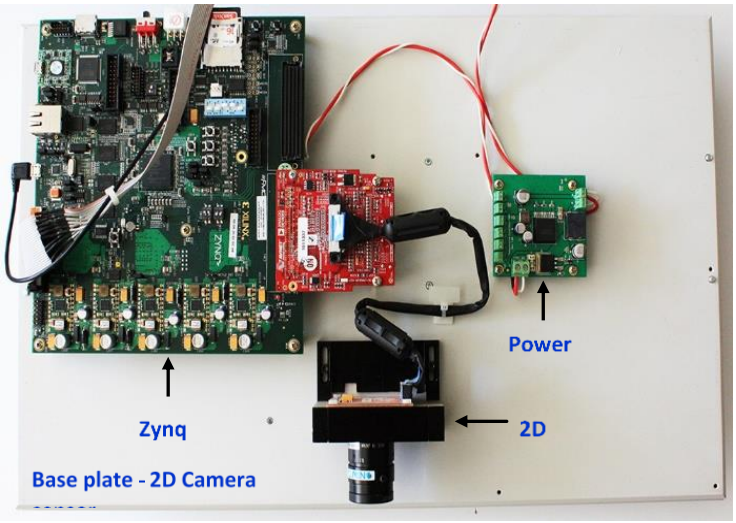


Figure 6. A Xilinx Zynq ZC702 development board with a VITA 2000 camera sensor installed in the Testbed system to capture plant images

Then we captured plant images under different weather conditions by using a weed sensing trolley as shown in Figure 7.

- FT_BRC (published online) [75]

Figure 7. A weed sensing trolley with a Zynq ZC702 development board

Lastly, we collected plant images from the field by using an integrated weed sensing system as can be seen in Figure 8. This system was designed to combine spectral reflectance and digital images in order to optimize the potential herbicide savings and accurate detections in real time.

- Fieldtrip_can_weeds dataset (published online) [76]

Figure 8. An integrated weed sensing system with the combination of multispectral and spatial sensors in the field.

The objective of this thesis is to classify and detect crops/weeds with visual similarities at different growth stages and under various environments such as weather conditions and occlusions. To the best of our knowledge, camera-based weed identification has been an attractive research topic for many years, but it has not achieved widespread adoption in agriculture. Therefore, robust algorithms, with fast execution time, size invariance and high discrimination accuracy, are developed and contributed to meet practical working requirements including real-time deployment and detections in the complicated environments. The performance of our robust algorithm has been validated through all our plant datasets and a comparison with DCNNs.

1.5 Organisation of the Dissertation

This thesis consists of seven chapters as follows:

Chapter 1 provides an introduction to the current knowledge about weeds and their detrimental effects on cereal crops. A literature review summarises knowledge about previous methods for analysing plant images including pre-processing, segmentation, feature extraction and classification. Through this literature, current and promising techniques are identified to further develop in detecting and classify crops/weeds with similar appearance in the field. Next, the rationale, aims and contributions of this project are presented in this chapter.

Chapter 2 motivates the promising potential of LBP-based method in weed recognition by combining various LBP operators with different number of neighbours and radius. This chapter has already been published.

Chapter 3 provides details on a novel method to address the limitations of the previous LBP-based method in Chapter 2. This new method enhances the accurate classification and detection of crops and weeds with similar morphology under simulated field conditions This chapter has already been published.

Chapter 4 provides the comparison of the performance of our proposed method and deep convolutional neural networks (DCNNs). Another experiment is conducted to compare execution time (including training time and testing time) of these methods in a weed detection task. The paper submission on Sensors Journal is currently under review.

Chapter 5 provides additional work in detecting different types of weeds in barley fields with cloudy, windy and shadow weather conditions by fine tuning different Faster-RCNN models to achieve the high performance in automatic weed detection under complex field environments. The manuscript is submitted to Plant Biology Journal.

Chapter 6 provides a synthesis of the findings and contributions of our study as shown in this chapter. In addition, we also discuss the further development and directions of our advanced algorithms in the future work.

1.6 References

[1] A. Tellaeche, X. P. BurgosArtizzu, G. Pajares, A. Ribeiro, and C. Fernández-Quintanilla, "A new vision-based approach to differential spraying in precision agriculture," *computers and electronics in agriculture*, vol. 60, no. 2, pp. 144-155, 2008.

[2] F. Liu and N. V. O'Connell, "Off-site movement of surface-applied simazine from a citrus orchard as affected by irrigation incorporation," *Weed science*, vol. 50, no. 5, pp. 672-676, 2002.

[3] K. H. Ghazali, M. M. Mustafa, and A. Hussain, "Machine vision system for automatic weeding strategy using image processing technique," *American-Eurasian Journal of Agricultural and Environmental Science*, vol. 3, no. 3, pp. 451-458, 2008.

[4] S. H. Hlaing and A. S. Khaing, "Weed and crop segmentation and classification using area thresholding," *IJRET*, vol. 3, pp. 375-382, 2014.

[5] K. Porker and R. Wheeler, "Getting the best from barley–agronomy and management," *South Australia*, 2014.

[6] T. C. Sam Trengove. "New technology for improved herbicide use efficiency." Grains Research and Development Corporation (GRDC). https://grdc.com.au/Research-and-

Development/GRDC-Update-Papers/2015/02/New-technology-for-improved-herbicide-use-efficiency (accessed 21/7/2016.

[7] T. Heisel, S. Christensen, A. Walter, and J. Stafford, "Whole-field experiments with site-specific weed management," in *Precision agriculture'99, Part 2. Papers presented at the 2nd European Conference on Precision Agriculture, Odense, Denmark, 11-15 July 1999.*, 1999: Sheffield Academic Press, pp. 759-768.

[8] (2016). *Western Australian wheat industry*. [Online] Available: https://www.agric.wa.gov.au/grains-research-development/western-australian-wheat-industry

[9] S. S. a. G. M. Rick Graham, NSW DPI Tamworth. "Barley agronomy and varieties." https://grdc.com.au/Research-and-Development/GRDC-Update-Papers/2015/03/Barley-agronomy-and-varieties (accessed 18/7/2016.

[10] Agrifutures. "Canola." https://www.agrifutures.com.au/farm-diversity/canola/ (accessed.

[11] Aegic. "Australian canola." https://www.aegic.org.au/wp-content/uploads/2018/02/AEGIC-Grain-Note-canola_LR.pdf (accessed.

[12] M. J. Owen, N. J. Martinez, and S. B. Powles, "Multiple herbicide-resistant wild radish (Raphanus raphanistrum) populations dominate Western Australian cropping fields," *Crop and Pasture Science*, vol. 66, no. 10, pp. 1079-1085, 2015.

[13] M. J. Walsh, S. B. Powles, B. R. Beard, B. T. Parkin, and S. A. Porter, "Multiple-herbicide resistance across four modes of action in wild radish (Raphanus raphanistrum)," *Weed science*, vol. 52, no. 1, pp. 8-13, 2004.

[14] R. E. Blackshaw, D. Lemerle, R. Mailer, and K. R. Young, "Influence of wild radish on yield and quality of canola," *Weed Science*, vol. 50, no. 3, pp. 344-349, 2002.

[15] K. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, "Machine learning in agriculture: A review," *Sensors*, vol. 18, no. 8, p. 2674, 2018.

[16] C. M. Onyango and J. Marchant, "Segmentation of row crop plants from weeds using colour and morphology," *Computers and electronics in agriculture*, vol. 39, no. 3, pp. 141-155, 2003.

[17] H. T. Søgaard, "Weed classification by active shape models," *Biosystems engineering*, vol. 91, no. 3, pp. 271-281, 2005.

[18] I. Schuster, H. Nordmeyer, and T. Rath, "Comparison of vision-based and manual weed mapping in sugar beet," *Biosystems engineering,* vol. 98, no. 1, pp. 17-25, 2007.

[19] G. Jeon, "Color image enhancement by histogram equalization in heterogeneous color space," *Int. J. Multimedia Ubiquitous Eng,* vol. 9, no. 7, pp. 309-318, 2014.

[20] G. Pajares, J. J. Ruz, and J. M. de la Cruz, "Performance analysis of homomorphic systems for image change detection," in *Iberian Conference on Pattern Recognition and Image Analysis*, 2005: Springer, pp. 563-570.

[21] L. I. Voicu, H. R. Myler, and A. R. Weeks, "Practical considerations on color image enhancement using homomorphic filtering," *Journal of Electronic Imaging,* vol. 6, no. 1, pp. 108-114, 1997.

[22] S. Sankaran, A. Mishra, R. Ehsani, and C. Davis, "A review of advanced techniques for detecting plant diseases," *Computers and Electronics in Agriculture,* vol. 72, no. 1, pp. 1-13, 2010.

[23] E. Hamuda, M. Glavin, and E. Jones, "A survey of image processing techniques for plant extraction and segmentation in the field," *Computers and Electronics in Agriculture,* vol. 125, pp. 184-199, 2016.

[24] M. U. Ahmed, T. Chanwimalueang, S. Thayyil, and D. P. Mandic, "A Multivariate Multiscale Fuzzy Entropy Algorithm with Application to Uterine EMG Complexity Analysis," *Entropy,* vol. 19, no. 1, p. 2, 2016.

[25] S. A. Medjahed, "A Comparative Study of Feature Extraction Methods in Images Classification," *International Journal of Image, Graphics and Signal Processing,* vol. 7, no. 3, p. 16, 2015.

[26] M. Yang, K. Kpalma, and J. Ronsin, "A survey of shape feature extraction techniques," ed: IN-TECH, 2008.

[27] L. Tang, L. Tian, B. Steward, and J. Reid, "Texture-based weed classification using Gabor wavelets and neural network for real-time selective herbicide applications," *Urbana,* vol. 51, p. 61801, 1999.

[28] D. ping Tian, "A review on image feature extraction and representation techniques," *International Journal of Multimedia and Ubiquitous Engineering,* vol. 8, no. 4, pp. 385-396, 2013.

[29] A. Materka and M. Strzelecki, "Texture analysis methods–a review," *Technical university of lodz, institute of electronics, COST B11 report, Brussels*, vol. 10, no. 1.97, p. 4968, 1998.

[30] T. Tuytelaars and K. Mikolajczyk, "Local invariant feature detectors: a survey," *Foundations and trends® in computer graphics and vision*, vol. 3, no. 3, pp. 177-280, 2008.

[31] C. Schmid, R. Mohr, and C. Bauckhage, "Evaluation of interest point detectors," *International Journal of computer vision*, vol. 37, no. 2, pp. 151-172, 2000.

[32] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, no. 2, pp. 91-110, 2004.

[33] J. Wu, Z. Cui, V. S. Sheng, P. Zhao, D. Su, and S. Gong, "A Comparative Study of SIFT and its Variants," *Measurement Science Review*, vol. 13, no. 3, pp. 122-131, 2013.

[34] Y. Ke and R. Sukthankar, "PCA-SIFT: A more distinctive representation for local image descriptors," in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, 2004, vol. 2: IEEE, pp. II-II.

[35] E. N. Mortensen, H. Deng, and L. Shapiro, "A SIFT descriptor with global context," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, 2005, vol. 1: IEEE, pp. 184-190.

[36] A. E. Abdel-Hakim and A. A. Farag, "CSIFT: A SIFT descriptor with color invariant characteristics," in *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, 2006, vol. 2: IEEE, pp. 1978-1983.

[37] J.-M. Morel and G. Yu, "ASIFT: A new framework for fully affine invariant image comparison," *SIAM Journal on Imaging Sciences*, vol. 2, no. 2, pp. 438-469, 2009.

[38] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Computer vision and image understanding*, vol. 110, no. 3, pp. 346-359, 2008.

[39] K. Xia, Y. Wu, X. Ren, and Y. Jin, "Research in clustering algorithm for diseases analysis," *Journal of Networks*, vol. 8, no. 7, pp. 1632-1640, 2013.

[40] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern recognition*, vol. 29, no. 1, pp. 51-59, 1996.

[41] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE transactions on pattern analysis and machine intelligence,* vol. 28, no. 12, pp. 2037-2041, 2006.

[42] G. Zhao and M. Pietikainen, "Dynamic texture recognition using local binary patterns with an application to facial expressions," *IEEE transactions on pattern analysis and machine intelligence,* vol. 29, no. 6, pp. 915-928, 2007.

[43] F. Ahmed, A. H. Bari, A. Shihavuddin, H. A. Al-Mamun, and P. Kwan, "A study on local binary pattern for automated weed classification using template matching and support vector machine," in *Computational Intelligence and Informatics (CINTI), 2011 IEEE 12th International Symposium on*, 2011: IEEE, pp. 329-334.

[44] F. Ahmed, M. H. Kabir, S. Bhuyan, H. Bari, and E. Hossain, "Automated weed classification with local pattern-based texture descriptors," *Int. Arab J. Inf. Technol.,* vol. 11, no. 1, pp. 87-94, 2014.

[45] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on pattern analysis and machine intelligence,* vol. 24, no. 7, pp. 971-987, 2002.

[46] D. Huang, C. Shan, M. Ardabilian, Y. Wang, and L. Chen, "Local binary patterns and its application to facial image analysis: a survey," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews),* vol. 41, no. 6, pp. 765-781, 2011.

[47] L. Liu, S. Lao, P. W. Fieguth, Y. Guo, X. Wang, and M. Pietikäinen, "Median robust extended local binary pattern for texture classification," *IEEE Transactions on Image Processing,* vol. 25, no. 3, pp. 1368-1381, 2016.

[48] A. Bakhshipour and A. Jafari, "Evaluation of support vector machine and artificial neural networks in weed detection using shape features," *Computers and Electronics in Agriculture,* vol. 145, pp. 153-160, 2018.

[49] L. Wu and Y. Wen, "Weed/corn seedling recognition by support vector machine using texture features," *African Journal of Agricultural Research,* vol. 4, no. 9, pp. 840-846, 2009.

[50] T. Rumpf, A.-K. Mahlein, U. Steiner, E.-C. Oerke, H.-W. Dehne, and L. Plümer, "Early detection and classification of plant diseases with Support Vector Machines based on hyperspectral reflectance," *Computers and Electronics in Agriculture,* vol. 74, no. 1, pp. 91-99, 2010.

[51] C. Gliever and D. C. Slaughter, "Crop versus weed recognition with artificial neural networks," in *ASAE Meeting Paper*, 2001, no. 01-3104.

[52] M. T. Lamata and S. Moral, "Classification of fuzzy measures," *Fuzzy Sets and Systems*, vol. 33, no. 2, pp. 243-253, 1989.

[53] E.-H. S. Han, G. Karypis, and V. Kumar, "Text categorization using weight adjusted k-nearest neighbor classification," in *Pacific-asia conference on knowledge discovery and data mining*, 2001: Springer, pp. 53-65.

[54] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE transactions on information theory*, vol. 13, no. 1, pp. 21-27, 1967.

[55] A. Ghulam, I. Porton, and K. Freeman, "Detecting subcanopy invasive plant species in tropical rainforest by integrating optical and microwave (InSAR/PolInSAR) remote sensing data, and a decision tree algorithm," *ISPRS journal of photogrammetry and remote sensing*, vol. 88, pp. 174-192, 2014.

[56] J. Gao, D. Nuyttens, P. Lootens, Y. He, and J. G. Pieters, "Recognising weeds in a maize crop using a random forest machine-learning algorithm and near-infrared snapshot mosaic hyperspectral imagery," *Biosystems Engineering*, vol. 170, pp. 39-50, 2018.

[57] K. G. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, "Machine learning in agriculture: A review," *Sensors*, vol. 18, no. 8, p. 2674, 2018.

[58] F. Ahmed, H. A. Al-Mamun, A. H. Bari, E. Hossain, and P. Kwan, "Classification of crops and weeds from digital images: A support vector machine approach," *Crop Protection*, vol. 40, pp. 98-104, 2012.

[59] M. F. Kazerouni, J. Schlemperz, and K.-D. Kuhnert, "Efficient Modern Description Methods by Using SURF Algorithm for Recognition of Plant Species," *Advances in Image and Video Processing*, vol. 3, no. 2, p. 10, 2015.

[60] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, 2005, vol. 1: IEEE, pp. 886-893.

[61] O. Lahdenoja, J. Poikonen, and M. Laiho, "Towards understanding the formation of uniform local binary patterns," *ISRN Machine Vision*, 2013.

[62] S. Happy, A. George, and A. Routray, "A real time facial expression classification system using Local Binary Patterns," in *Intelligent Human Computer Interaction (IHCI), 4th International Conference*, 2012: IEEE, pp. 1-5.

[63] M. Pietikäinen and G. Zhao, "Two decades of local binary patterns: A survey," *Advances in Independent Component Analysis and Learning Machines*, pp. 175-210, 2015.

[64] H. Jin, Q. Liu, H. Lu, and X. Tong, "Face detection using improved LBP under Bayesian framework," in *Image and Graphics (ICIG'04), Third International Conference*, 2004: IEEE, pp. 306-309.

[65] W. Louis and K. N. Plataniotis, "Co-occurrence of local binary patterns features for frontal face detection in surveillance applications," *EURASIP Journal on Image and Video Processing*, vol. 2011, no. 1, p. 745487, 2010.

[66] G. Zhao and M. Pietikainen, "Dynamic texture recognition using local binary patterns with an application to facial expressions," *IEEE transactions on pattern analysis and machine intelligence*, vol. 29, no. 6, 2007.

[67] C. Shan, S. Gong, and P. W. McOwan, "Facial expression recognition based on local binary patterns: A comprehensive study," *Image and Vision Computing*, vol. 27, no. 6, pp. 803-816, 2009.

[68] Z. Guo, L. Zhang, and D. Zhang, "A completed modeling of local binary pattern operator for texture classification," *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1657-1663, 2010.

[69] S. Liao, M. W. Law, and A. C. Chung, "Dominant local binary patterns for texture classification," *IEEE transactions on image processing*, vol. 18, no. 5, pp. 1107-1118, 2009.

[70] M. Heikkila and M. Pietikainen, "A texture-based method for modeling the background and detecting moving objects," *IEEE transactions on pattern analysis and machine intelligence*, vol. 28, no. 4, pp. 657-662, 2006.

[71] V. Kellokumpu, G. Zhao, and M. Pietikäinen, "Human activity recognition using a dynamic texture based method," in *BMVC*, 2008, vol. 1, p. 2.

[72] V. N. T. Le. Bccr-segset dataset [Online] Available: https://drive.google.com/file/d/1Oc0xncO_tvCNnSioy2IKTuKi6bEwgzkL/view?usp=sharing

[73] V. N. T. Le. Can-rad dataset [Online] Available: https://drive.google.com/file/d/1azyjPjGR8AP_BUvzBYPho2AZckB-c7Ao/view?usp=sharing

[74] V. N. T. Le. Mixed-plants dataset [Online] Available: https://drive.google.com/file/d/12dIYUbxJM2Envcc8tRrVLq0Y2Rh_IaT8/view?usp=sharing

[75] V. N. T. Le. FT_BRC dataset [Online] Available: https://drive.google.com/open?id=179WA2Ky6aWh4zU8eb_MuQ_F7jWkc4Mj_

[76] V. N. T. Le. Fieldtrip_can_weeds dataset [Online] Available: https://drive.google.com/file/d/1mxH5TAK2p8--fzg9PAzxAtQNb6O5mSX4/view?usp=sharing

Chapter 2 – Effective plant discrimination based on the combination of local binary pattern operators and multiclass support vector machine methods

This chapter was published as an article in the journal of Information Processing in Agriculture, 2019, vol. 6, pp 116-131. DOI: 10.1016/j.inpa.2018.08.002 This article appears as it does in print, with the exception of minor changes to the layout, number formats, font size and font style, which was implemented to maintain consistency in the formatting of this thesis.

2.1 Abstract

Accurate crop and weed discrimination play a critical role in addressing the challenges of weed management in agriculture. The use of herbicides is currently the most common approach to weed control. However, herbicide resistant plants have long been recognised as a major concern due to the excessive use of herbicides. Effective weed detection techniques can reduce the cost of weed management and improve crop quality and yield. A computationally efficient and robust plant classification algorithm is developed and applied to the classification of three crops: Brassica napus (canola), Zea mays (maize/corn), and radish. The developed algorithm is based on the combination of Local Binary Pattern (LBP) operators, for the extraction of crop leaf textural features and Support Vector Machine (SVM) method, for multiclass plant classification. This paper presents the first investigation of the accuracy of the combined LBP algorithms, trained using a large dataset of canola, radish and corn leaf images captured by a testing facility under simulated field conditions. The dataset has four subclasses, background, canola, corn, and radish, with 24,000 images used for training and 6000 images, for validation. The dataset is referred herein as "bccr-segset" and published online. In each subclass, plant images are collected at four crop growth stages. Experimentally, the algorithm demonstrates plant classification accuracy as high as 91.85%, for the four classes.

Keywords: Plant discrimination, Classification, LBP, PCA and SVM

2.2 Introduction

Weed infestation has always been a critical issue that limits the productivity of farms and the yield of crops. The ability to accurately discriminate weeds from crops in real-time will advance precision crop and weed management, whereby weeds in a field are prevented from competing for light water and nutrients required by the crops. Blanket herbicide spraying is currently the most common practice used for weed control. The worthwhile objective of precision weed control is to bring down the cost of weed

management. To enhance the longevity of the current range of agricultural chemicals, it is important to deter the increase in herbicide resistant weeds.

Cereal crops such as wheat, rice, maize (corn), oats, barley, rye and sorghum, represent a large portion of the crops grown worldwide [1]. Hence, detecting dominant weeds in cereal crop fields and controlling them in real-time will enable effective site-specific weed management, resulting in substantial economic benefits [2]. A variety of weed detection approaches based on feature extraction have been proposed, these include shape-based analysis [3, 4], colour-based analysis [5], texture-based image analysis [6, 7] and spectral analysis [8-10]. However, the accuracy of the above-mentioned approaches has been limited due to the complexity of the field environment, the wide variety of species and the morphological variation of plants at various growth stages.

Numerous approaches to the discrimination of crops and weeds have been reported. Over the last two decades, spectral techniques based on the calculation of the Normalised Difference Vegetation Indices (NDVIs) [11, 12] have been proposed for distinguishing between plant species. However, these spectral techniques have some limitations, especially when the spectral characteristics of weeds and crops are similar over the operational wavelengths. In addition, in typical farming field conditions, the wind, shadowing, and background illumination may change the spectral features of plants, thus reducing the discrimination accuracy of NDVI-based weed sensors [13, 14]. The limitations of such spectral-reflectance sensors have triggered research on the development of spatial sensors, based on the use of image processing techniques, for the classification of plant species and detection of weeds in real time.

A variety of feature extraction operators have been proposed for detecting robust features in images, based on the Scale Invariant Feature Transform (SIFT) [15], Speeded Up Robust Features (SURF) [16], the Histogram of Oriented Gradients (HOG) [17], LBP, Gabor filters [18] to name a few. In this paper, we adopt the LBP technique for plant feature extraction for several reasons. Firstly, LBP method is very flexible and robust to monotonic grey-level transformation, illumination, scaling, viewpoint, and rotation variance [6]. Secondly, the LBP method enables image analysis in challenging real-time settings, due to computational simplicity [19]. In fact, the LPB is computationally less complex than its SIFT or SURF counterparts, exhibiting high discrimination capability [20]. Finally, the LBP has exhibited superior performance in several applications, such as face recognition [21-23], face expression analysis [24, 25], texture classification [6, 26, 27], and motion analysis [28, 29].

The optimization of LBP methods for discriminating crops and weeds has proved difficult in special scenarios [30, 31]. In particular, Ahmed et al. used 400 colour images (taken at an angle of 45 degrees with respect to the ground) in natural lighting conditions, 200 samples were of broadleaves and 200 of grass weeds [31]. From observation the number of images and the types of plants collected in the dataset is limited. Reduced accuracy was attained in the field due to the relatively small number of plant images and viewpoints, variable lighting conditions and change in plant aspect ratios for each growth stage.

Furthermore, several extended LBP methods have used common and published texture databases including Outex [32], Brodatz [33], UIUC [34], UMD [35] and CUReT [36] to validate, evaluate or compare classification results [37]. However, databases for the detection and classification of plant textures have not been commonly published.

Typically, after extracting good features from plant images, the next process is to classify plant species. Previous research has mainly focused on the use of artificial neural networks (ANN) [38, 39], Bayesian classifiers [40-42], k-nearest neighbour (KNN) classifiers [43], discriminant analysis [44, 45] and SVM classifiers [46-49] for weed identification and discrimination. According to [50-52], SVM has been regarded as a robust technique for difficult classification tasks. This paper focuses on applying the LBP method in conjunction with SVM for plant feature extraction and classification of various plants images.

The main contributions of the work in this paper are summarized as follows:

- A large plant dataset was captured by using a Testbed with around 30000 plant images. This large dataset contains four classes, a variety of plant images at four defined growth stages, with rotation, scale and viewpoint variance in order to evaluate the robustness and performance of the method.
- Due to the low dimensionality of the plant representation and the low tolerance to illumination changes, LBP was especially investigated with different parameters for plant detection and combined with SVM-based classification to investigate its capability to operate in real-time.

The paper consists of four sections and is structured as follows. Section 2.2 explains why weed detection plays a crucial role in agricultural precision. It also introduces the selected method and presents a brief review of LBP analysis, together with the advantages and disadvantages of the proposed weed detection and classification approach. Section 2.3 describes the principles of the LBP technique and the rationale of combining LBP operators with SVM for the extraction of key features from plant images and the classification of different types of plants in a large dataset. Performance measures for classification and data collection are also presented in Section 2.3. A detailed flowchart for training and validating the dataset are covered in Section 2.4. Results are presented in Section 2.5, indicating that performance is best achieved by using segmented images (i.e. working with the green plant material extracted from images and converting it to greyscale). Based on these initial results, the data set “bccr-segset” is collected in the form of greyscale segmented images. Then, the classification accuracy and F1 scores of groups with different plant classes are discussed in detail, illustrating the effectiveness of the methodology in regard to plant detection and classification. Finally, conclusions and future work are discussed in Section 2.6.

2.3 Materials

This section describes the background and performance metrics that lead to the generation of the results shown in Section 2.5. The theoretical concept and principle of the selected methods in segmentation, feature extraction and classification processes are detailed in Sections 2.3.1-2.3.3 Classification accuracy and F1 scores measures are presented in Section 2.3.4. Data collection is explained in detail in Section 2.3.5.

2.3.1 Segmentation

Image segmentation refers to the process of partitioning an image into multiple segments or regions. In terms of weed detection, this process is based on the segmentation of green plant material (crops and weeds) and non-green background areas (i.e. soil and residues). Removing the background areas of the images enables better plant feature extraction and classification.

In this paper, the ExG-ExR (Excess Green minus Excess Red Indices) method is used to segment green plant regions with ExG−ExR = 3×g − 2.4×r − b (g: green, r: red and b: blue). This colour index-based method has exhibited adequate robustness and high accuracy compared to other methods, such as ExG (Excess Green Index)+Otsu and NDI (Normalised Difference vegetation Index)+Otsu under greenhouse field lighting conditions and natural field lighting conditions [53]. Typically, the ExG component extracts green information, while the ExR component eliminates the background noise [54]. An example of image segmentation is illustrated in Figure 1, which shows canola, corn and radish plants that were randomly arranged along the testing trays of a test bed. The vegetation indices of the RGB plant image were first extracted by applying the ExG-ExR approach, then, the image was converted to a greyscale image before applying feature extraction and classification.

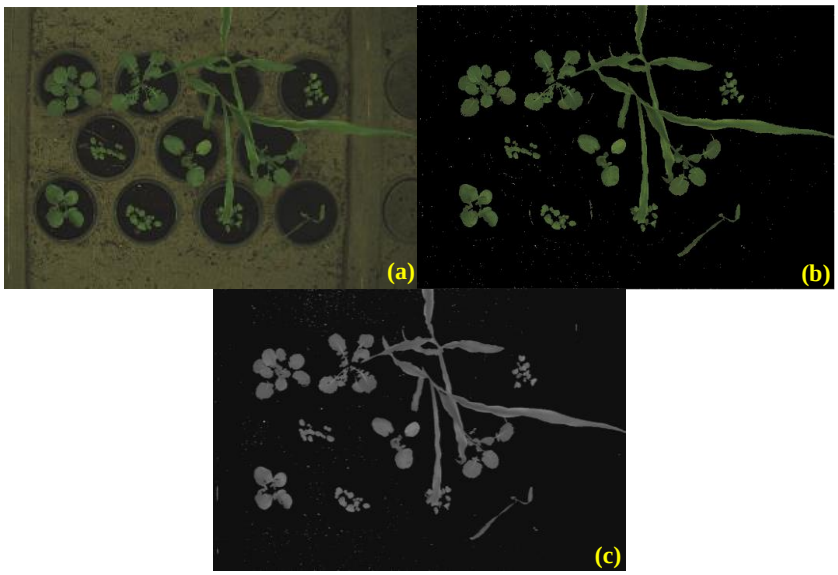


Figure 1. Images of canola, corn and radish: (a) full RGB image, (b) image with extracted green material (plants) by applying the ExG-ExR method, (c) greyscale image of (b).

2.3.2 Local Binary Pattern Operators

To better understand how LBP is applied for weed detection, a brief background on LBP is presented. The LBP method has been regarded as a powerful tool for extracting robust features from texture-based image analysis and classifying objects based on local image texture properties. The first LBP algorithm was reported in 1996 [55], since then, various LBP algorithms have been developed to primarily detect textures or objects in images. A very small local neighbourhood of a pixel is used to calculate a feature vector. Basically, the LBP operator labels the pixels of an image by thresholding the local structure around each pixel and considering the result as a binary number. Figure 2 illustrates an example of computing LBP in a 3×3 neighbourhood by comparing the intensities of the eight neighbours around each pixel with the intensity of the centre pixel. When the intensity of the centre pixel is greater than that of a neighbour, it is considered to be '0', otherwise '1'. A binary chain is obtained by combining every single binary code in a clockwise direction. For Figure 2, the binary code is 11110001, or 241 in decimal [55]. The binary number is used to build a histogram, which can be regarded as representing the texture of an image.

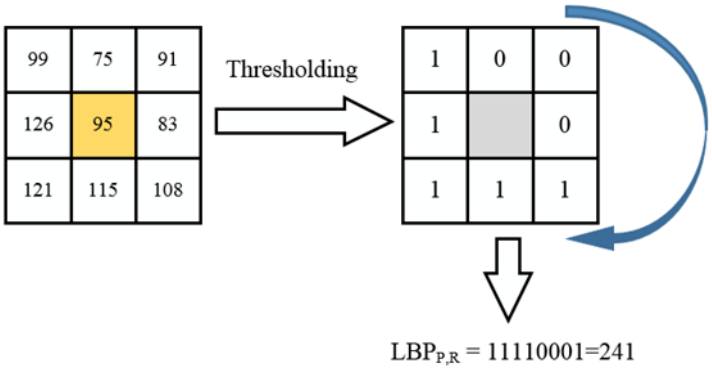


Figure 2. An example of computing LBP codes. A binary code is obtained by comparing the intensity of the centre pixel with those of the eight neighbours in a 3×3 neighbourhood.

The main limitation of the LBP operator presented above is that it only covers a small area of the neighbourhood. For a small 3×3 neighbourhood the LBP fails to capture dominant textural features in an image. As a result, the LBP operator was improved upon by increasing the number of pixels and the radius in the circular neighbourhood [6]. Note that it is typically more flexible and effective to improve LBP operators using textures of different scales. Generally, the value of the LBP code of a pixel (x_c, y_c) can be calculated as follows [6]:

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p \quad \text{where} \quad s(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases} \tag{1}$$

where:

g_c: is the grey value of the centre pixel.

g_p: represent the grey values of the circularly symmetric neighbourhood from $p = 0$ to $P - 1$ and $g_p = x_{P,R,p}$.

P: is the number of surrounding pixels in the circular neighbourhood with the radius R.

$s(x)$: is the thresholding step function which helps the LBP algorithm to gain illumination invariance against any monotonic transformation.

According to Eq. 1, the $LBP_{P,R}$ operator produces 2^P different output values. If the image is rotated, the grey values, g_p, of the circularly symmetric neighbourhood will move correspondingly along the perimeter of the circle. This generates a different LBP value, except for patterns with only the value '0' or '1'. In order to eliminate rotation effects, a rotation-invariant LBP is defined as follows [6]:

$$LBP_{P,R}^{ri} = \min\{ROR(LBP_{P,R}, i) \quad | \quad i = 0,1, \dots, P-1\} \tag{2}$$

where $ROR(x, i)$ performs an i-step circular bit-wise right shift on the P-bit number x.

To choose good and quality features, feature space dimensionality needs to be reduced by keeping only the rotationally-unique patterns. Accordingly, Ojala et al. named these patterns uniform patterns. The patterns denoted as $LBP_{P,R}^{u2}$ stand for the number of spatial transitions in the patterns meaning that the uniform patterns need to have two bitwise transitions from 0 to 1 or vice versa. For instance, uniform patterns with eight pixels in the circular neighbourhood, 00000000 (0 transitions), 11111111 (0 transitions), or 01110000 (2 transitions) are uniform because the parameter U that measures the uniformity has at most 2 transitions. Examples of non-uniform patterns are: 00000101 (4 transitions) and 01000101 (6 transitions). Consequently the rotation invariant uniform descriptor $LBP_{P,R}^{riu2}$ can be defined as follows [6]:

$$LBP_{P,R}^{riu2} = \begin{cases} \sum_{p=0}^{P-1} s(x_{P,R,p} - x_c), & \text{if } U(LBP_{P,R}) \leq 2 \\ P + 1, & \text{if } U(LBP_{P,R}) > 2 \end{cases} \tag{3}$$

The uniform descriptor has $P(P-1) + 3$ patterns including $P(P-1) + 2$ distinct uniform patterns and all non-uniform patterns assigned to a group $(P + 1)$. According to Ojala et al., the rotation invariant uniform descriptor has $(P + 2)$ distinct output patterns [6]. This reduces the feature space and helps increase the speed of LBP. For example, if the number of pixels in the circular neighbourhood is 8, the number of uniform patterns is 58 and the number of rotation invariant uniform patterns is 10.

2.3.3 Support Vector Machine

The final stage in the image processing is classification. There are different machine learning methods such as decision trees, SVM, neural networks, k-nearest neighbour method, and the Bayesian classifier. For a classifier to achieve good performance, sufficient data needs to be acquired and the training

performance analysed. The SVM can deal with pattern classification and eliminate over-fitting, and it is robust to noise [47, 56]. SVM was first introduced in 1992 [57]. SVM performs classification more accurately than other algorithms in many applications, especially those applications involving very high dimensional data [42, 46, 47, 58, 59]. This high performance makes the SVM classifier a preferred option for many applications, such as face recognition, weed identification and disease detection in plant leaves. Therefore, the optimal combination of the LBP descriptors and SVM classification can result in high plant discrimination accuracy. In particular, SVM generates an optimal hyper-plane that maximizes the margin between the classes.

To be a good discriminative classifier, SVM needs to use an appropriate kernel function. Due to the separation of the learning algorithm and kernel functions, kernels can be studied independently of the learning algorithm. One can design and experiment with different kernel functions without touching the underlying learning algorithm. Commonly, polynomial or Gaussian RBF (Radial Basis Function) kernels are used in most applications, depending on the types of data. In this paper, 2^{nd} order polynomials and 5-fold cross validation are used. Specifically, the training set is firstly divided into five subsets of equal size, and four parts of the data are iteratively used for training, with the remaining part of data used for testing. This cross-validation procedure helps to prevent data overfitting and subsequent loss of generalization.

2.3.4 Performance metrics for plant classification

The common way of assessing a classification algorithm is to calculate its classification accuracy, which is defined as

$$\text{Classification Accuracy } (\%) = \frac{\text{Number of correct classifications}}{\text{Total number of samples}} . 100\% \qquad (4)$$

However, in order to assess the performance of the SVM classifier for each class, confusion matrices are evaluated by computing main metrics, namely: precision, recall and F1 score, from the measured true positives, false positives, true negatives and false negatives. All parameters differentiate the correct classification of labels within different classes [60, 61]. A basic confusion matrix comprises 4 entries: True Positive (TP), False Negative (FN), False Positive (FP) and True Negative (TN). According to [61], we can calculate the average of precision, recall and F1 score for multi-class classification by firstly computing these parameters based on TP, TN, FN, and FP in each class as follows:

$$\text{Recall (class)} = \frac{\text{TP(class)}}{\text{TP(class)} + \text{FN(class)}} \qquad (5)$$

$$\text{Precision (class)} = \frac{\text{TP(class)}}{\text{TP(class)} + \text{FP(class)}} \tag{6}$$

$$\text{F1 score (class)} = \frac{2 \times \text{Precision(class)} \times \text{Recall(class)}}{\text{Precision(class)} + \text{Recall(class)}} = \frac{2\text{TP(class)}}{2\text{TP(class)} + \text{FN(class)} + \text{FP(class)}} \tag{7}$$

Precision in each class is defined as the number of correctly classified positive plant images divided by the total number of plant images in the data. Recall in each class is the ratio of the number of correctly classified positive plant images to the number of positive plant images in the data. F1 score in each class is a composite measure of precision and recall in each class.

2.3.5 Data Collection

In this study all the data was captured on a custom-built testing facility at ESRI (Electron Science Research Institute), Edith Cowan University, Australia, which is shown in Figure 3 and Figure 4. The hardware comprises a Xilinx Zynq ZC702 development platform [62] that captures HD images (1920 x 1080 pixels) at 60 frames per second using an On-Semi VITA 2000 camera sensor. The Zynq development board and camera are mounted on a moveable trolley with the camera optical axis perpendicular to the ground and move on a linear drive across the frame of the Testbed. The captured images have a spatial resolution of ≈1mm/pixel and a size of 228×228 pixels, which is down-sampled by a factor of 2 from a size of 456×456 pixels. In addition, the vertical height of the camera above the surface of the plant pots is 980mm and 9mm is the camera focal length.

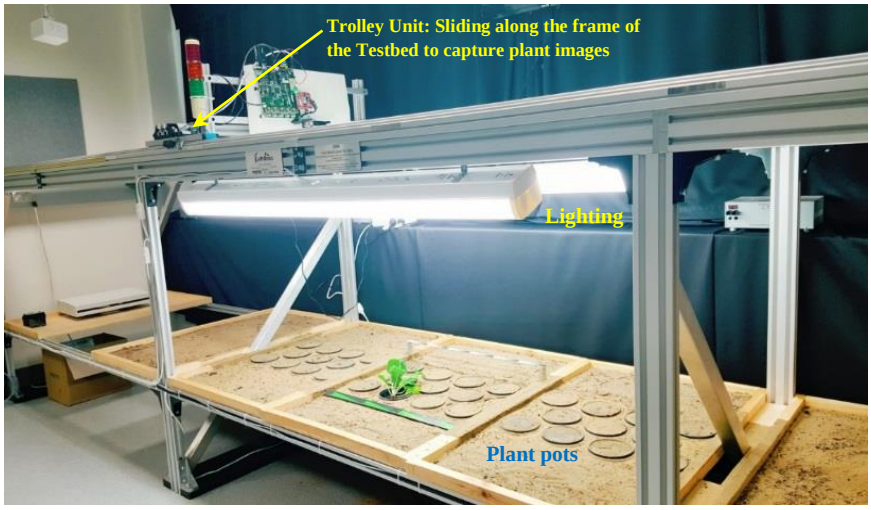


Figure 3. High-speed testbed used for controlled data capture.

As can be seen in Figure 3, individual trays are capable of holding 11 potted plants, with each tray filled with soil to provide a uniform background that can be used to simulate a West Australian wheat belt

farming environment. For experimental purposes, only the outer pot plant holders of the middle tray were used.

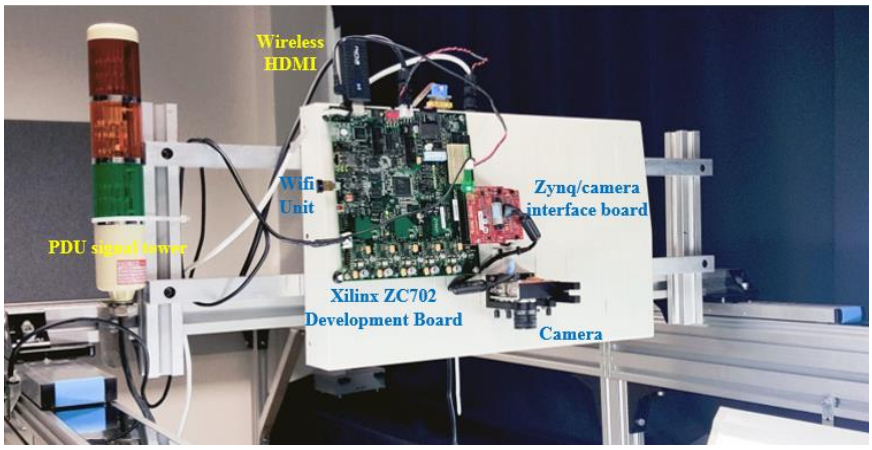


Figure 4. Zynq board with integrated VITA 2000 camera mounted on a moveable trolley.

The maximum allowable speed of the trolley is 5m/s, with the system capable of capturing images in real-time. The Testbed is also equipped with two fluorescent tube lamps as illustrated in Figure 3. The artificial lighting is there to provide uniform illumination for the purposes of data capture. For the purposes of the experimental work presented herein, all data was captured at a speed of 1m/s (3.6km/h) to capture high quality images.

Data capture runs comprised collecting multiple images of the individual test plants placed in the centre Testbed tray, Figure 3, with image variation obtained through manual plant rotation. The segmented greyscale images collectively formed the large data set used in the experimental work. This data set is referred to herein as "bccr-segset" and published online.

Data labelling

Data labelling was conducted by providing the ground truth in regard to which types of plants were identified in images. In the context of continuous runs on the Testbed, images comprised just background, partial plant with background or full plant with background., making the detection and classification processes challenging. Whilst the partial plant images could be removed from the dataset altogether, this would introduce a dataset bias. On the other hand, the human labelling error was quite high when attempts were made to decide among the labels that contained little plant information (i.e. "is this background or plant?"). Therefore, a semi-automatic way was adopted to solve this problem by thresholding the amount of green plant material according to their growth stages. If an image did not contain enough green plant material, then it was labelled as background.

First of all, as a pre-processing stage, images were filtered by using open and close morphological operations in order to remove the background noise. Then, binary images were segmented and

thresholded according to the amount of corresponding plant area found. Initial experiments showed that it was not sufficient to do a green threshold on the entire image, therefore images were divided into 7 equal areas (Top left, Top right, Bottom left and Bottom Right, Centre left, Centre and Centre right) as shown in Figure 5.

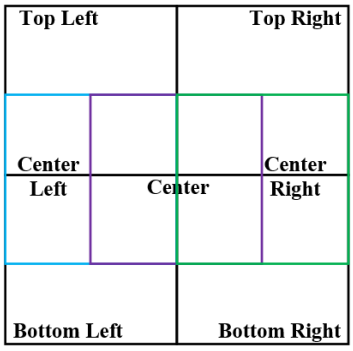


Figure 5. Thresholding areas used in collected images to filter partial plants with insufficient information for classification.

The thresholding test was applied for each of the square areas shown in Figure 5. The image was labelled as a plant class if the thresholding test passed for any of the areas. Lastly, an edge area threshold was also defined in order to allow for partial plants to have enough green material for identification. All the thresholds were experimentally derived and are shown in Table 1.

Table 1. Default thresholds for canola, corn and radish plants

Thresholds for plants (cm^2)	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**
Threshold (Inner, Edge) - Canola	(1.4, 3.3)	(3.0, 6.7)	(7.0, 10.0)	(8.0, 12.2)
Threshold (Inner, Edge) - Corn	(2.2, 5.7)	(3.0, 6.7)	(4.2, 9.2)	(7.5, 13.9)
Threshold (Inner, Edge) - Radish	(2.5, 4.0)	(3.2, 6.7)	(7.0, 10.0)	(8.0, 13.8)

As can be seen in Figure 6, partial plants in some growth stages with insufficient information were considered as a background class in the dataset. This allowed a more reliable labelling process without removing images from the dataset. In turn, this assured that the input sample distribution did not change.

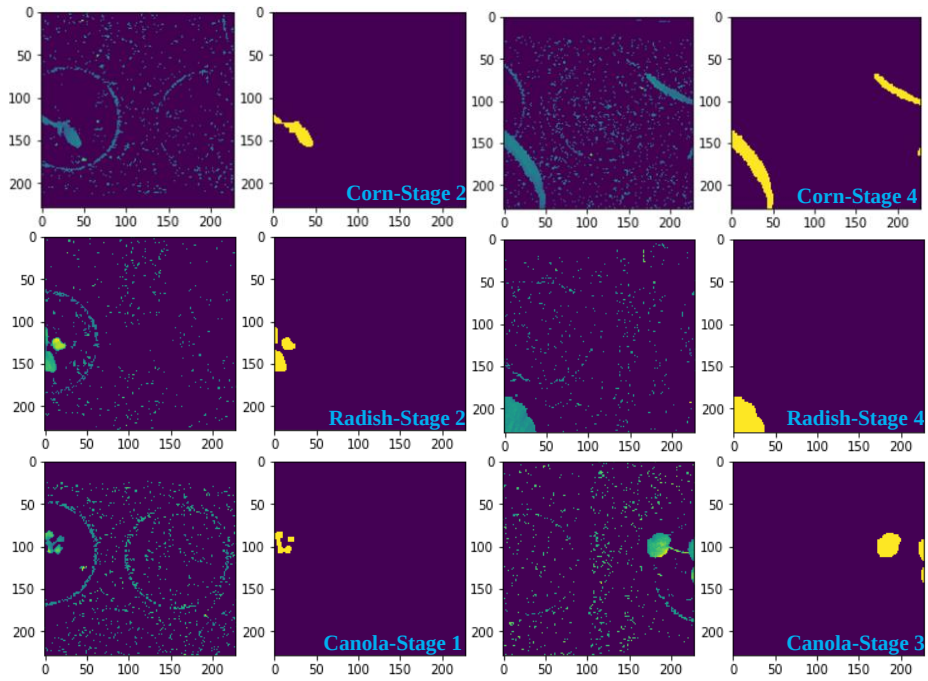


Figure 6. Examples of filtered and segmented images of 3 different partial plants (Canola, Corn and Radish) removed from the dataset at three different growth stages.

2.4 Methodology

All of the plant images went through the following processing steps: pre-processing, segmentation, feature extraction and classification. The extracted LBP features were stored in a database. Pre-processing was the same for both training and validation phases. The training dataset was trained by using the SVM and then the prediction model was exported to compare with textural features in the validation set for recognising and classifying different types of plants.

Figure 7 shows the flowchart that illustrates the training, testing and validation of the dataset through the combination of LBP operators and SVM for three-plant classification.

The steps shown in Figure 7 are summarised as follows:

1. The dataset with greyscale segmented images is provided to start the process.
2. In order to read all plant images, the location of the dataset is input.
3. The dataset is divided into the training and validation phases.
4. The LBP hyper-parameters are set, including the number of neighbours (P) and the radius (R), and a rotation invariant uniform ($riu2$) descriptor. In the preliminary results, $LBP_{8,1}^{riu2}$, $LBP_{16,2}^{riu2}$, $LBP_{24,3}^{riu2}$ and combined LBP operators ($LBP_{8,1}^{riu2} + LBP_{16,2}^{riu2} + LBP_{24,3}^{riu2}$) are applied to extract robust features from plant images.
5. The LBP method is initialised by inputting hyper-parameters then run to extract features from plant images.

6. Labelling images corresponding to what they represent (the classes selected in this paper are Canola, Corn, Radish and Background). After that, a table of features and labels is generated to prepare for the training process by programming in MATLAB®.
7. The table of robust features and labels is regarded as an input dataset for training.
8. Apply the SVM approach with 5-fold cross validation to classify different types of plants. After training the dataset, a model is exported to make predictions for the plant images in a validation dataset.
9. The classification accuracy and *F1* score are calculated. When other hyper-parameters are to be tested, this model is restarted at step 4.

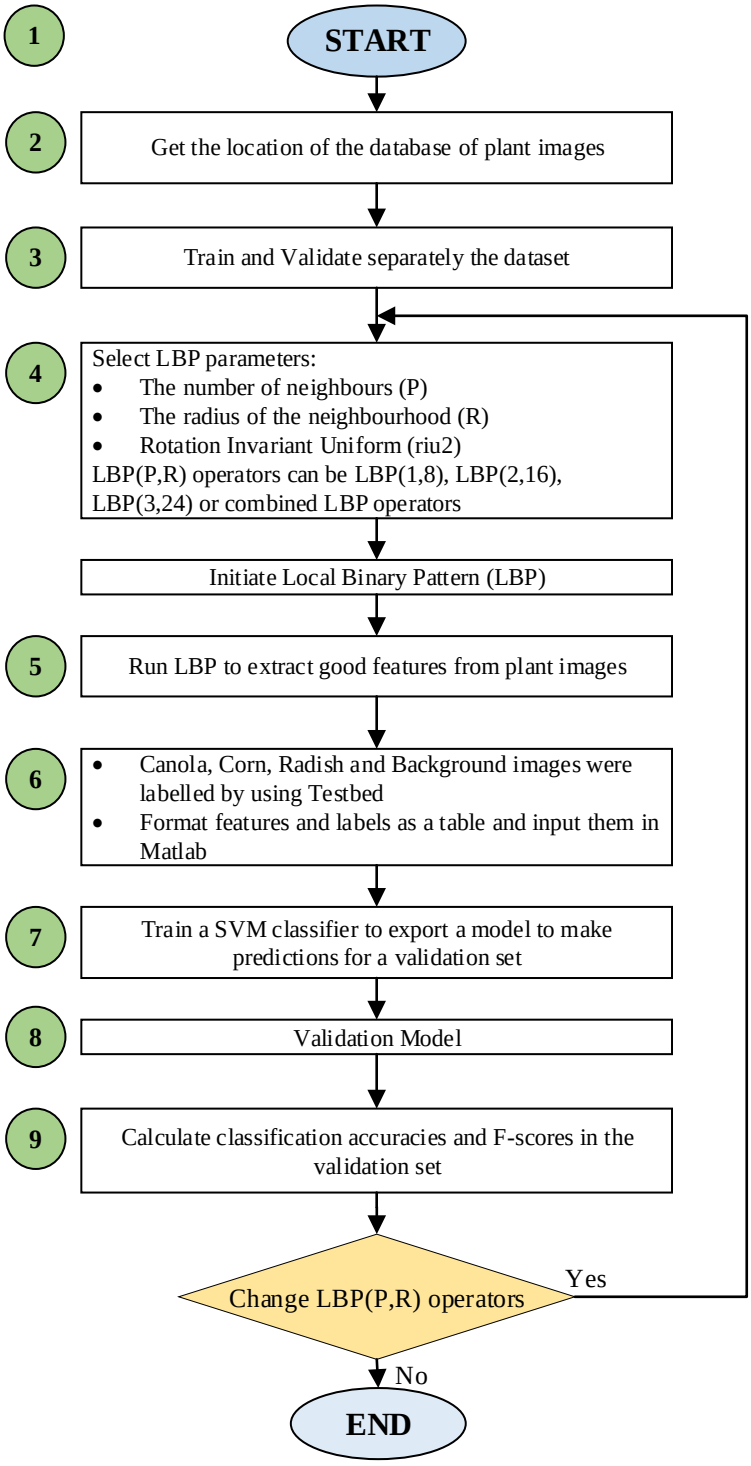


Figure 7. A flowchart for training, testing and validating the dataset

2.5 Results and Discussion

The results are divided into two sections: (i) the accuracies of classification models are evaluated based on comparing an unsegmented validation dataset with a validation segmented dataset, and (ii) the classification accuracy of the LBP operators and the SVM in the large dataset is reported. As noted in the data collection section, plant images were captured at the same height from the camera to the plant

pots. Therefore, the scales of the images of the plants taken during the four growth stages corresponded to the actual sizes of the plants. The computer used in these experiments had a 3.4GHz processor, 16GB RAM and ran MATLAB 2016b.

2.5.1 Initial results of the comparison between classification accuracies of an unsegmented dataset and a segmented dataset

In this section, an initial performance comparison is made between segmented and unsegmented greyscale images. With regard the current experimental setup, the effort required to capture and label the unsegmented greyscale images is greater than that of capturing segmented images. Experiments are conducted by selecting unsegmented and segmented datasets with 4032 images in each dataset. The detailed parameters of the two datasets are listed in Table 2. All plant samples consisted of canola and corn species taken, as previously mentioned, at three growth stages. The number of canola samples was equal to the number of corn samples in the training sets and the validation sets. Typical plant images in the unsegmented and segmented dataset for three different growth stages are shown in Figure 8.

Table 2. Parameters of unsegmented and segmented datasets

Parameters	Greyscale unsegmented and segmented images
Total images	4032 images in each dataset
Train set	3360 images in each dataset
Validation set	672 images in each dataset
Number of classes	2 classes (canola and corn plants)
Image size	228 × 228 pixels

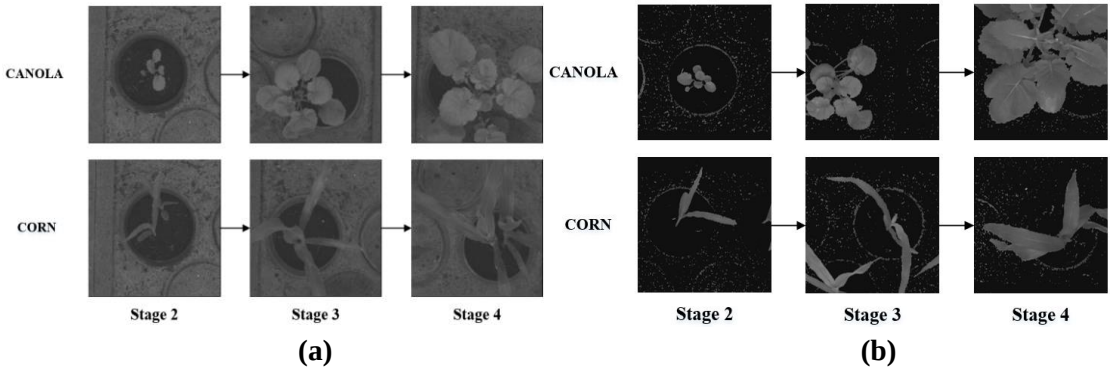


Figure 8. Greyscale unsegmented (a) and segmented (b) plant images at three different growth stages of canola and corn plants

The results of the classification accuracy were assessed against the percentages of correct classified plants. It can be observed from Table 3 that the combination of LBP operators significantly improves the classification accuracies in the validation sets. According to Ojala et al., the performance of the

combined LBP operators outperformed that of single LBP operators [6]. In this experiment, it was obviously true that the classification accuracies achieved using the combination of $LBP_{8,1}^{riu2}$, $LBP_{16,2}^{riu2}$ and $LBP_{24,3}^{riu2}$ was also higher than those attained using single LBP operators. This demonstrates that robust features extracted through the combined-LBP operators can increase the classification accuracy and *F1* scores. In comparison with using the greyscale unsegmented dataset, the accuracy of classification models using the validation segmented dataset is generally higher.

Table 3. Classification accuracies attained by using LBP operators with SVM for two different validation datasets.

LBP operators with 5-fold cross validation	Number of bins	Unsegmented dataset accuracy	Segmented dataset accuracy
(8,1)	10	79.91%	75.45%
(16,2)	18	91.52%	95.98%
(24,3)	26	93.01%	97.02%
(8,1) + (16,2)	28	94.20%	98.07%
(8,1) + (24,3)	36	96.28%	99.40%
(16,2) + (24,3)	44	95.83%	98.51%
(8,1) + (16,2) + (24,3)	54	97.32%	99.26%
(8,1) + (16,2) + (24,3) + PCA	16	95.24%	98.07%

The experimental results shown in Table 3 show that converting RGB plant images into greyscale without segmentation does not increase the classification accuracy. Whereas, by segmenting RGB images using the ExG-ExR method and then converting them to greyscale results in higher classification accuracy. Furthermore, experimental results show that by applying the above-mentioned pre-segmentation steps an increase of 2-4% in accuracy is attained, for the detection and discrimination of plant species.

Principal Component Analysis (PCA) is a useful tool for reducing the dimensionality of data. Typically, PCA produces the principal components of an image and extracts the relevant features from the data matrix of the image by calculating the eigenvalues. Note, however, in some cases, many significant features could be eliminated when PCA is applied, thereby reducing plant discrimination accuracy [63, 64]. Therefore, optimising the number of retained principal components is important for increasing plant discrimination accuracy. In our experiments, PCA was used in conjunction with the combined-LBP operators and SVM, and the optimum number of principal components for our algorithms was found to be 16. This optimum number was deduced experimentally and is offered herein for completion.

Note that classification accuracy is not a sufficient indicator to claim that the model is acceptable for plant classification [60]. In fact, three other indicators (*Precision*, *Recall*, and *F1* score) are typical to validate the suitability of the model for plant classification. Table 4 shows the *F1* scores of the

classification models for the validation unsegmented and validation segmented datasets, for canola and corn plants. As seen from Table 4, the *F1* scores for canola and corn plants are relatively similar. It is obvious from Table 4 that the highest *F1* scores (>99%) are attained with segmented data and the combination of $LBP_{8,1}^{riu2}$ and $LBP_{24,3}^{riu2}$.

Table 4. F1 scores of the classification models for the validation unsegmented and validation segmented datasets.

LBP operators with 5-fold cross validation	F1 scores of the unsegmented dataset		F1 scores of the segmented dataset	
	Canola	Corn	Canola	Corn
(8,1)	79.88%	79.94%	74.67%	74.44%
(16,2)	91.45%	91.58%	95.96%	96.00%
(24,3)	92.97%	93.04%	97.07%	96.98%
(8,1) + (16,2)	94.24%	94.15%	98.07%	98.06%
(8,1) + (24,3)	96.26%	96.30%	99.41%	99.40%
(16,2) + (24,3)	95.77%	95.89%	98.52%	98.51%
(8,1) + (16,2) + (24,3)	97.28%	97.36%	99.26%	99.25%
(8,1) + (16,2) + (24,3) +PCA	95.18%	95.30%	98.01%	98.12%

2.5.2 Classification accuracies and F1 scores of a multi-class dataset

Having investigated the performance of the greyscale segmented images (in Section 2.5.1), we discuss in this section the performance of the method based on the combination of the LBP operators and SVM for a larger dataset, using only greyscale segmented images.

In these experiments, canola, corn and radish plants were collected at four different growth stages, using the custom-built testbed. Images were segmented and converted to greyscale with the size of 228×228 pixels. The datasets were divided into training and validation, as illustrated in Figure 9.

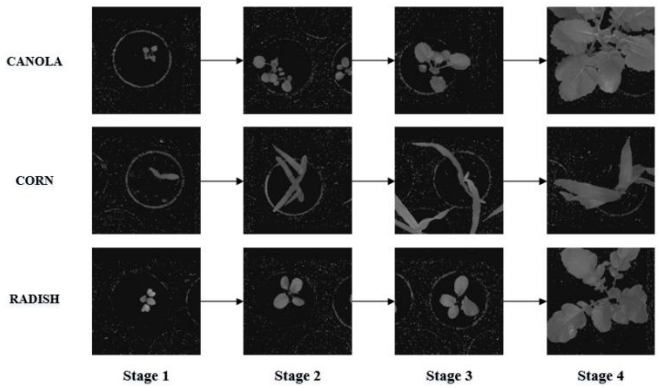


Figure 9. Segmented greyscale images of canola, corn and radish, at four different growth stages.

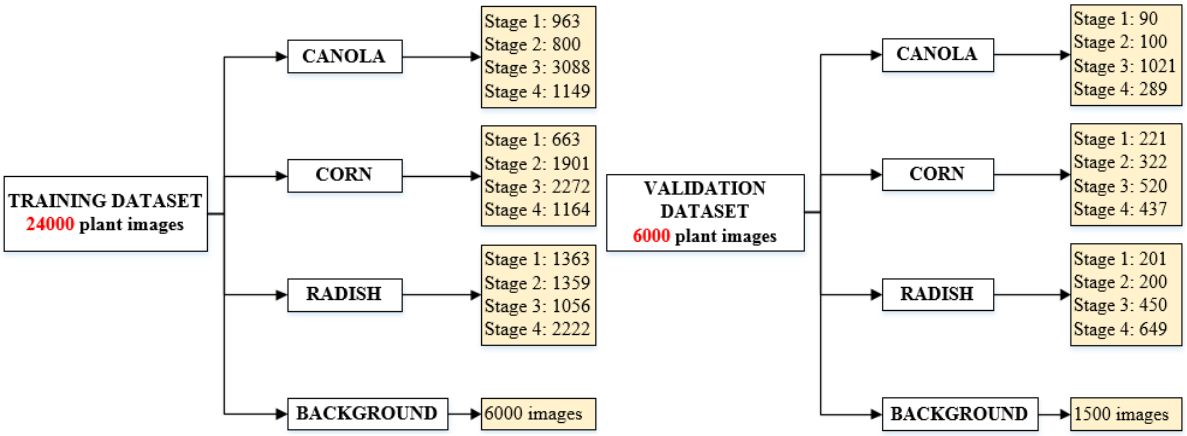


Figure 10. Illustration of the partitioning of the big dataset into training and validation datasets for canola, corn, radish and background.

The training dataset was used to train the SVM classifier with 5-fold cross validation to generate a prediction model for the validation dataset. Kernel functions were introduced to enhance efficient non-linear classification. Note that polynomial kernels and radial basis functions are widely used with SVM [65]. Different kernels were trialled in the experiments with the quadratic kernel was found to be more effective for SVM and LBP combination, the quadratic kernel generating the best and most consistent results. The "one against one" SVM strategy was selected in this scenario due to the large number of training images [66]. This obtained the optimum compromise between training time and accuracy performance. MATLAB was used to visualize the distribution of the LBP textural features.

Figure 10 shows the distribution of the training dataset for canola, corn, radish and background, using LBP operators ($LBP_{8,1}^{riu2}$, $LBP_{16,2}^{riu2}$, $LBP_{24,3}^{riu2}$) and the SVM classifier. The scatter plot shown in Figure 11 illustrates the distribution of two selected features out of a total of 54 features. From the plant images, it is obvious that the texture of the corn leaves is completely different to that of the leaves of canola and radish. Corn is categorised as a narrow leaf plant, whilst canola and radish are broad leaf plants. The distributions of canola and radish plant features overlap, mainly because their measured textural features are similar, making their discrimination challenging. Intuitively, these plants have the same botanical family (Brassicaceae or Cruciferae) and corn belongs to grass family (Poaceae). However, this plot is limited by the distribution of 2 selected features.

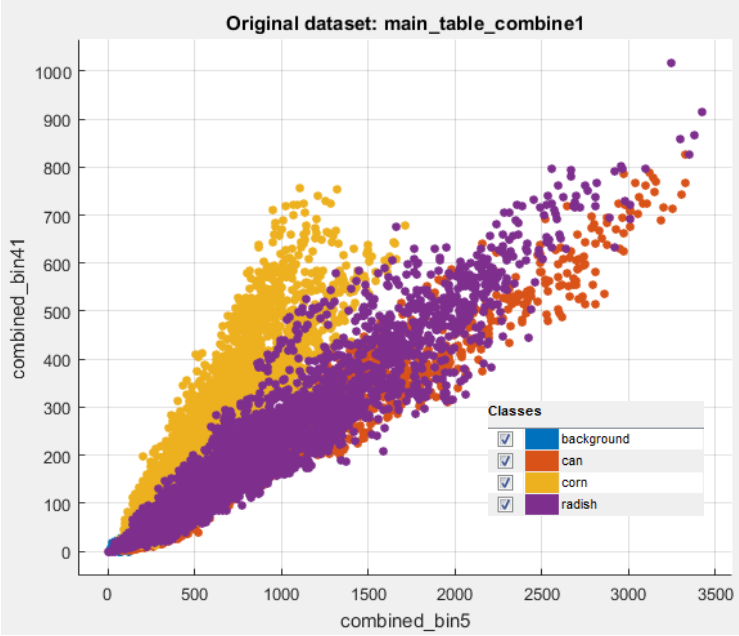


Figure 11. Typical textural feature distribution of the training dataset for canola, corn, radish and background. Based on the LBP operators ($LBP_{8,1}^{riu2}, LBP_{16,2}^{riu2}, LBP_{24,3}^{riu2}$) and the SVM classifier. Textural feature distribution is shown for two selected features out of a total of 54 features.

In order to visualize the structure of the "bccr-segset" large dataset in a two-dimensional map, we used t-SNE technique [67] for the train dataset (24000 plant images) and test dataset (6000 plant images). According to the article and user's guide for t-SNE, we implemented this technique by using Matlab with main parameters such as two-dimensional visualization, dimensionality reduction of the data (the value was 50), perplexity of the Gaussian distributions (the value was 30). As can be observed from Figure 12, the distribution of background class is totally separated from other classes. Meanwhile, the distribution of corn, canola and radish images was classified into many small groups and had some overlapping patterns. This leads to the increased misclassification among canola, corn and radish images.

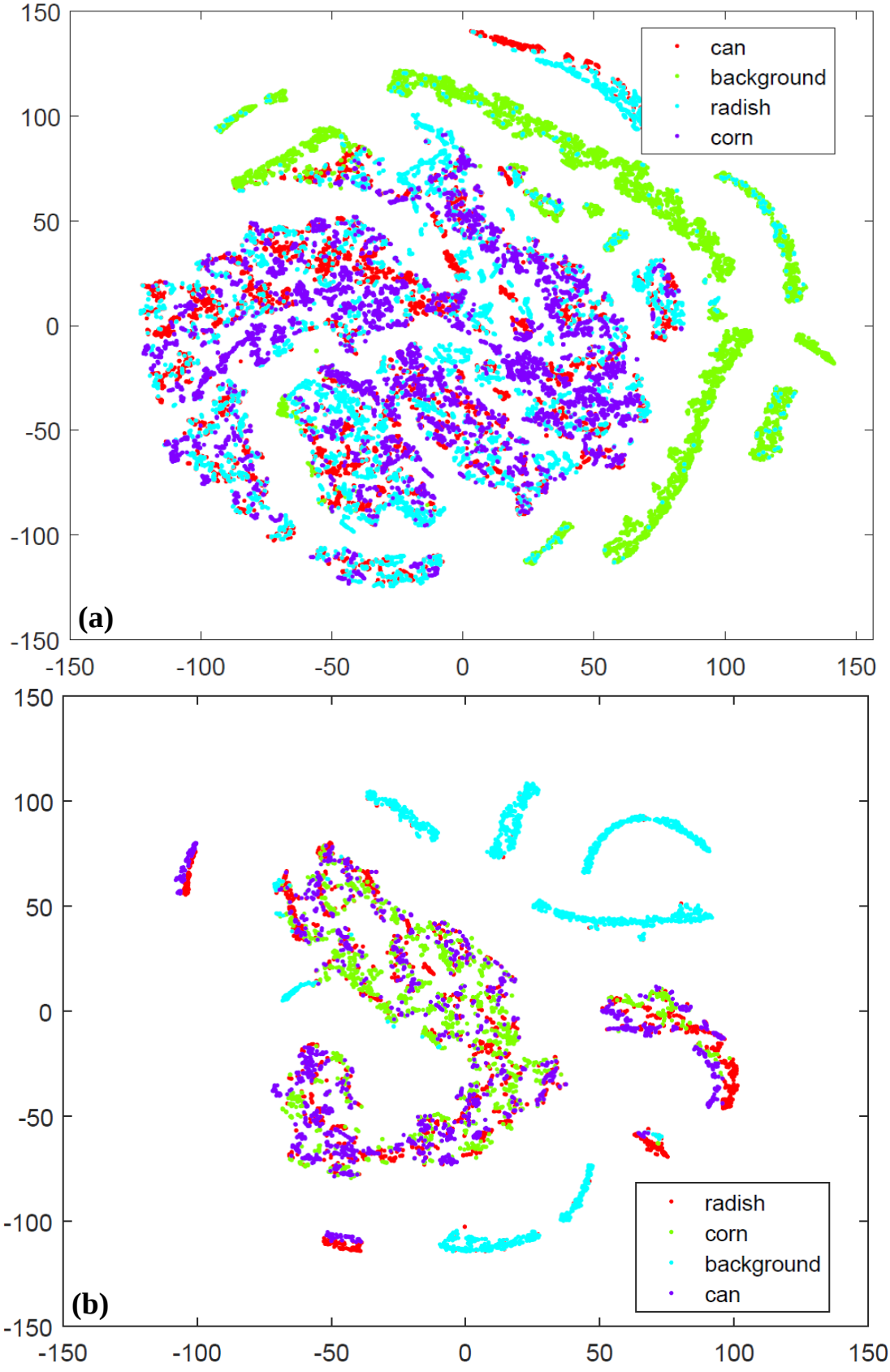


Figure 12. Visualization of (a) the train dataset **(**24000 plant images) and (b) the test dataset (6000 plant images) with 4 classes (background, canola, corn and radish).

For the validation set, the generated prediction model was applied to evaluate the robustness of this model by evaluating the classification accuracies for scenarios of two classes, three classes and four classes. To evaluate the quality of classification of the model, we applied performance measures to calculate the confusion matrices described in Section 2.4.

Performance metrics for multi-class classification were computed by applying the general formulas from Sokolova and Lapalme [61]. After training the 24000-plant-image dataset, Table 5 shows the average classification accuracy results obtained on the test dataset (6000 plant images) by using the combination LBP operators ($LBP_{8,1}^{riu2}, LBP_{16,2}^{riu2}, LBP_{24,3}^{riu2}$) with PCA (16 principle components) and without PCA. The classification accuracy of LBP operators without PCA shown in Table 5 was relatively higher than the one with PCA. However, a slight improvement in execution time was obtained by applying PCA, due to reduction of features considered to 16 dominant features.

Table 5. Classification accuracies of an algorithm combining LBP operators ($LBP_{8,1}^{riu2}, LBP_{16,2}^{riu2}, LBP_{24,3}^{riu2}$) and SVM for different scenarios. Execution time and PCA is shown herein for completion.

LBP operators with 5-fold cross validation	**Average classification accuracy of LBP operators (8,1) + (16,2) + (24,3)**	**Execution time (Milliseconds/ Image)**	**Average classification accuracy of LBP operators (8,1) + (16,2) + (24,3) with PCA (16 principle components)**	**Execution time (Milliseconds/ Image)**
Four classes (Canola, Corn, Radish & Background)	91.85%	47.898	91.08%	45.418

To have a better understanding of classification for classes, Table 6 shows the confusion matrix of the test dataset for four classes which was obtained by using SVM (polynomial kernel, order 2) without PCA. After calculating the number of correctly and falsely classified images in the confusion matrix, *TP*, *FP* and *FN* parameters in each class were calculated. We applied performance measures to calculate the confusion matrix, precision, recall and *F1*-score of the test dataset described in Section 2.3.4 by using the SVM classifier (polynomial kernel, order 2) were computed as shown in Table 7.

Table 6. The number of plant images in the test dataset correctly and incorrectly recognized using the confusion matrix, for a group of three plants (canola, corn and radish) and background.

		Predicted classes				Parameters		
		Background	Canola	Corn	Radish	TP	FP	FN
Actual classes	Background	1479	3	0	18	1479	58	21
	Canola	0	1253	15	232	1253	154	247
	Corn	2	24	1471	3	1471	24	29
	Radish	56	127	9	1308	1308	253	192

The evaluation of the performance of different SVM kernels is presented in Table 7. According to a comparison of the *F1* scores for multi-class classification, the classification performance of SVM (polynomial kernel, order 2) with 91.83% was higher than SVM (polynomial kernel, order 3) and SVM (RBF kernel) with 90.66% and 90.78% respectively. Furthermore, corn and background classes were classified with high accuracy. In contrast, for groups with many similar features (canola and radish), the algorithm displayed reduced discrimination capability.

The distinctions in the leaf texture of plants and the number of green pixels in images provided significant information for the reliability of classification results. In particular, the differences between narrow-leaf and broadleaf plants enhanced the classification rates. Therefore, background and corn images were classified with higher accuracy compared to canola and radish images. As for the similarity between canola and radish plants, the *F1* score of differentiating between them in Table 7 were considerably lower. These plants with round shaped leaves can be discriminated by simply recognizing the edges of canola plants, which generally look like outward-pointing teeth. In addition, one of the main obstacles for the relatively high misclassification rates is that plant leaves may look unexpectedly deformed and twisted after imaging, since these plants are not always perpendicular to the camera lens. Overall, the algorithm combining LBP operators with SVM produced consistently robust classification, scale and rotation invariance.

Table 7. Precision, Recall and F1-score of the test dataset with different SVM kernels

SVM kernels	Train the dataset	Classes	Precision	Recall	F1-score
Quadratic SVM	95.20%±0.25	Background	96.23%	98.60%	97.40%
		Canola	89.05%	83.53%	86.21%
		Corn	98.39%	98.07%	98.23%
		Radish	83.79%	87.20%	85.46%
	The average of parameters		**91.87%**	**91.85%**	**91.83%**
Cubic SVM	96.00%±1.11	Background	96.41%	98.33%	97.36%
		Canola	86.59%	82.20%	84.34%
		Corn	98.04%	96.93%	97.49%
		Radish	81.77%	85.20%	83.45%
	The average of parameters		**90.70%**	**90.67%**	**90.66%**
RBF kernel	94.90%±0.37	Background	96.17%	98.87%	97.50%
		Canola	83.64%	85.20%	84.41%
		Corn	98.64%	96.87%	97.75%
		Radish	84.69%	82.27%	83.46%
	The average of parameters		**90.79%**	**90.80%**	**90.78%**

To investigate the performance of SVM kernels, we conducted a comparative study of the *F1* scores for SVM classifier and K-Nearest Neighbour (KNN) classifier. KNN is an algorithm for classifying classes based on a similarity measure (distance functions) [68]. This method has two types of distance functions including distance metric and distance weight [69]. Particularly, three distance metrics including Euclidean, Minkowski and Cosine were used in this experiment and the results were computed by using Matlab. It is generally observed in Table 8 that the average F1 score in the case of using weight KNN (86.73%) was higher than other KNN techniques such as Coarse KNN (82.67%), Cosine KNN (83.79%), Fine KNN (85.78%), Cubic KNN (86.26%) and Medium KNN (86.50%). Based on the results from Table 7 and Table 8, the SVM classifier outperformed the KNN classifier for the test dataset (6000 images).

Table 8. Precision, Recall and F1-score of the test dataset with different types of KNN

KNN	Classes	Precision	Recall	F1-score
Fine KNN	Background	95.75%	96.20%	95.98%
Number of neighbours:1	Canola	77.37%	76.80%	77.08%
Distance metric: Euclidean	Corn	96.98%	91.93%	94.39%
Distance metric: Equal	Radish	73.70%	77.73%	75.67%
The average of parameters		**85.95%**	**85.67%**	**85.78%**
Medium KNN	Background	96.11%	98.87%	97.47%
Number of neighbours:10	Canola	74.10%	83.93%	78.71%
Distance metric: Euclidean	Corn	96.65%	92.40%	94.48%
Distance metric: Equal	Radish	80.36%	70.93%	75.35%
The average of parameters		**86.81%**	**86.53%**	**86.50%**
Coarse KNN	Background	95.55%	98.80%	97.15%
Number of neighbours:100	Canola	66.56%	81.33%	73.21%
Distance metric: Euclidean	Corn	95.50%	89.20%	92.24%
Distance metric: Equal	Radish	76.05%	61.60%	68.07%
The average of parameters		**83.42%**	**82.73%**	**82.67%**
Cosine KNN	Background	85.31%	99.13%	91.71%
Number of neighbours:10	Canola	77.69%	72.67%	75.09%
Distance metric: Cosine	Corn	95.80%	88.13%	91.81%
Distance metric: Equal	Radish	77.20%	75.87%	76.53%
The average of parameters		**84.00%**	**83.95%**	**83.79%**
Cubic KNN	Background	96.05%	98.80%	97.40%
Number of neighbours:10	Canola	73.52%	83.87%	78.36%
Distance metric: Minkowski	Corn	96.58%	92.13%	94.30%
Distance metric: Equal	Radish	80.23%	70.33%	74.96%
The average of parameters		**86.60%**	**86.28%**	**86.26%**
Weighted KNN	Background	96.11%	98.87%	97.47%
Number of neighbours:10	Canola	76.05%	80.67%	78.29%
Distance metric: Euclidean	Corn	96.54%	93.07%	94.77%
Distance metric: Squared inverse	Radish	78.52%	74.33%	76.37%
The average of parameters		**86.81%**	**86.74%**	**86.73%**

We used the dataset with four-growth stages, where leaves in each stage were captured with the difference of size and morphology. However, the number of collected images as mentioned in Figure 10 was not equal in each stage. In order to evaluate the performance of the classification of 4 different plant classes in each stage, we divided and equalised the train dataset (3200 plant images with 800 images in each class) and the test dataset (320 images with 80 images in each class). In addition, the effectiveness of the classified plant images was evaluated by the *F1* scores in the case of three different SVM kernels. As can be observed in Table 9, the *F1* score at stage 1 was higher than those at other stages. The morphology of canola and radish in stage 1 is distinctly different. Specifically, the two-heart shape of radish leaves in stage 1 has a distinctive appearance compared to the shape of canola leaves. As for the stage 2 and 3, the classification performance of SVM (RBF kernel) was higher than

of SVM (polynomial kernel, order 2 and 3). However, the number of correctly classified plant images based on the *F1* score was higher for the SVM (polynomial kernel, order 2) in comparison with the SVM (RBF kernel).

Table 9. Precision, Recall and F1-score of the test dataset at four-growth stages with different SVM kernels

		SVM (Polynomial, order 2)	SVM (Polynomial, order 3)	SVM (RBF kernel)
Stages	**Plant Categories**	**F1-score**	**F1-score**	**F1-score**
Stage 1	Background	98.73%	98.73%	98.73%
	Canola	98.16%	97.53%	98.77%
	Corn	100.00%	100.00%	100.00%
	Radish	98.11%	97.50%	97.50%
Average F1-score in Stage 1		**98.75%**	**98.44%**	**98.75%**
Stage 2	Background	99.37%	99.37%	99.37%
	Canola	68.15%	85.71%	86.75%
	Corn	90.91%	98.77%	97.53%
	Radish	80.00%	86.08%	86.27%
Average F1-score in Stage 2		**84.61%**	**92.48%**	**92.48%**
Stage 3	Background	96.10%	84.89%	99.37%
	Canola	85.71%	92.50%	88.05%
	Corn	98.14%	99.37%	99.37%
	Radish	83.04%	82.42%	88.34%
Average F1-score in Stage 3		**90.75%**	**89.80%**	**93.78%**
Stage 4	Background	98.14%	98.73%	98.09%
	Canola	92.22%	87.43%	86.96%
	Corn	98.73%	98.75%	98.11%
	Radish	93.51%	84.89%	84.29%
Average F1-score in Stage 4		**95.65%**	**92.45%**	**91.86%**

The capability of discriminating between canola and radish images in Table 9 was always lower than for background and corn images. Consequently, improving the LBP method is crucial to discriminate plant species with relatively similar features. A possible way to achieve this is to combine the uniform rotation invariant LBP features with significant non-uniform LBP features. Another potential approach is to take all features of the LBP method to acquire vital information of microscopic images of the plant species [70]. These are promising approaches that enable the development of LBP algorithms for the discrimination of plant species of similar features.

2.6 Conclusions

An algorithm based on the combination of LBP operators and an SVM classifier has been investigated, and its performance experimentally evaluated for the discrimination of different types of plants. An

initial comparison of unsegmented and segmented dataset types has been carried out in order to identify the type that yields higher classification accuracy. This comparison has shown that the green segmentation pre-processing step is beneficial for feature extraction and classification. A large segmented dataset has been collected using a high-speed Testbed that enabled the methods to be assessed and validated. A dataset has been made available (published online), which can be flexibly used by other researchers for information and comparison. Particularly, eight cases have been created using the large dataset and the experimental results have demonstrated that the combined LBP algorithm can attain a discrimination accuracy greater than 91% for corn, canola and radish plants and background. Results have also shown that if the shapes of canola and radish leaves are similar, the classification accuracy of the LBP algorithm decreases significantly. Furthermore, results have shown that the current execution time of plant classification is short, making the combined LBP algorithm a promising candidate for real-time weed detection.

Future work is focusing on the extension of the LBP method using colour images (instead of grey-level) and the introduction of identification techniques based on the use of non-uniform patterns in order to increase the weed detection accuracy. In addition, further investigations are required for improving the classification of broad leaves (e.g., radish and canola) and assessing the LBP algorithm in scenarios in which weeds and crops are partially occluded.

2.7 References

[1] B. Leff, N. Ramankutty, and J. A. Foley, "Geographic distribution of major crops across the world," *Global Biogeochemical Cycles*, vol. 18, no. 1, pp. 1-27, 2004.

[2] J. M. DiTomaso, "Invasive weeds in rangelands: species, impacts, and management," *Weed science*, vol. 48, no. 2, pp. 255-265, 2000.

[3] A. Perez, F. Lopez, J. Benlloch, and S. Christensen, "Colour and shape analysis techniques for weed detection in cereal fields," *Computers and electronics in agriculture*, vol. 25, no. 3, pp. 197-212, 2000.

[4] H. T. Søgaard, "Weed classification by active shape models," *Biosystems engineering*, vol. 91, no. 3, pp. 271-281, 2005.

[5] D. Woebbecke, G. Meyer, K. Von Bargen, and D. Mortensen, "Color indices for weed identification under various soil, residue, and lighting conditions," *Transactions of the ASAE*, vol. 38, no. 1, pp. 259-269, 1995.

[6] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on pattern analysis and machine intelligence,* vol. 24, no. 7, pp. 971-987, 2002.

[7] L. Liu, S. Lao, P. W. Fieguth, Y. Guo, X. Wang, and M. Pietikäinen, "Median robust extended local binary pattern for texture classification," *IEEE Transactions on Image Processing,* vol. 25, no. 3, pp. 1368-1381, 2016.

[8] L. Tang, L. Tian, B. Steward, and J. Reid, "Texture-based weed classification using Gabor wavelets and neural network for real-time selective herbicide applications," *Urbana,* vol. 51, p. 61801, 1999.

[9] A. J. Ishak, A. Hussain, and M. M. Mustafa, "Weed image classification using Gabor wavelet and gradient field distribution," *Computers and Electronics in Agriculture,* vol. 66, no. 1, pp. 53-61, 2009.

[10] J. Bossu, C. Gée, G. Jones, and F. Truchetet, "Wavelet transform to discriminate between crop and weed in perspective agronomic images," *computers and electronics in agriculture,* vol. 65, no. 1, pp. 133-143, 2009.

[11] P. Hansen and J. Schjoerring, "Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression," *Remote sensing of environment,* vol. 86, no. 4, pp. 542-553, 2003.

[12] K. Thorp and L. Tian, "A review on remote sensing of weeds in agriculture," *Precision Agriculture,* vol. 5, no. 5, pp. 477-508, 2004.

[13] A. Huete, K. Didan, T. Miura, E. P. Rodriguez, X. Gao, and L. G. Ferreira, "Overview of the radiometric and biophysical performance of the MODIS vegetation indices," *Remote sensing of environment,* vol. 83, no. 1, pp. 195-213, 2002.

[14] M. Ozdogan, Y. Yang, G. Allez, and C. Cervantes, "Remote sensing of irrigated agriculture: Opportunities and challenges," *Remote sensing,* vol. 2, no. 9, pp. 2274-2304, 2010.

[15] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision,* vol. 60, no. 2, pp. 91-110, 2004.

[16] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Computer vision and image understanding,* vol. 110, no. 3, pp. 346-359, 2008.

[17] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, 2005, vol. 1: IEEE, pp. 886-893.

[18] C. Liu and H. Wechsler, "Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition," *IEEE Transactions on Image processing,* vol. 11, no. 4, pp. 467-476, 2002.

[19] S. Happy, A. George, and A. Routray, "A real time facial expression classification system using Local Binary Patterns," in *Intelligent Human Computer Interaction (IHCI), 4th International Conference*, 2012: IEEE, pp. 1-5.

[20] O. Lahdenoja, J. Poikonen, and M. Laiho, "Towards understanding the formation of uniform local binary patterns," *ISRN Machine Vision,* 2013.

[21] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE transactions on pattern analysis and machine intelligence,* vol. 28, no. 12, pp. 2037-2041, 2006.

[22] H. Jin, Q. Liu, H. Lu, and X. Tong, "Face detection using improved LBP under Bayesian framework," in *Image and Graphics (ICIG'04), Third International Conference*, 2004: IEEE, pp. 306-309.

[23] W. Louis and K. N. Plataniotis, "Co-occurrence of local binary patterns features for frontal face detection in surveillance applications," *EURASIP Journal on Image and Video Processing,* vol. 2011, no. 1, p. 745487, 2010.

[24] G. Zhao and M. Pietikainen, "Dynamic texture recognition using local binary patterns with an application to facial expressions," *IEEE transactions on pattern analysis and machine intelligence,* vol. 29, no. 6, 2007.

[25] C. Shan, S. Gong, and P. W. McOwan, "Facial expression recognition based on local binary patterns: A comprehensive study," *Image and Vision Computing,* vol. 27, no. 6, pp. 803-816, 2009.

[26] Z. Guo, L. Zhang, and D. Zhang, "A completed modeling of local binary pattern operator for texture classification," *IEEE Transactions on Image Processing,* vol. 19, no. 6, pp. 1657-1663, 2010.

[27] S. Liao, M. W. Law, and A. C. Chung, "Dominant local binary patterns for texture classification," *IEEE transactions on image processing*, vol. 18, no. 5, pp. 1107-1118, 2009.

[28] M. Heikkila and M. Pietikainen, "A texture-based method for modeling the background and detecting moving objects," *IEEE transactions on pattern analysis and machine intelligence*, vol. 28, no. 4, pp. 657-662, 2006.

[29] V. Kellokumpu, G. Zhao, and M. Pietikäinen, "Human activity recognition using a dynamic texture based method," in *BMVC*, 2008, vol. 1, p. 2.

[30] F. Ahmed, A. H. Bari, A. Shihavuddin, H. A. Al-Mamun, and P. Kwan, "A study on local binary pattern for automated weed classification using template matching and support vector machine," in *Computational Intelligence and Informatics (CINTI), 2011 IEEE 12th International Symposium on*, 2011: IEEE, pp. 329-334.

[31] F. Ahmed, M. H. Kabir, S. Bhuyan, H. Bari, and E. Hossain, "Automated weed classification with local pattern-based texture descriptors," *Int. Arab J. Inf. Technol.*, vol. 11, no. 1, pp. 87-94, 2014.

[32] T. Ojala, T. Maenpaa, M. Pietikainen, J. Viertola, J. Kyllonen, and S. Huovinen, "Outex-new framework for empirical evaluation of texture analysis algorithms," in *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, 2002, vol. 1: IEEE, pp. 701-706.

[33] P. Brodatz, *Textures: a photographic album for artists and designers*. Dover Pubns, 1966.

[34] S. Lazebnik, C. Schmid, and J. Ponce, "A sparse texture representation using local affine regions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1265-1278, 2005.

[35] Y. Xu, X. Yang, H. Ling, and H. Ji, "A new texture descriptor using multifractal analysis in multi-orientation wavelet pyramid," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, 2010: IEEE, pp. 161-168.

[36] M. Varma and A. Zisserman, "A statistical approach to material classification using image patch exemplars," *IEEE transactions on pattern analysis and machine intelligence*, vol. 31, no. 11, pp. 2032-2047, 2009.

[37] L. Liu, P. Fieguth, Y. Guo, X. Wang, and M. Pietikäinen, "Local binary features for texture classification: Taxonomy and experimental study," *Pattern Recognition*, vol. 62, pp. 135-160, 2017.

[38] T. Burks, S. Shearer, J. Heath, and K. Donohue, "Evaluation of neural-network classifiers for weed species discrimination," *Biosystems Engineering,* vol. 91, no. 3, pp. 293-304, 2005.

[39] L. Tang, L. Tian, and B. L. Steward, "Classification of broadleaf and grass weeds using Gabor wavelets and an artificial neural network," *Transactions of the ASAE,* vol. 46, no. 4, p. 1247, 2003.

[40] C. M. Onyango and J. Marchant, "Segmentation of row crop plants from weeds using colour and morphology," *Computers and electronics in agriculture,* vol. 39, no. 3, pp. 141-155, 2003.

[41] F.-M. De Rainville *et al.*, "Bayesian classification and unsupervised learning for isolating weeds in row crops," *Pattern Analysis and Applications,* vol. 17, no. 2, pp. 401-414, 2014.

[42] F. Ahmed, A. H. Bari, E. Hossain, H. A. Al-Mamun, and P. Kwan, "Performance analysis of support vector machine and Bayesian classifier for crop and weed classification from digital images," *World Applied Sciences Journal,* vol. 12, no. 4, pp. 432-440, 2011.

[43] I. Ahmad, M. H. Siddiqi, I. Fatima, S. Lee, and Y.-K. Lee, "Weed classification based on Haar wavelet transform via k-nearest neighbor (k-NN) for real-time automatic sprayer control system," in *Proceedings of the 5th International Conference on Ubiquitous Information Management and Communication*, 2011: ACM, p. 17.

[44] T. Burks, S. Shearer, and F. Payne, "Classification of weed species using color texture features and discriminant analysis," *Transactions of the ASAE,* vol. 43, no. 2, p. 441, 2000.

[45] R. Pydipati, T. Burks, and W. Lee, "Identification of citrus disease using color texture features and discriminant analysis," *Computers and electronics in agriculture,* vol. 52, no. 1, pp. 49-59, 2006.

[46] C. Pulido, L. Solaque, and N. Velasco, "Weed recognition by SVM texture feature classification in outdoor vegetable crop images," *IngenIería e InvestIgacIón,* vol. 37, no. 1, pp. 68-74, 2017.

[47] F. Ahmed, H. A. Al-Mamun, A. H. Bari, E. Hossain, and P. Kwan, "Classification of crops and weeds from digital images: A support vector machine approach," *Crop Protection,* vol. 40, pp. 98-104, 2012.

[48] J. Liu, S. Zhang, and S. Deng, "A method of plant classification based on wavelet transforms and support vector machines," in *International Conference on Intelligent Computing*, 2009: Springer, pp. 253-260.

[49] J. M. Guerrero, G. Pajares, M. Montalvo, J. Romeo, and M. Guijarro, "Support vector machines for crop/weeds identification in maize fields," *Expert Systems with Applications,* vol. 39, no. 12, pp. 11149-11155, 2012.

[50] S. Arivazhagan, R. N. Shebiah, S. Ananthi, and S. V. Varthini, "Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features," *Agricultural Engineering International: CIGR Journal,* vol. 15, no. 1, pp. 211-217, 2013.

[51] A. Camargo and J. Smith, "Image pattern classification for the identification of disease causing agents in plants," *Computers and Electronics in Agriculture,* vol. 66, no. 2, pp. 121-125, 2009.

[52] T. Rumpf, A.-K. Mahlein, U. Steiner, E.-C. Oerke, H.-W. Dehne, and L. Plümer, "Early detection and classification of plant diseases with Support Vector Machines based on hyperspectral reflectance," *Computers and Electronics in Agriculture,* vol. 74, no. 1, pp. 91-99, 2010.

[53] G. E. Meyer and J. C. Neto, "Verification of color vegetation indices for automated crop imaging applications," *Computers and electronics in agriculture,* vol. 63, no. 2, pp. 282-293, 2008.

[54] J. Camargo Neto, "A combined statistical-soft computing approach for classification and mapping weed species in minimum-tillage systems," 2004.

[55] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern recognition,* vol. 29, no. 1, pp. 51-59, 1996.

[56] L. Wu and Y. Wen, "Weed/corn seedling recognition by support vector machine using texture features," *African Journal of Agricultural Research,* vol. 4, no. 9, pp. 840-846, 2009.

[57] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proceedings of the fifth annual workshop on Computational learning theory*, 1992: ACM, pp. 144-152.

[58] A. Mathur and G. M. Foody, "Crop classification by support vector machine with intelligently selected training data for an operational application," *International Journal of Remote Sensing,* vol. 29, no. 8, pp. 2227-2240, 2008.

[59] C. Lin, "A support vector machine embedded weed identification system," University of Illinois at Urbana-Champaign, 2009.

[60] M. Sokolova, N. Japkowicz, and S. Szpakowicz, "Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation," in *Australian conference on artificial intelligence*, 2006, vol. 4304, pp. 1015-1021.

[61] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Information Processing & Management*, vol. 45, no. 4, pp. 427-437, 2009.

[62] Xilinx. "Xilinx Zynq-7000 All Programmable SoC ZC702 Evaluation Kit." Xilinx. https://www.xilinx.com/products/boards-and-kits/ek-z7-zc702-g.html (accessed 2006).

[63] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley interdisciplinary reviews: computational statistics*, vol. 2, no. 4, pp. 433-459, 2010.

[64] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics and intelligent laboratory systems*, vol. 2, no. 1-3, pp. 37-52, 1987.

[65] T. Hofmann, B. Schölkopf, and A. J. Smola, "Kernel methods in machine learning," *The annals of statistics*, pp. 1171-1220, 2008.

[66] J. Milgram, M. Cheriet, and R. Sabourin, ""One against one" or "one against all": Which one is better for handwriting recognition with SVMs?," in *Tenth international workshop on frontiers in handwriting recognition*, 2006: Suvisoft.

[67] L. v. d. Maaten and G. Hinton, "Visualizing data using t-SNE," *Journal of machine learning research*, vol. 9, no. Nov, pp. 2579-2605, 2008.

[68] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE transactions on information theory*, vol. 13, no. 1, pp. 21-27, 1967.

[69] K. Q. Weinberger and L. K. Saul, "Distance metric learning for large margin nearest neighbor classification," *Journal of Machine Learning Research*, vol. 10, no. Feb, pp. 207-244, 2009.

[70] A. R. Yadav, R. Anand, M. Dewal, and S. Gupta, "Multiresolution local binary pattern variants based texture feature extraction techniques for efficient classification of microscopic images of hardwood species," *Applied Soft Computing*, vol. 32, pp. 101-112, 2015.

Chapter 3 – A novel method for detecting morphologically similar crops and weeds based on the combination of contour masks and Local Binary Pattern operators

This chapter was published as an article in the GigaScience Journal, vol. 9, no. 3, 2020, DOI: 10.1093/gigascience/giaa017. This article appears as it does in print, with the exception of minor changes to the layout, number formats, font size and font style, which was implemented to maintain consistency in the formatting of this thesis.

3.1 Abstract

Weeds are a major cause of low agricultural productivity. Some weeds have morphological features similar to crops making them difficult to discriminate. This paper proposes a novel method using a combination of filtered-features extracted by combined Local Binary Pattern operators and features extracted by plant-leaf contour masks to improve the discrimination rate between broadleaf plants. Opening and closing morphological operators were applied to filter noise in plant images. The images at four stages of growth were collected using a testbed system. Mask-based Local Binary Pattern features were combined with filtered-features and a coefficient k. The classification of crops and weeds was achieved using support-vector-machine with radial basis function kernel. By investigating optimal parameters, this method reached a classification accuracy of 98.63% with four classes in the "bccr-segset" dataset published online in comparison with an accuracy of 91.85% attained by a previously reported method. The proposed method enhances the identification of crops and weeds with similar appearance and demonstrates its capabilities in real-time weed detection.

Keywords: Precision agriculture; Morphological operators; Feature extraction; Local Binary Patterns; Contour masks; Plant classification; Computer vision.

3.2 Introduction

Weed infestation poses a threat to the environment, crop yields and quality. Weeds in a field retard crop growth by competing for access to sunshine, water and nutrients. In particular, the density, spreading time and growth characteristics are important factors for weed management [1]. One of the most invasive and serious weeds is wild radish, which causes significant crop yield losses and low-quality crops due to its fast growth rate, contaminants, multiple-herbicide resistance and vigorous competition [2-4]. Currently, blanket herbicide spraying is the most common practice used to eradicate weeds. However, the excessive use of herbicides has negative impacts on the environment in addition to the development of herbicide-resistance properties in weeds. The dramatic challenge for controlling weeds

is to attain an optimal eradication efficacy with minimum herbicide usage. Note that, reducing the herbicide application rates brings down the cost of weed management. Hence, it is a worthwhile objective in precision agriculture.

Spraying selective weeds automatically in vegetation fields is considered as a potential method to reduce the environmental and economic costs of weed management. Wild radish is a dominant weed in all broadacre field crops, including wheat, barley, sorghum, maize and canola. Canola is the most difficult crop to discriminate against wild radish because of their morphological similarity [5]. Therefore, canola, corn and wild radish are selected for experimental investigation in this study. Classifying crops and wild radish plants is a vital practical problem in agriculture. The ability to accurately detect and classify weeds in row crops in real time enables the selective application of herbicides, thus enhancing the quality and productivity of crops.

There have been numerous studies on weed-from-crop discrimination. Spectral techniques based on the calculation of the Normalised Difference Vegetation Indices (NDVIs) [6, 7] have long been proposed for identifying plant species. However, this method has some deficiencies. In typical farm field conditions, the wind, shadowing, and soil background brightness may change the spectral features of plants, leading to the reduction of the discrimination accuracy of NDVI-based weed sensors [8, 9]. Due to the drawbacks of such spectral-reflectance sensors, research on spatial sensors based on the use of image processing techniques for the classification of plant species and weeds in real time have been conducted [10]. One such spatial technique is "texture analysis" in image processing, which has been applied in many fields, such as industrial inspection systems, medical image analysis, face recognition and content-based image retrieval [11]. There are significant challenges in image texture analysis, such as noise sensitivity, grey scale variation, rotation sensitivity and illumination and brightness conditions. One of the discriminative and computationally effective local texture descriptors that can potentially overcome these issues is local binary patterns (LBP) [12-14]. The important role of extracting dominant features is emphasized, as poor features combining with even the best classifier are unlikely to achieve good identification results.

In this paper, the LBP method is applied to extract plant features due to its flexibility and robustness in monotonic grey-level transformation, illumination, scaling, viewpoint, and rotation variance. Furthermore, the LBP method is also a robust tool for identifying the relationship among the pixels in plant images and detecting microstructures including lines, spots, edges and flat areas [14]. Another attractive feature of the LBP method is low computational complexity [15]. In fact, the LPB is computationally less complex than its SIFT or SURF counterparts [16]. Finally, it has exhibited superior performance in various applications, such as motion analysis [17, 18], texture recognition [12, 14, 19], face recognition [20-22], face expression analysis [23, 24], fingerprint recognition [25] and image retrieval [26, 27].

Numerous studies on the LBP method have been developed to enhance its discriminative power including Completed LBP [12], Extended LBP [28, 29], Discriminative completed LBP [30], Dominant LBP with Gabor filtering features [19], Pairwise rotation invariant co-occurrence LBP [31], Fuzzy LBP [32], Robust LBP [33], Noise-tolerant LBP [34] and Noise resistant LBP [35]. However, these methods still have unsatisfying tolerance to noise in images and increased feature dimensionality, leading to high computational complexity [36].

In the agricultural context, the complex and similar morphologies of plant leaves are one of the key challenges to find effective and discriminative plant descriptors. Combining LBP features with other features from different methods has become an interesting research topic in plant recognition. There have been several approaches based on applying the LBP method for the identification and classification of plants. For example, using LBP, in conjunction with template matching and SVM, was proposed to classify broadleaf and grass weed images [37]. These weed images having broad and narrow leaf shapes were easily distinguished. Similarly, another study on combining LBP, Local Ternary Pattern and Local Directional to classify broadleaf and narrow grass weeds [38]. Another statistical method for separating sugar beets and weeds has been proposed, based on using shape features [39]. However, this method was considered accurate only because the sugar beet sizes were significantly different from those of the weeds. The LBP method has also been used for crop segmentation in order to detect occluded crops (sweet pepper) [40]. However, the detection accuracy was quite limited (just 66.8%). The detection and classification of apple fruit diseases using Global Colour Histogram, Colour Coherence Vector, LBP and Complete LBP has been investigated [41]. The classification accuracy of this method was just above 93%. Identifying medicinal plants was conducted by combining morphological, LBP variance and colour features and the classification accuracy of this method was 72.16% [42]. In addition, canola, corn and radish plants have been classified using the combined LBP operators and SVM with a classification accuracy of 91.85% [43]. These methods are still deemed unsatisfactorily due to their low classification accuracy.

Some studies have investigated a promising approach to reducing noise and increasing classification accuracy is the combination of the LBP operators and contours that mask LBP images. LBP-guided active contour approaches have only been proposed for texture segmentation [44]. The active contour can identify the position of the initial curve anywhere in the captured image and then automatically detect interior contours. By combining scalar and vector LBP active contours, reduced computational cost and high segmentation quality can be achieved. However, typically, this method has been applied in the segmentation process. LBP-based edge-texture features for object recognition has also been proposed [45]. Particularly, discriminative LBP (DLBP) and Local Ternary Pattern (DLTP) were focused on differentiating a bright object against a dark background by combining edge and texture information. Another method for detecting humans based on non-redundant LBP shape descriptor has been implemented by concatenating a set of local appearance descriptors extracted at a set of key points.

However, occlusion was the main limitation that made this method impractical [46]. Another LBP edge-mapped descriptor for face recognition has been investigated [47], whereby LBP was applied on the edge contours (eyes, nose, and mouth) instead of the whole image, then the LBP intensity was combined with the edge pixel array around the feature points.

The above-mentioned methods have their own drawbacks, such as having unsatisfactory classification accuracy, computational complexity, application-specific recognition and not dealing with occlusion. In the context of this paper, we address the challenge of discriminating broadleaf plants species of relatively similar morphology by proposing a novel method called "filtered LBP method with contour mask and coefficient k (k-FLBPCM)", which enhances the plant discrimination capability. The k-FLBPCM is based on combining filtered LBP features and contour mask-based features to precisely identify and classify broad-leaf plants in the field. The current k-FLBPCM method has particularly been applied for the classification of two broad-leaf plants, namely canola (crop) and wild radish (weed), which significantly improves on the accuracy of our previously published paper [43]. This paper still employs a support vector machine (SVM) classifier due to its good accuracy and relevance to real-life datasets [48, 49]. The "bccr-segset" dataset, which comprises a variety of plant images at four defined growth stages, with rotation, scale and viewpoint variance, is used in this paper in order to compare the present results with our previously reported results.

3.3 Materials

3.3.1 Morphological operations

The Excess Green minus Excess Red Indices (ExG-ExR) method was used to segment green plant regions in the bccr-segset dataset [43]. During segmentation, the noise in plant images creates issues in the process of edge detection. However, reducing the noise level in these plant images plays an important role in image enhancement for the next stages of feature extraction and classification.

Morphological image processing is particularly investigated in this paper [50]. Morphological operators are introduced and extended to analyse images by Serra [51]. Particularly, in morphological analysis, images are treated as sets that illustrate the plant shapes, represented in greyscale or binary images. Morphological transformations are a tool that helps extract features from images using Minkowski addition and subtraction [52]. The morphological process needs two inputs including grey-scale images and structuring elements. The function of morphology operators is to transform from one set to another with the aim of searching the special structure of the original set. Then, the special structure information is stored in the transformed set and the transformation is recognized by special structuring elements. As a result, there is a correlation among some characteristics of the structuring elements.

There are two basic morphological operations for binary and grey-scale images including erosion and dilation. Erosion is defined as a shrinking transformation, which reduces the size of regions within the

image, while expanding the size of holes within the regions. As for dilation, it is defined as an expansion transformation, which increases the size of the regions within the image while reducing the size of the holes in the regions and gaps between the regions. It is important to note that the erosion operator filters the inner image, while the dilation operator filters the outer image. Opening and closing morphological operators, which are an extension of erosion and dilation operators are also used, to find specific shapes in an image. Specifically, the opening operation comprises the erosion operation followed by the dilation operation and helps to smooth the contour of an image and eliminate small objects. On the other hand, the closing operation tends to remove small holes and fill gaps in the contours [53]. Note that morphological operations have gained popularity because they are useful for the detection of the edge of an image and suppression of noise.

In this paper, opening and closing morphological operators are applied on grey-scale images, mainly to filter noise [53], while erosion and dilation operations are used for processing image edges. *I(x,y)* is considered as a grey-scale two-dimensional image and *S* is referred as structuring element. The erosion of a grey-scale image *I(x,y)* by a structuring element *S(a,b)* is defined as [52, 54]:

$$I \ominus S = \min\{I(x+a, y+b) - S(a,b)\} \tag{8}$$

The dilation of a grey-scale image, *I(x,y)*, is denoted by

$$I \oplus S = \max\{I(x-a, y-b) + S(a,b)\} \tag{9}$$

Based on the erosion and dilation operators, the opening and closing of the image *I* by the structuring element *S* are respectively defined as follows:

$$I \circ S = (I \ominus S) \oplus I \tag{10}$$

$$I \bullet S = (I \oplus S) \ominus S \tag{11}$$

In this paper, the first step is to select structuring elements which are regarded as matrices and able to measure the shape of the image. In addition, choosing the shape and size of the structuring element is based on the condition and processing demand of the image. In this paper, we used a 5×5 square structuring element to input in the opening and closing morphological operators for filtering. The opened and closed images were then converted to binary images by using thresholds for next features extraction and classification processes.

3.3.2 Local Binary Pattern Operators

The LBP algorithm was introduced by Ojala et al. in 1996 [55]. The LBP operator has been developed to detect textures or objects in images for a long time. It is considered a robust texture descriptor for analysing images, because of its capability to represent plant discriminative information and computational efficiency [55]. It is also one of the best texture descriptors and has been effectively used in various applications. The potentials and effectiveness of LBP have been presented in identifying

objects, recognizing faces and facial expressions and classifying demographics. In this paper, the LBP operator is particularly used for leaf description due to its effectiveness in pattern description.

The main limitation of the previously reported LBP operator was to only cover a small 3×3 neighbourhood, thus failing to capture dominant textural features in images with large-scale structures. To overcome this drawback (i.e., improve the LBP operators), the number of pixels and the radius in the circular neighbourhood have been increased [14]. Typically, it is more flexible and effective to enhance the performance of the LBP method by using textures of different scales. Generally, the value of the LBP code of a centre pixel (x_c, y_c) can be calculated as follows [14]:

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p \quad \text{where} \quad s(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases} \tag{12}$$

where g_c is the grey value of the central pixel and g_p indicates the grey values of the circularly symmetric neighbourhood from $p = 0$ to $P - 1$ and $g_p = x_{P,R,p}$. In addition, P stands for the number of surrounding pixels in the circular neighbourhood with the spatial resolution of the neighbourhood R. Also, $s(x)$ symbolizes the thresholding function, which helps the LBP algorithm to gain illumination invariance against any monotonic transformation. The probability distribution of the 2^p LBP patterns represents the characteristic of the texture image. The mentioned parameters of the LBP algorithm control how patterns are computed for each pixel in input images.

Rotating an image causes diverse LBP codes. Therefore, LBP codes need to rotate back to the position of the reference pixel in order to invalidate the results of translating a pixel location and generate multiple identical versions of binary codes. To address the problem of the image rotation effect, a rotation-invariant LBP has been defined as follows [14, 56]:

$$LBP_{P,R}^{ri} = \min\{ROR(LBP_{P,R}, i) \quad | \quad i = 0, 1, \dots, P - 1\} \tag{13}$$

where the function $ROR(x, i)$ performs an i-step circular bit-wise right shift on the P-bit number x. The rotation invariant LBP is formed by circularly rotating the basic LBP code and keeping the rotationally unique patterns that result in a significant reduction in feature dimensionality.

For uniform patterns, $LBP_{P,R}$refers to the number of spatial transitions in the patterns and the $LBP_{P,R}^{u2}$ patterns need to have at most two bitwise transitions from 0 to 1 or vice versa. As for a given pattern of P bits, the uniform descriptor produces $P(P - 1) + 3$ output bins, which consist of $P(P - 1) + 2$ bins for distinct uniform patterns, and a single bin $(P + 1)$ assigned to all non-uniform patterns. To overcome poor discrimination, due to the crude quantization of angular space at 45° intervals, the rotation invariant uniform descriptor $LBP_{P,R}^{riu2}$, which has a U value of at most 2, is defined as follows [14]:

$$LBP_{P,R}^{riu2} = \begin{cases} \sum_{p=0}^{P-1} s(g_p - g_c), & \text{if } U(LBP_{P,R}) \leq 2 \\ P+1, & \text{if } U(LBP_{P,R}) > 2 \end{cases} \quad (14)$$

The other patterns are marked as "miscellaneous" label and grouped into a single value. To map from $LBP_{P,R}$ to $LBP_{P,R}^{riu2}$, the rotation invariant uniform descriptor has (P+2) distinct output patterns. Correspondently, the $LBP_{8,1}^{riu2}$, $LBP_{16,2}^{riu2}$ and $LBP_{24,3}^{riu2}$ operators have 10, 18 and 26 bins, respectively.

3.3.3 Support Vector Machines (SVM)

After the dominant features are extracted using the LBP method, the next stage is classification. There are several different classification methods, including decision trees, SVM, neural networks, k-nearest neighbour method and the Bayesian classifier. One of the efficient classification methods is SVM, due to its high performance in many applications, such as face recognition [57, 58], weed identification [59, 60] and disease detection in plant leaves [61, 62]. Therefore, the optimal combination of the LBP descriptors and the SVM classifier can lead to high plant discrimination accuracy. Furthermore, the SVM method has become widespread for classifying objects. It is also regarded as an effective and robust supervised classifier due to its capability of dealing with pattern recognition problems in image processing and preventing over-fitting and noise data [63, 64]. SVM was originally introduced in 1992 [65] and then significantly extended by many other researchers. A binary classification SVM was first proposed [66]. A given training dataset of images (x_i, y_i) where $x_i \in \mathbb{R}^d$ for i = 1, 2, 3… N (images) with a label $y_i \in \{-1, 1\}$, the SVM binary classifier $f(x)$ predicts a label y as follows [66]:

$$f(x_i) \begin{cases} \geq 0 & y_i = +1 \\ < 0 & y_i = -1 \end{cases} \quad (15)$$

For example, $y_i f(x_i) > 0$ is considered as a correct classification. The optimization problem solved for binary classification is formulated as follows [65, 67]:

$$\min_{w,b,\xi} = \frac{1}{2} w^T w + C \sum_{i=1}^{l} \xi_i \quad (16)$$

subject to the constraint $y_i(w^T \phi(x_i) + b) \geq 1 - \xi_i$ with $\xi_i \geq 0, i = 1, \ldots, l$

According to Eq. (9), the training data x_i are mapped into a higher dimensional space by the function ϕ and every constraint can be satisfied if ξ_i is sufficiently large. In addition, C > 0 is the regularization parameter, w is known as the weight vector and b is the bias. The SVM method generates an optimal hyperplane with the maximal margin between classes in the higher dimensional space. A kernel function $K(x_i, x_j)$ is represented as $\phi(x_i)^T \phi(x_j)$ and two kernels including polynomial and radial basis function (RBF) are applied in this paper. The polynomial and RBF kernels with kernel parameters γ, r, d are given by [68]

$$\text{Polynomial SVM: } K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0 \quad (17)$$

$$\text{RBF SVM: } K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0 \quad (18)$$

Kernel selection has long been a problem. In this paper, a study is conducted using independent test sets to compare kernels and select the best one.

3.3.4 Data Collection

As mentioned in the article [43], all data was captured on a custom-built testing facility in Figure 1 at ESRI (Electron Science Research Institute), Edith Cowan University, Australia. Particularly, a Xilinx Zynq ZC702 development platform [65] captured HD images (1920×1080 pixels) at 60 frames per second and used an On-Semi VITA 2000 camera sensor. All images captured by the camera had a spatial resolution of ≈1mm/pixel and size of 228×228 pixels, which were down-sampled by a factor of 2 from a size of 456×456 pixels. Moreover, the vertical height of the camera above the surface of the plant pots was 980 mm and the camera focal length was 9mm.

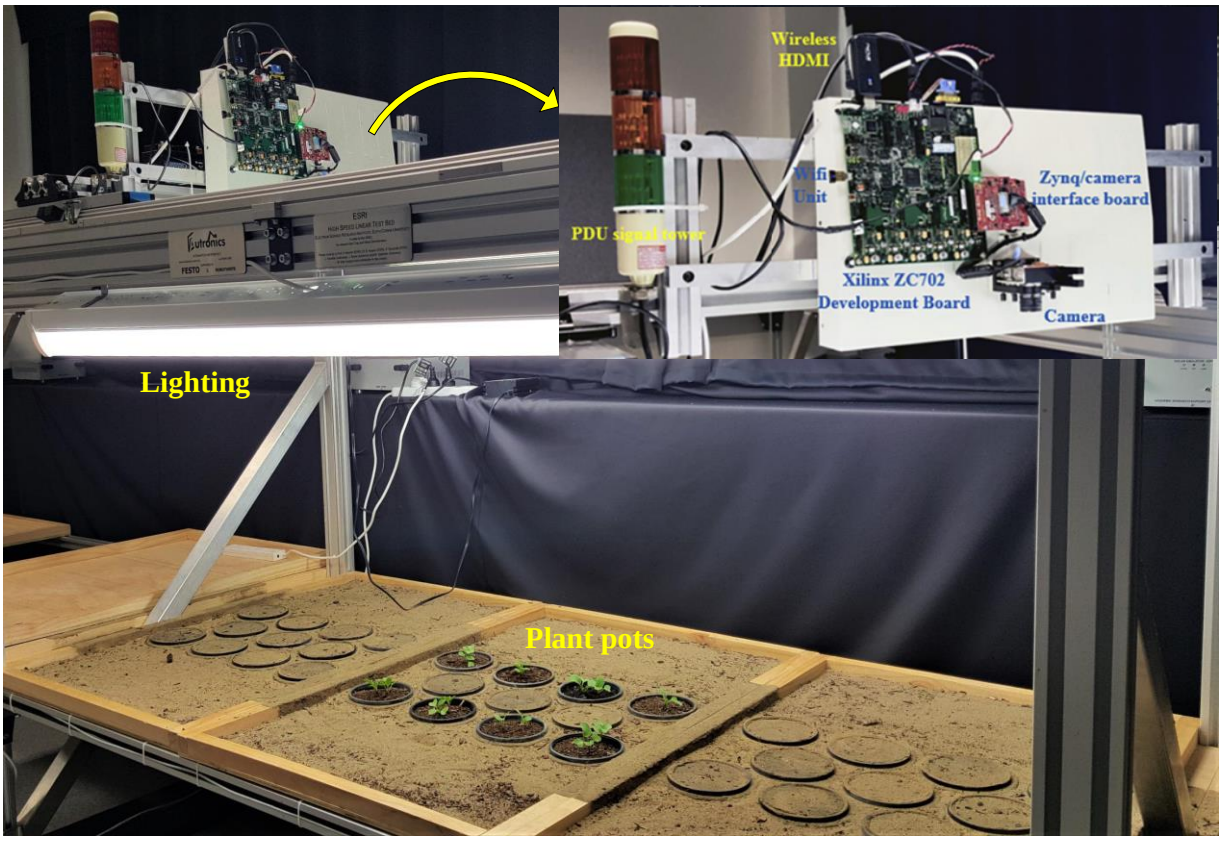


Figure 1. A high-speed testbed system used for controlled data capture [43]. This system has two components including (Plant Discrimination Unit) PDU based on spectral reflectance techniques and a Xilinx Zynq ZC702 development platform.

In this paper, we continue to use the bccr-segset dataset to compare the performance of the novel combination of the LBP algorithm and contoured mask with coefficient k with that of the combined LBP operators reported in [43]. In addition, a new dataset of broadleaf images including only canola and radish leaves is captured to objectively evaluate the detection capability of the proposed approach.

3.4 Methodology

In the previous paper [43], three different LBP operators $LBP_{8,1}^{riu2}$, $LBP_{16,2}^{riu2}$ and $LBP_{24,3}^{riu2}$ and the SVM method were combined to detect and classify broadleaf and narrow-leaf plants. The results confirmed that the classification accuracies between broad and narrow leaves were higher than the ones between broadleaf groups. The recognition of leaves is based on the observation of their morphological features such as texture and shape. According to our "bccr-segset" dataset, canola and radish plants belong to the broadleaf group, develop as a rosette and have lobes. However, there are some differences between leaf shapes on the canola and radish plants. When the edge of each leaf is observed closely at the third stage in Figure 2, canola leaves have outward-pointing teeth and radish leaves have a rounded shape with curved-toothed edge. In other words, from the glossary of leaf morphology, the leaf margin of canola is sinuate while the edge of radish is undulate with a wavy edge, shallower than sinuate [69]. For canola leaves at the fourth growth stage, their lobes are often completely separated towards the base of the leaf. With regard to older radish leaves, they have larger rounded lobe at the tip of the leaf, some pairs of side lobes and each set is progressively smaller toward the base.

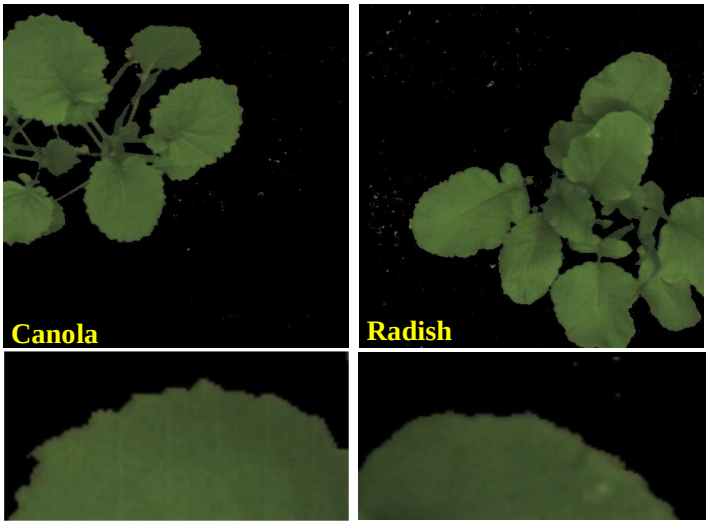


Figure 2. Full and zoomed-in images of canola and radish leaves in the third stage.

To overcome the limitation of the combined LBP operators in the previous paper, a novel method has been developed for amplifying the dominant features of canola and radish leaves. The flowchart below describes this method in detail.

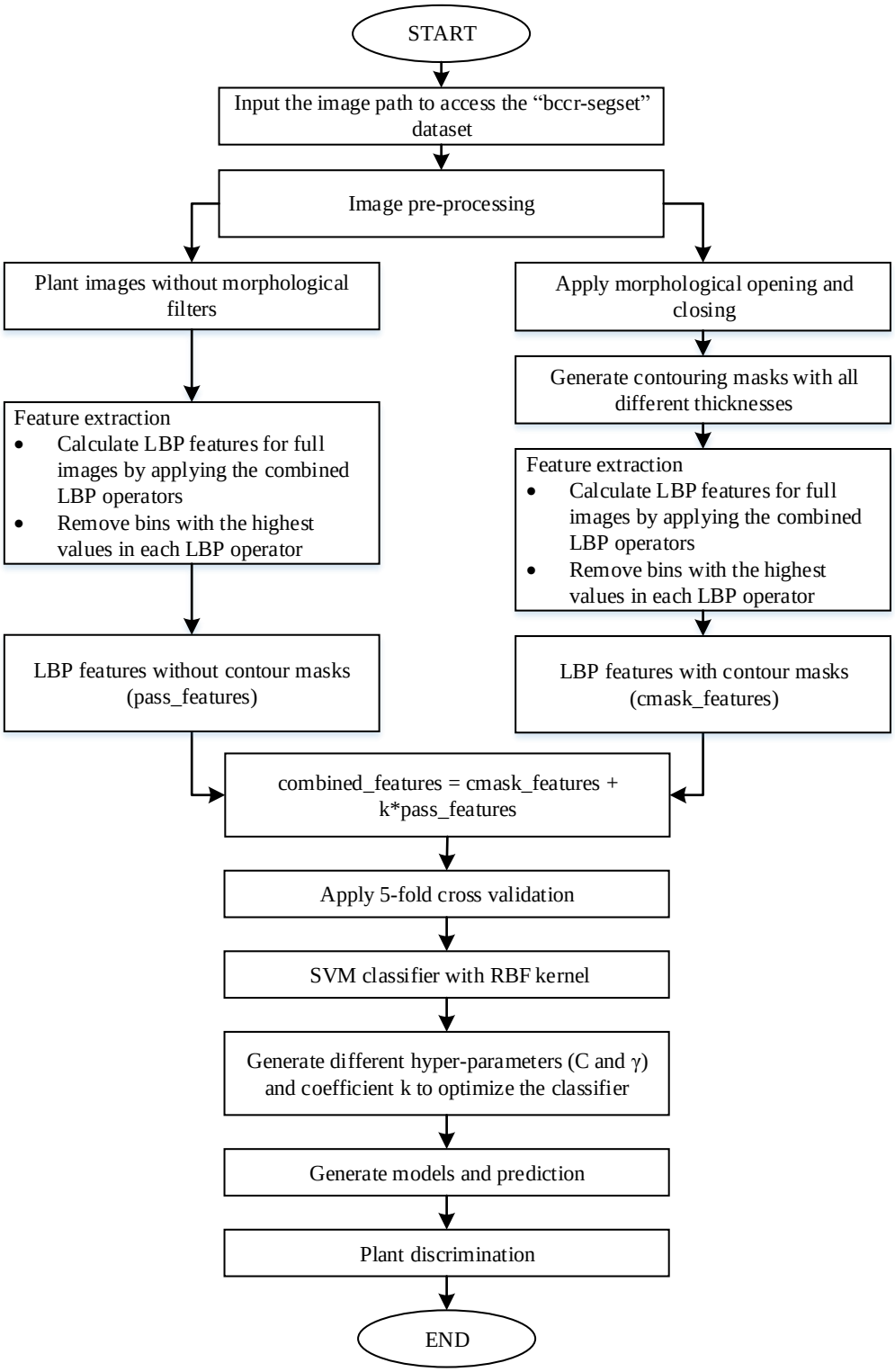


Figure 3. A flowchart describing the procedures of the novel method through steps, namely, filtering LBP bins, extracting features, masking images based on contours and classifying plant leaves.

To begin with, we input the "bccr-segset" dataset into the plant classification program. The dataset was processed in two branches: (i) the dataset was input to the feature extraction block without applying the morphological operations, and (ii) the dataset applied the morphological opening and closing, and generated contour masks with different thicknesses as shown in Figure 3. To be more specific in the second branch, a 5x5 morphological filter was created to implement the morphological opening and closing on all plant images in the dataset. By selecting a threshold, grayscale images were converted into binary images to get better accuracy. Here, we masked all plant images with contours, i.e., boundaries around selected plant images. The *findContours* function and *drawContours* function in OpenCV were used, and then all the masks of plant images of different thicknesses were stored. This eliminates the need to recalculate when the thickness was changed.

The next stage of both branches was going through the feature extraction block. Particularly, LBP features were computed for full images in the mentioned dataset by incorporating $\mathrm{LBP}_{8,1}^{\mathrm{riu2}} + \mathrm{LBP}_{16,2}^{\mathrm{riu2}} + \mathrm{LBP}_{24,3}^{\mathrm{riu2}}$ operators, which are accumulated into a histogram of P+2 bins (with P=8, 16, 24 corresponding to each LBP operator). Each bin denotes an estimate of the probability of encountering the corresponding pattern in the plant image. The discrete histograms of the $\mathrm{LBP}_{\mathrm{P,R}}^{\mathrm{riu2}}$ operators were calculated over plant images. Note that it is not necessary for all bins in the LBP histogram to contain useful information for plant leaf detection. It is observed that for the LBP histograms of plant images at the bin level, the 9th bin of $\mathrm{LBP}_{8,1}^{\mathrm{riu2}}$, the 17th bin of $\mathrm{LBP}_{16,2}^{\mathrm{riu2}}$ and the 25th bin of $\mathrm{LBP}_{24,3}^{\mathrm{riu2}}$ contain a much higher number of hits when compared to the remaining bins from the LBP histogram. A further investigation shows that the LBP values for these bins correspond to patterns which have no pixel variations. For example, all pixels are constant values such as the values of background pixels. However, the remaining bins correspond to LBP patterns which mainly capture the intensity variations of green pixels (plant leaves). Therefore, bins P+1 (the 9th bin of $\mathrm{LBP}_{8,1}^{\mathrm{riu2}}$, the 17th bin of $\mathrm{LBP}_{16,2}^{\mathrm{riu2}}$ and the 25th bin of $\mathrm{LBP}_{24,3}^{\mathrm{riu2}}$) were removed from each LBP histogram in order to better scale the remaining bins. According to the combination of three different spatial resolutions and different angular resolutions in LBP operators, three bins including 9th, 27th and 53rd were removed in the joint histogram of $\mathrm{LBP}_{8,1}^{\mathrm{riu2}} + \mathrm{LBP}_{16,2}^{\mathrm{riu2}} + \mathrm{LBP}_{24,3}^{\mathrm{riu2}}$ operator (10 bins +18 bins + 26bins = 54 bins). After applying the $\mathrm{LBP}_{8,1}^{\mathrm{riu2}} + \mathrm{LBP}_{16,2}^{\mathrm{riu2}} + \mathrm{LBP}_{24,3}^{\mathrm{riu2}}$ operator for the plant images, the resultant images were called as LBP images.

As can be seen in Figure 4, it illustrates an example of the process shown in the flowchart (Figure 3). In Figure 4 a) we show an original canola leaf image and its three histograms corresponding to $\mathrm{LBP}_{8,1}^{\mathrm{riu2}}, \mathrm{LBP}_{16,2}^{\mathrm{riu2}}$ and $\mathrm{LBP}_{24,3}^{\mathrm{riu2}}$ operators. The 9th, 17th and 25th bins in each operator have the highest level of the distribution of patterns. The LBP-based canola leaf image and contour mask, the original histogram and the filtered histogram of the contour masks are shown in Figure 4 b), c), d) with the

$LBP_{8,1}^{riu2}$, $LBP_{16,2}^{riu2}$ and $LBP_{24,3}^{riu2}$ operators, respectively. It is apparent that the feature distribution is easily observed in the other bins of the LBP histogram with bin removal. Interestingly, dominant features such as edge and corner patterns in other bins can be seen clearly by removing some specific bins (9th, 17th, and 25th bins) in the LBP histograms. Similarly, plant features in the histogram of the LBP based contour mask with bin removal also present their significance. It is noted that the bin number of the LBP histogram in Figure 4, calculated in a Python code, has an index range from 0 to [(P+2) - 1] bins. Note that the bin number mentioned in this paper starts from 1 to P+2. For example, the $LBP_{8,1}^{riu2}$ operator has an index range from 0 to 9 but the bin number from 1 to 10.

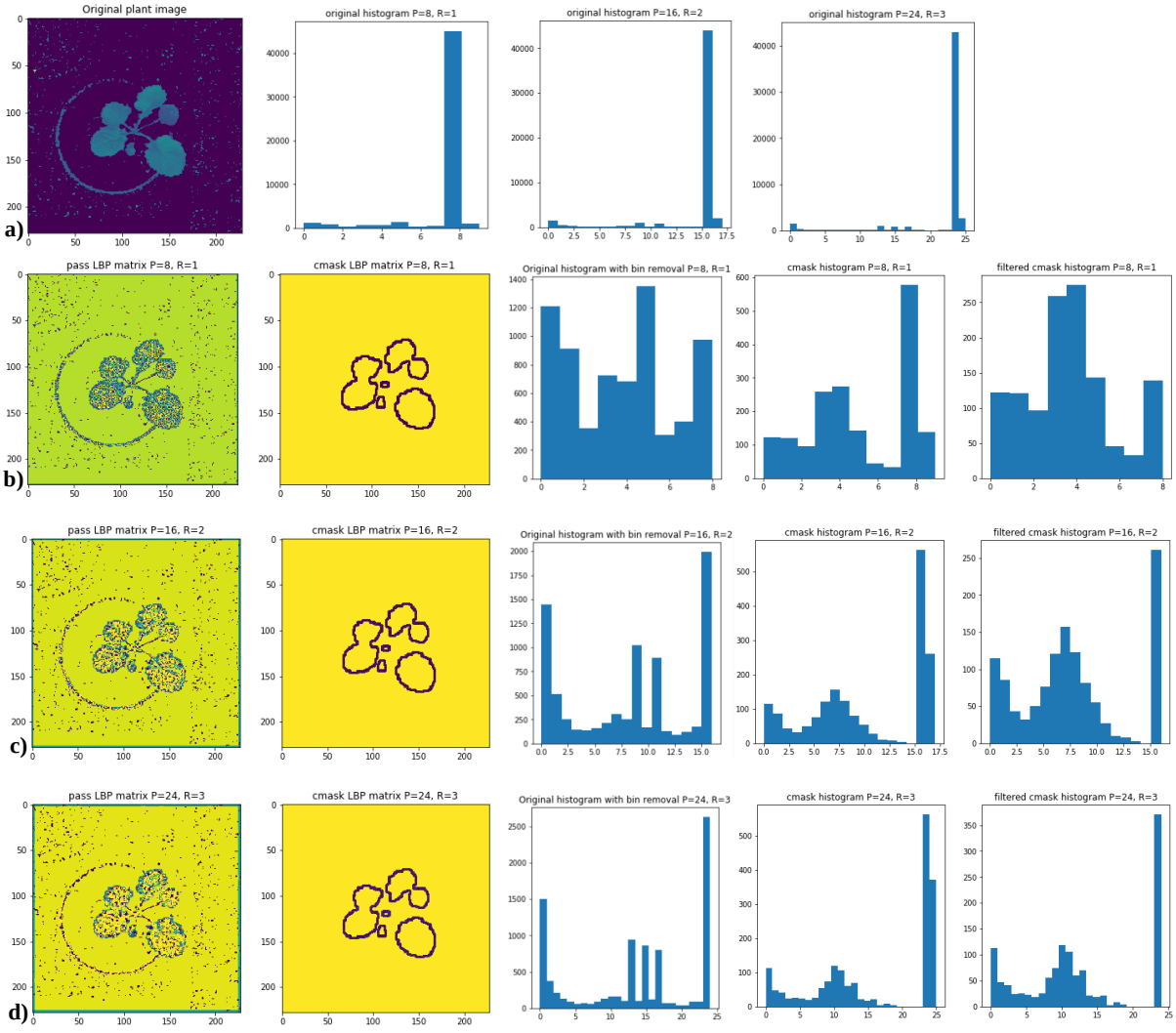


Figure 4. a) An original canola leaf image and its LBP histograms corresponding to $LBP_{8,1}^{riu2}$, $LBP_{16,2}^{riu2}$ and $LBP_{24,3}^{riu2}$ operators, (**b-d)** LBP images, LBP images with contour masks, and their original LBP histograms and filtered LBP histograms are presented by implementing $LBP_{8,1}^{riu2}$, $LBP_{16,2}^{riu2}$ and $LBP_{24,3}^{riu2}$ operators, respectively. Multiresolution analysis can be achieved by altering P and R of LBP operators and then combining these operators as shown in Figure 5.

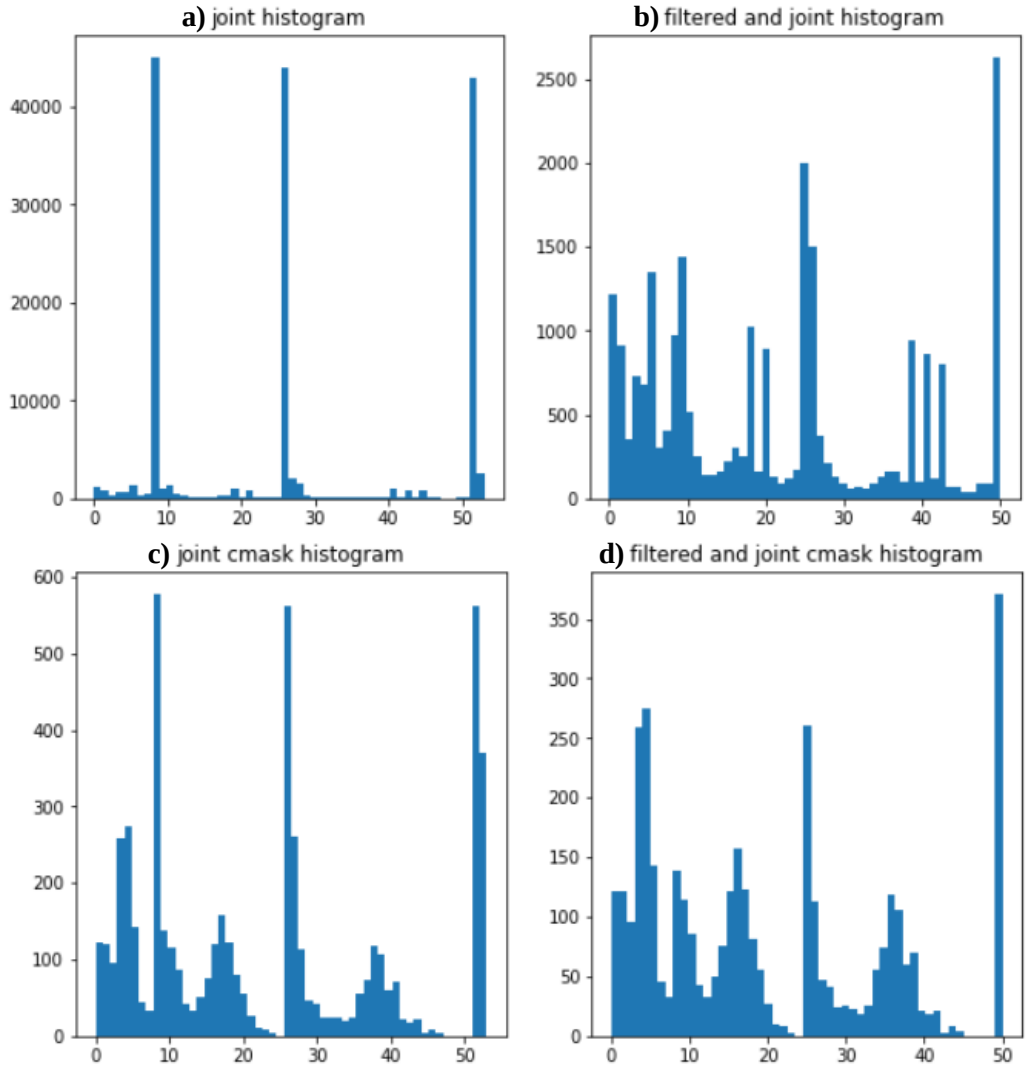


Figure 5. Four different LBP histograms of a canola leaf image. a) Combining three operators ($LBP_{8,1}^{riu2}, LBP_{16,2}^{riu2}$ and$LBP_{24,3}^{riu2}$). b) A filtered and joint histogram is generated by eliminating 9th, 27th and 53rd bins in the joint histogram. c) A joint cmask histogram is generated by applying the LBP method with a contour mask. d) Removing 9th, 27th and 53rd bins in the joint cmask histogram.

As shown in Figure 3, the filtered LBP features without contour mask in plant images are denoted as ***pass_features***. The method used to generate images is referred to as Filtered LBP method (FLBP). The FLBP method is applied to the plant images, and results in 51 features are calculated over the entire image. The FLBP based contour masks are denoted as ***cmask_features***. The method used to create images consisting of cmask_features is referred to as the Filtered LBP based Contour Mask (FLBPbCM). Applying the FLBPbCM method to the plant images also results in 51 features computed only on the contours. The remaining region in the image is set to the maximum value (255) in the LBP matrix and ignored when generating the LBP histogram.

The novelty of the current ***k-FLBPCM method*** filtered LBP method with Contour Mask and coefficient k (k-FLBPCM) is a combination of ***pass_features*** and ***cmask_features***. Due to the high bin values in

the FLBP method as shown in Figure 4 and Figure 5, cmask_features are scaled by multiplying pass_features by coefficient k in k-FLBPCM method. For example, Table 1 shows the distributions of patterns (bin values) in a typical canola image. It demonstrates that by combining the pass_features (in FLBP method) and cmask_features (in FLBPbCM method), the bin values of the k-FLBPCM method have better balance between these two feature sets. The purpose of multiplying coefficient k ($k \leq 1$) with pass_features is to reduce the gap between the bin values of the cmask_features and pass_features.

Table 1. The bin values of a typical canola image using FLBP, FLBPbCM and the combined k-FLBPCM methods

Bin values of different methods	**Bin 1**	**Bin 2**	**Bin 3**	**Bin 4**	**Bin 5**	**Bin 6**	**Bin 7**	**Bin 8**	**Bin 10**
FLBP	1212	913	355	727	680	1351	305	402	974
FLBPbCM	122	121	96	259	275	143	45	33	139
k-FLBPCM with k=0.5	728	577.5	273.5	622.5	615	818.5	197.5	234	626

After the feature extraction step, the plant images are classified by using SVM kernels. Initially, 5-fold-cross validation was used to divide the dataset into five subsets. Due to the different plant growth stages in the dataset, images at each growth stage are equally divided in each subset as well. A single subset of the dataset is used for testing while the remaining four subsets of the dataset are used for training. The cross-validation process was iteratively applied five times, with the test subset changed each time. This procedure helps to prevent overfitting. After generating the training model by selecting RBF kernel in SVM and making predictions, the classification accuracies of the methods was calculated by using the performance metrics such as accuracy, precision, recall and F1-score.

3.5 Results and Discussions

The results are divided into two sections: the first section presents the average classification accuracies of the broadleaf classes consisting of canola and radish. The effectiveness of the proposed k-FLBPCM method is evaluated based on factors including feature extraction (by comparing among the FLBP, FLBPbCM, and k-FLBPCM methods), different SVM kernels (the second order polynomial kernel and RBF kernel), contour thickness, LBP parameters P (the total number of the neighbouring pixels) and R (the radius) as well as the coefficient k. In the second section, the parameters (C, Gamma (γ), coefficient k and thickness) for the classification of all four classes in the "bccr-segset" dataset including canola, corn, radish and background are optimized to obtain improved classification accuracy. The computer used in these experiments had a 3.4GHz processor, 16GB RAM and ran Python 2.7.13.

3.5.1 Results of the k-FLBPCM, FLBPbCM and FLBP methods in classifying two different broadleaf plants

Canola and radish images were taken from the "bccr-segset" dataset. The train and test sets of canola and radish classes consist of 15000 images (7500 images in each class). After applying the FLBP, FLBPbCM, or k-FLBPCM methods, SVM was used to classify the two broadleaf classes including canola and radish plants. The classification accuracies of the second order polynomial kernel and the RBF kernel were compared. In this experiment, C = 10, 60, $\gamma = 10^{-5}, 10^{-6}$ and thickness =2 were selected. The values of C and γ selected were typical values, before any optimization had been performed.

The results of using two SVM kernels (the second order polynomial and RBF kernels) on the given dataset for classification are summarised in Table 2. In particular, the average classification accuracy of the k-FLBPCM method (C=10, $\gamma = 10^{-5}$, k=0.5 and 0.2) with the RBF kernel was 97.32%, followed by 96.40% corresponding to k-FLBPCM method with coefficient k=0.1. Meanwhile, the average classification accuracy of the k-FLBPCM method (C=10, $\gamma = 10^{-5}$, k=0.5) with the second order polynomial kernel was just 95.46%. Similarly, the case (C=60, $\gamma = 10^{-6}$) of the k-FLBPCM method with the RBF kernel was also higher than the polynomial kernel of degree 2. In addition, the FLBP method with the RBF kernel had higher classification rate than the polynomial kernel. As for the FLBPbCM method (C=10, $\gamma = 10^{-5}$), the RBF kernel had the classification accuracy of 94.07% in comparison to the second order polynomial kernel at 88.53%. These results show that the RBF kernel, which nonlinearly maps features into a higher dimensional space, resulting in higher classification accuracy for all three methods (FLBP, FLBPbCM and k-FLBPCM methods).

Table 2. The average classification accuracy score of the k-FLBPCM, FLBPbCM and FLBP methods with the second order polynomial and RBF kernels.

C	γ	Thickness	Methods	Accuracy Score	
				Polynomial kernel of degree 2	RBF kernel
10	1E-05	2	k-FLBPCM method, k=0.5	95.46%	97.32%
10	1E-05	2	k-FLBPCM method, k=0.2	94.91%	97.32%
10	1E-05	2	k-FLBPCM method, k=0.1	94.27%	96.40%
60	1E-06	2	k-FLBPCM method, k=1	94.92%	97.50%
60	1E-06	2	k-FLBPCM method, k=0.5	94.56%	96.89%
60	1E-06	2	k-FLBPCM method, k=0.2	93.55%	96.06%
10	1E-05	No thickness	FLBP method	93.53%	95.36%
60	1E-06	No thickness	FLBP method	93.74%	96.72%
10	1E-05	2	FLBPCM method	88.53%	94.07%
60	1E-06	2	FLBPCM method	88.26%	94.83%

A second experiment was conducted to investigate the effects of the hyper-parameters C and γ, as well as the coefficient k on the classification accuracy of canola and radish images. Various pairs of (C, γ) values were tried and good results were obtained with exponentially growing sequences of C and γ [70]. Therefore, we chose the ranges of C, γ and coefficient k as follows: C =1, 10, 30, 60, 100, 1000, γ = $10^{-4}, 10^{-5}, 10^{-6}, 10^{-7}$. In addition, as mentioned in the method section, we selected k (k≤1) randomly from 0.1 to 1 (k =0.1, 0.2, 0.5, 0.7, 0.8, and 1.0). We tested all these values in the experiments in order to observe the variation of values and chose an optimal set k, C and Gamma when these parameters reach the highest classification accuracy. As shown in

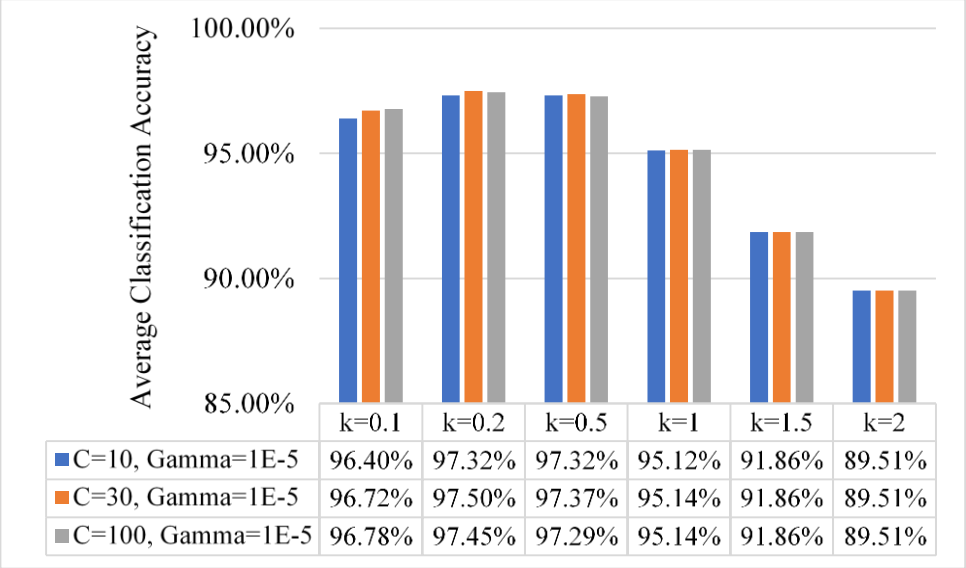
Table **3**, the k-FLBPCM method had the highest classification accuracy, averaged over the 5-folds of the cross validation, in the first pair (C=30, γ =10^{-5}, thickness=2, k=0.2) and the second pair (C=60, γ =10^{-6}, thickness=2, k=1), at 97.50%. In addition, the average classification accuracies of the k-FLBPCM method with different parameters were sorted from high to low. Due to the large number of combinations possible, only the top 10 cases are listed in

Table **3**. Due to the low accuracy of using γ =10^{-4}, the parameter γ should be less than 10^{-5} to improve the classification accuracy of the k-FLBPCM method.

Table 3. The average accuracy scores of the k-FLBPCM method with the RBF kernel, varying C, γ and the coefficient k.

C	γ	Thickness	k-FLBPCM method	Accuracy score
30	1E-05	2	k=0.2	97.50%
60	1E-06	2	k=1	97.50%
60	1E-05	2	k=0.2	97.49%
100	1E-05	2	k=0.2	97.45%
100	1E-06	2	k=1	97.42%
30	1E-06	2	k=1	97.42%
100	1E-06	2	k=0.7	97.40%
30	1E-05	2	k=0.5	97.37%
100	1E-06	2	k=0.8	97.35%
60	1E-06	2	k=0.8	97.34%

Although all experiments were conducted with different coefficients k, this parameter should be less than or equal to 1. We find that (k ≤ 1) results in optimal accuracy. As shown in Figure 6, the average classification accuracies of the proposed k-FLBPCM method with k ≤ 1 were higher than the ones with k>1.



	k=0.1	k=0.2	k=0.5	k=1	k=1.5	k=2
C=10, Gamma=1E-5	96.40%	97.32%	97.32%	95.12%	91.86%	89.51%
C=30, Gamma=1E-5	96.72%	97.50%	97.37%	95.14%	91.86%	89.51%
C=100, Gamma=1E-5	96.78%	97.45%	97.29%	95.14%	91.86%	89.51%

Figure 6. The average classification accuracies of the k-FLBPCM method with different coefficients k.

3.5.1.1 Comparing the FLBP, FLBPbCM, and k-FLBPCM methods

To check the effectiveness of the k-FLBPCM method in a different dataset, a new set of canola and radish images in four different growth stages was collected and designated "can-rad" dataset (published online). A total of 19600 broadleaf images (9800 images in each class) were collected at four different growth stages. The parameters C =10, 30, 60, 100, 1000, $\gamma = 10^{-5}, 10^{-6}$, and thicknesses from 1 to 8 were selected. Note that the SVM classifier was used only with the RBF kernel in the remaining parts of experiments. Further, only the 10 highest classification accuracies for each method are listed in Tables 3.3-5 and the average classification accuracy scores are sorted from high to low.

Table 4. Classification accuracy of the FLBP method.

C	γ	Classification Accuracy of the FLBP method
100	**1E-06**	**95.13%**
1000	1E-06	95.03%
60	1E-06	94.96%
30	1E-06	94.92%
10	1E-06	94.31%
1000	1E-07	93.92%
10	1E-05	93.78%
30	1E-05	93.67%

60	1E-05	93.62%
100	1E-05	93.61%

Table 5. Classification accuracy of the FLBPbCM method.

C	γ	Thickness	Classification Accuracy of the FLBPbCM method
30	**1E-05**	**8**	**93.95%**
30	**1E-05**	**7**	**93.95%**
100	1E-05	2	93.94%
30	1E-05	6	93.88%
30	1E-05	5	93.88%
10	1E-05	8	93.88%
10	1E-05	7	93.88%
1000	1E-05	2	93.87%
60	1E-05	6	93.87%
100	1E-05	6	93.81%

As can be seen from Table 4 and Table 5, the classification accuracy of the FLBP method was 95.13% with C =100 and $\gamma = 10^{-6}$, while that of the FLBPbCM method was 93.95%, lower than the FLBP method. However, when combining the FLBP and FLBPbCM methods (in k-FLBPCM method), the classification accuracy was significantly higher. Table 6 shows that the highest average classification accuracy of the k-FLBPCM method was 96.21%.

Table 6. Classification accuracy of the k-FLBPCM method

C	γ	Thickness	k-FLBPCM method	Classification Accuracy
1000	**1E-06**	**2**	**k=0.5**	**96.21%**
30	1E-05	2	k=0.5	96.19%
10	1E-05	2	k=0.5	96.18%
30	1E-05	4	k=0.5	96.16%
30	1E-05	3	k=0.5	96.16%
60	1E-05	2	k=0.5	96.15%
10	1E-05	4	k=0.5	96.14%
10	1E-05	3	k=0.5	96.14%
30	1E-05	2	k=0.2	96.13%

30	1E-05	4	k=0.2	96.11%

3.5.1.2 Effects of the contour thickness on the classification accuracy

Next, we evaluated the average classification accuracy of the k-FLBPCM method for varying the thicknesses of the contour lines. The "can-rad" dataset was used for this investigation. We selected C=10, 30, 100, $\gamma = 10^{-5}$, coefficient k = 0.5 and thickness from 1 to 8. As can be seen in Figure 7, two images of canola and radish with varying contour thickness are presented at the third growth stage.

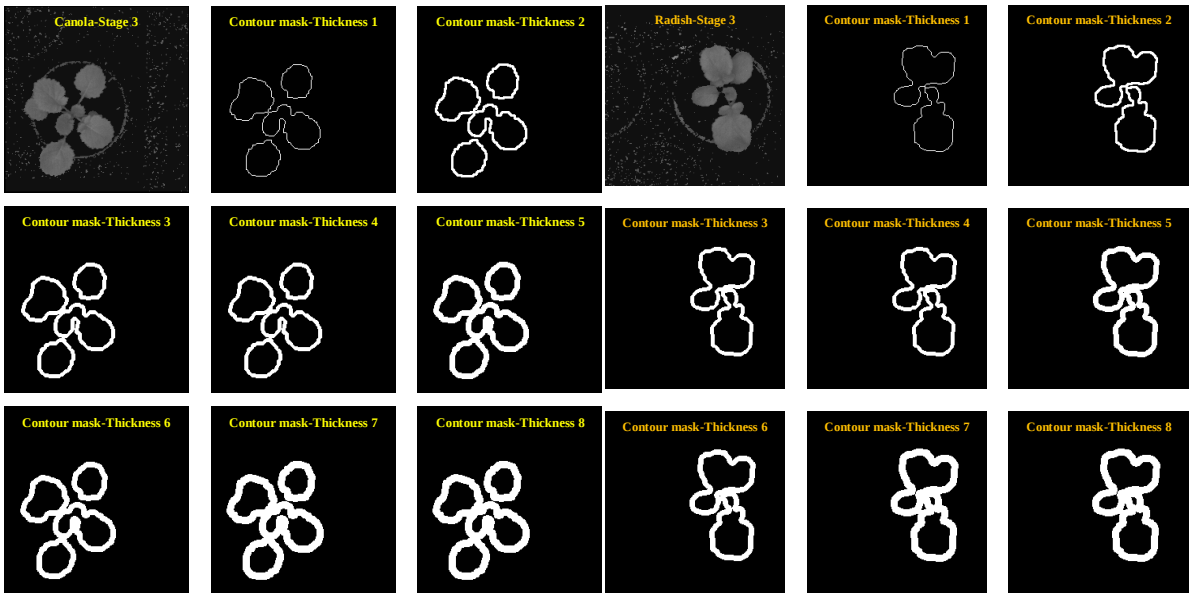
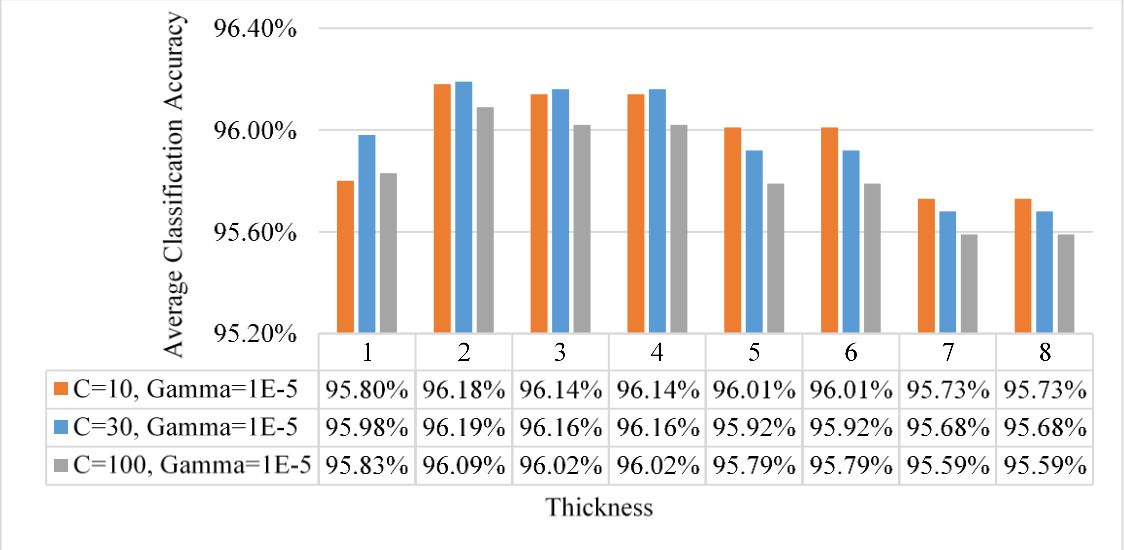


Figure 7. Canola and Radish at the third stage with varying thicknesses of the contour lines.

The average classification accuracies of the k-FLBPCM method for different thicknesses are reported in Figure 8. Our proposed k-FLBPCM method attained optimal discrimination between canola and radish at contour thickness of 2 with the accuracy of 96.19% (C=30, $\gamma = 10^{-5}$), while the lowest accuracy was 95.73% with thicknesses of 7 and 8. These two broadleaf plants displayed morphological similarity at a contour thickness of 2. As shown in Figure 7, for the thickness greater than 2, the leaf features were smoothed by the thick edge, while for the thickness of 1, the edge features were too thin to fully show the difference between the undulate and sinuate patterns.



	1	2	3	4	5	6	7	8
C=10, Gamma=1E-5	95.80%	96.18%	96.14%	96.14%	96.01%	96.01%	95.73%	95.73%
C=30, Gamma=1E-5	95.98%	96.19%	96.16%	96.16%	95.92%	95.92%	95.68%	95.68%
C=100, Gamma=1E-5	95.83%	96.09%	96.02%	96.02%	95.79%	95.79%	95.59%	95.59%

Figure 8. The average classification accuracies of the k-FLBPCM method (coefficient k=0.5) for different contour line thicknesses and four growth stages

3.5.2 Classification capabilities of the k-FLBPCM, FLBPbCM and FLBP methods

The k-FLBPCM method was evaluated on the full "bccr-segset" dataset, which included 30,000 plant images in four classes (canola, corn, radish and background) under different rotations, scales and illumination conditions. Plant images were taken under different rotation angles (45°, 90°, 135°, 180°, 225°, 270°, 315°, 360°), lighting conditions (sunlight and fluorescent), sizes and morphologies of plants through four growth stages, as illustrated in Figure 9. The number of plant images at each class and each growth stage is indicated in Figure 9 [43].

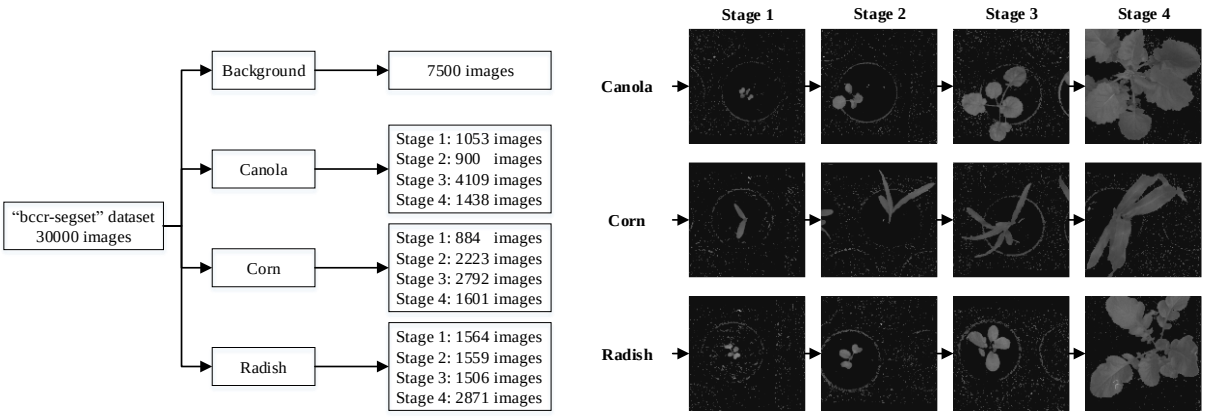


Figure 9. The "bccr-segset" dataset and its four-growth stages.

The average classification accuracies of the FLBP, FLBPCM and k-FLBPCM methods are listed in Table 7. Note that, in this investigation, the following typical values were selected: C=30, 60, 100 and

γ=10^{-5}, 10^{-6}. The k-FLBPCM method again achieved the highest accuracies among all compared methods, confirming the results in the given "can-rad" dataset.

Table 7. Comparison of the average classification accuracies of the FLBP, FLBPCM and k-FLBPCM methods.

C	γ	Thickness	Method	Accuracy score
30	1E-05	2	k-FLBPCM, k=0.2	98.63%
60	1E-05	2	k-FLBPCM, k=0.2	98.61%
100	1E-06	2	k-FLBPCM, k=0.8	98.61%
30	1E-05	No thickness	FLBP	97.23%
60	1E-05	No thickness	FLBP	97.22%
100	1E-06	No thickness	FLBP	98.17%
30	1E-05	2	FLBPCM	97.04%
60	1E-05	2	FLBPCM	97.14%
100	1E-06	2	FLBPCM	96.01%

In order to find optimal (C, γ) pairs, we investigated the following parameter ranges: C = 1, 10, 30, 60, 100, 1000, $\gamma = 10^{-5}, 10^{-6}$, k = 0.1, 0.2, 0.5, 0.8, 1 and thickness of 2. Only the 10 highest classification accuracies of the k-FLBPCM method were listed in Table 8. This method attained the highest classification accuracy of 98.63% with C =30, $\gamma = 10^{-5}$ and coefficient k=0.2.

Table 8. Average classification accuracies of the k-FLBPCM method for different C and γ parameters and coefficients k.

C	γ	Thickness	k-FLBPCM Method	Accuracy score
30	1E-05	2	k=0.2	98.63%
100	1E-06	2	k=0.8	98.61%
100	1E-05	2	k=0.2	98.61%
60	1E-05	2	k=0.2	98.61%
100	1E-06	2	k=1	98.60%
60	1E-06	2	k=0.8	98.58%
60	1E-06	2	k=1	98.57%
1000	1E-06	2	k=0.5	98.56%
30	1E-06	2	k=1	98.56%
1000	1E-06	2	k=1	98.51%

The k-FLBPCM method can classify plant images with different conditions, as shown in our two datasets, and improve the classification accuracies achieved previously [43]. Particularly, there is a significant improvement in performance when combining LBP features with a contour based mask. The average classification accuracies of the k-FLBPCM method have increased over the previously described method by up to 6.78% [45].

The F1-score results for each class are indicated in Table 9. Particularly, the F1 scores of the k-FLBPCM method significantly increased to 97.40% and 97.40% for canola and radish, from 84.41% and 83.43% respectively, which had used combined LBP operators in the previously published paper [45]. In addition, the testing time (millisecond/image) of the k-FLBPCM method was faster than the combined LBP method [45].

Table 9. Comparison of performance metrics between the k-FLBPCM and combined LBP methods for each class.

Method	**SVM kernel**	**Classes**	**Precision**	**Recall**	**F1-score**	**Testing time (ms/image)**
k-FLBPCM	RBF kernel	Background	100%	100%	100%	0.491
		Canola	96.80%	97.60%	97.40%	
		Corn	100%	100%	100%	
		Radish	97.60%	97.20%	97.40%	
Combined LBP operators LBP (8,1) +LBP (16,2) +LBP (24,3)	RBF kernel	Background	96.17%	98.87%	97.50%	1.419
		Canola	83.64%	85.20%	84.41%	
		Corn	98.64%	96.87%	97.75%	
		Radish	84.69%	82.27%	83.46%	

With the aim of reducing the misclassification, we investigated the misclassified images through visual inspection as shown in Figure 10. The first stage plants Figure 10 (a), (b) and (c), appear to have been misclassified due to the close morphological similarities. In addition, deformity of the leaves and stems, especially arising from perspective distortions Figure 10 (e) (f) and leaf diseases Figure 10 (d) can also lead to the identification errors. However, the k-FLBPCM method considerably reduced the number of misclassified images and outperformed other methods by obtaining the high classification accuracy at 98.63%.

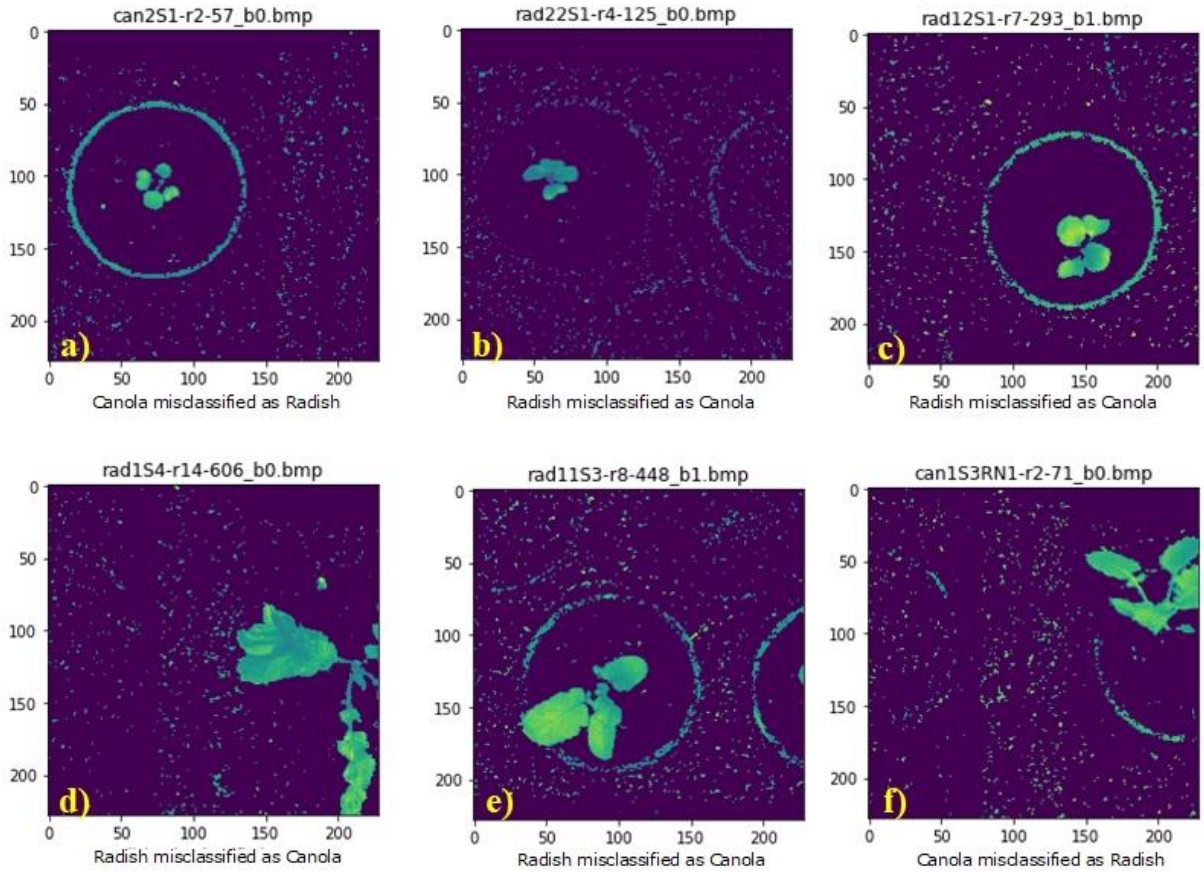


Figure 10. Misclassified images are printed from the model of the k-FLBPCM method with C =30, γ =10^{-5} and k=0.2

3.6 Conclusions

In this paper, the k-FLBPCM method combining LBP feature extraction with contour masks has been proposed for reducing the noise and improving the plant classification accuracy. Results have shown that various factors can reduce the weed identification accuracy, including outdoor scene complexity and morphological variability of plants. Based on the experimental results, the k-FLBPCM method had the best performance of 98.63% accuracy in identifying similar morphological plants. This method is particularly useful to discriminate between two classes with highly similar morphologies, while tolerating morphological variability within each class. Further, results have shown that the execution time of the proposed method is faster than the combined LBP method in the previous published paper. As a result, the proposed method helps to improve the plant classification with similar morphological features. Furthermore, the fast processing time of this method enhances the ability to implement the plant detection in the real time.

Future research might consider the potential of the k-FLBPCM method in diverse applications in order to identify objects of similar morphologies. Morphological cell analysis plays an significant role in supporting pathologists to accurately detect cancer cells [71, 72]. The advantages of the k-FLBPCM

method is that image data can be reused for extracting morphological features and identifying abnormal cells.

3.7 References

[1] D. Slaughter, D. Giles, and D. Downey, "Autonomous robotic weed control systems: A review," *Computers and electronics in agriculture,* vol. 61, no. 1, pp. 63-78, 2008.

[2] Q. Yu, X. Q. Zhang, A. Hashem, M. J. Walsh, and S. B. Powles, "ALS gene proline (197) mutations confer ALS herbicide resistance in eight separated wild radish (Raphanus raphanistrum) populations," *Weed Science,* vol. 51, no. 6, pp. 831-838, 2003.

[3] M. J. Walsh, S. B. Powles, B. R. Beard, B. T. Parkin, and S. A. Porter, "Multiple-herbicide resistance across four modes of action in wild radish (Raphanus raphanistrum)," *Weed science,* vol. 52, no. 1, pp. 8-13, 2004.

[4] R. E. Blackshaw, D. Lemerle, R. Mailer, and K. R. Young, "Influence of wild radish on yield and quality of canola," *Weed Science,* vol. 50, no. 3, pp. 344-349, 2002.

[5] M. Robertson, J. Whish, and F. Smith, "Simulating competition between canola and wild radish," *12th Australian research assembly on Brassicas,* pp. 2-5, 2001.

[6] P. Hansen and J. Schjoerring, "Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression," *Remote sensing of environment,* vol. 86, no. 4, pp. 542-553, 2003.

[7] K. Thorp and L. Tian, "A review on remote sensing of weeds in agriculture," *Precision Agriculture,* vol. 5, no. 5, pp. 477-508, 2004.

[8] A. Huete, K. Didan, T. Miura, E. P. Rodriguez, X. Gao, and L. G. Ferreira, "Overview of the radiometric and biophysical performance of the MODIS vegetation indices," *Remote sensing of environment,* vol. 83, no. 1, pp. 195-213, 2002.

[9] M. Ozdogan, Y. Yang, G. Allez, and C. Cervantes, "Remote sensing of irrigated agriculture: Opportunities and challenges," *Remote sensing,* vol. 2, no. 9, pp. 2274-2304, 2010.

[10] L. W. Lass, T. S. Prather, N. F. Glenn, K. T. Weber, J. T. Mundt, and J. Pettingill, "A review of remote sensing of invasive weeds and example of the early detection of spotted knapweed (Centaurea maculosa) and babysbreath (Gypsophila paniculata) with a hyperspectral sensor," *Weed Science,* vol. 53, no. 2, pp. 242-251, 2005.

[11] M. Pietikäinen, A. Hadid, G. Zhao, and T. Ahonen, *Computer vision using local binary patterns*. Springer Science & Business Media, 2011.

[12] Z. Guo, L. Zhang, and D. Zhang, "A completed modeling of local binary pattern operator for texture classification," *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1657-1663, 2010.

[13] M. Heikkilä, M. Pietikäinen, and C. Schmid, "Description of interest regions with local binary patterns," *Pattern recognition*, vol. 42, no. 3, pp. 425-436, 2009.

[14] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 7, pp. 971-987, 2002.

[15] S. Happy, A. George, and A. Routray, "A real time facial expression classification system using Local Binary Patterns," in *Intelligent Human Computer Interaction (IHCI), 4th International Conference*, 2012: IEEE, pp. 1-5.

[16] O. Lahdenoja, J. Poikonen, and M. Laiho, "Towards understanding the formation of uniform local binary patterns," *ISRN Machine Vision*, 2013.

[17] M. Heikkila and M. Pietikainen, "A texture-based method for modeling the background and detecting moving objects," *IEEE transactions on pattern analysis and machine intelligence*, vol. 28, no. 4, pp. 657-662, 2006.

[18] V. Kellokumpu, G. Zhao, and M. Pietikäinen, "Human activity recognition using a dynamic texture based method," in *BMVC*, 2008, vol. 1, p. 2.

[19] S. Liao, M. W. Law, and A. C. Chung, "Dominant local binary patterns for texture classification," *IEEE transactions on image processing*, vol. 18, no. 5, pp. 1107-1118, 2009.

[20] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 28, no. 12, pp. 2037-2041, 2006.

[21] H. Jin, Q. Liu, H. Lu, and X. Tong, "Face detection using improved LBP under Bayesian framework," in *Image and Graphics (ICIG'04), Third International Conference*, 2004: IEEE, pp. 306-309.

[22] W. Louis and K. N. Plataniotis, "Co-occurrence of local binary patterns features for frontal face detection in surveillance applications," *EURASIP Journal on Image and Video Processing,* vol. 2011, no. 1, p. 745487, 2010.

[23] G. Zhao and M. Pietikainen, "Dynamic texture recognition using local binary patterns with an application to facial expressions," *IEEE transactions on pattern analysis and machine intelligence,* vol. 29, no. 6, 2007.

[24] C. Shan, S. Gong, and P. W. McOwan, "Facial expression recognition based on local binary patterns: A comprehensive study," *Image and Vision Computing,* vol. 27, no. 6, pp. 803-816, 2009.

[25] L. Nanni and A. Lumini, "Local binary patterns for a hybrid fingerprint matcher," *Pattern recognition,* vol. 41, no. 11, pp. 3461-3466, 2008.

[26] V. Takala, T. Ahonen, and M. Pietikäinen, "Block-based methods for image retrieval using local binary patterns," in *Scandinavian Conference on Image Analysis*, 2005: Springer, pp. 882-891.

[27] M. Subrahmanyam, R. Maheshwari, and R. Balasubramanian, "Local maximum edge binary patterns: a new descriptor for image retrieval and object tracking," *Signal Processing,* vol. 92, no. 6, pp. 1467-1479, 2012.

[28] L. Liu, L. Zhao, Y. Long, G. Kuang, and P. Fieguth, "Extended local binary patterns for texture classification," *Image and Vision Computing,* vol. 30, no. 2, pp. 86-99, 2012.

[29] H. Zhou, R. Wang, and C. Wang, "A novel extended local-binary-pattern operator for texture analysis," *Information Sciences,* vol. 178, no. 22, pp. 4314-4325, 2008.

[30] Y. Guo, G. Zhao, and M. PietikäInen, "Discriminative features for texture description," *Pattern Recognition,* vol. 45, no. 10, pp. 3834-3843, 2012.

[31] X. Qi, R. Xiao, C.-G. Li, Y. Qiao, J. Guo, and X. Tang, "Pairwise rotation invariant co-occurrence local binary pattern," *IEEE transactions on pattern analysis and machine intelligence,* vol. 36, no. 11, pp. 2199-2213, 2014.

[32] D. K. Iakovidis, E. G. Keramidas, and D. Maroulis, "Fuzzy local binary patterns for ultrasound texture characterization," in *International conference image analysis and recognition*, 2008: Springer, pp. 750-759.

[33] J. Chen, V. Kellokumpu, G. Zhao, and M. Pietikäinen, "RLBP: Robust Local Binary Pattern," in *BMVC*, 2013.

[34] A. Fathi and A. R. Naghsh-Nilchi, "Noise tolerant local binary pattern operator for efficient texture analysis," *Pattern Recognition Letters*, vol. 33, no. 9, pp. 1093-1100, 2012.

[35] J. Ren, X. Jiang, and J. Yuan, "Noise-resistant local binary pattern with an embedded error-correction mechanism," *IEEE Transactions on Image Processing*, vol. 22, no. 10, pp. 4049-4060, 2013.

[36] L. Liu, S. Lao, P. W. Fieguth, Y. Guo, X. Wang, and M. Pietikäinen, "Median robust extended local binary pattern for texture classification," *IEEE Transactions on Image Processing*, vol. 25, no. 3, pp. 1368-1381, 2016.

[37] F. Ahmed, A. H. Bari, A. Shihavuddin, H. A. Al-Mamun, and P. Kwan, "A study on local binary pattern for automated weed classification using template matching and support vector machine," in *2011 IEEE 12th International Symposium on Computational Intelligence and Informatics (CINTI)*, 2011: IEEE, pp. 329-334.

[38] F. Ahmed, M. H. Kabir, S. Bhuyan, H. Bari, and E. Hossain, "Automated weed classification with local pattern-based texture descriptors," *International Arab Journal of Information Technology*, vol. 11, no. 1, pp. 87-94, 2014.

[39] P. Lottes, M. Hörferlin, S. Sander, and C. Stachniss, "Effective Vision-based Classification for Separating Sugar Beets and Weeds for Precision Farming," *Journal of Field Robotics*, vol. 34, no. 6, pp. 1160-1178, 2017.

[40] C. McCool, I. Sa, F. Dayoub, C. Lehnert, T. Perez, and B. Upcroft, "Visual detection of occluded crop: For automated harvesting," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, 2016: IEEE, pp. 2506-2512.

[41] S. R. Dubey and A. S. Jalal, "Detection and classification of apple fruit diseases using complete local binary patterns," in *2012 Third International Conference on Computer and Communication Technology*, 2012: IEEE, pp. 346-351.

[42] Y. Herdiyeni and M. M. Santoni, "Combination of morphological, local binary pattern variance and color moments features for indonesian medicinal plants identification," in *2012 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, 2012: IEEE, pp. 255-259.

[43] V. N. T. Le, B. Apopei, and K. Alameh, "Effective plant discrimination based on the combination of local binary pattern operators and multiclass support vector machine methods," *Information Processing in Agriculture*, 2018.

[44] M. A. Savelonas, D. K. Iakovidis, and D. Maroulis, "LBP-guided active contours," *Pattern Recognition Letters*, vol. 29, no. 9, pp. 1404-1415, 2008.

[45] A. Satpathy, X. Jiang, and H.-L. Eng, "LBP-based edge-texture features for object recognition," *IEEE Transactions on Image Processing*, vol. 23, no. 5, pp. 1953-1964, 2014.

[46] D. T. Nguyen, P. O. Ogunbona, and W. Li, "A novel shape-based non-redundant local binary pattern descriptor for object detection," *Pattern recognition*, vol. 46, no. 5, pp. 1485-1500, 2013.

[47] J. Lin and C.-T. Chiu, "Low-complexity face recognition using contour-based binary descriptor," *IET Image Processing*, vol. 11, no. 12, pp. 1179-1187, 2017.

[48] B. Scholkopf and A. J. Smola, *Learning with kernels: support vector machines, regularization, optimization, and beyond*. MIT press, 2001.

[49] Y. Ma and G. Guo, *Support vector machines applications*. Springer, 2014.

[50] M. R. Bai, V. V. Krishna, and J. SreeDevi, "A new morphological approach for noise removal cum edge detection," *International Journal of Computer Science Issues*, vol. 7, no. 6, p. 187, 2010.

[51] J. Serra, *Image analysis and mathematical morphology*. Academic Press, Inc., 1983.

[52] R. M. Haralick, S. R. Sternberg, and X. Zhuang, "Image analysis using mathematical morphology," *IEEE Transactions on pattern analysis and machine intelligence*, no. 4, pp. 532-550, 1987.

[53] Z. Yu-Qian, G. Wei-Hua, C. Zhen-Cheng, T. Jing-Tian, and L. Ling-Yun, "Medical images edge detection based on mathematical morphology," in *2005 IEEE engineering in medicine and biology 27th annual conference*, 2006: IEEE, pp. 6492-6495.

[54] L. Vincent, "Morphological area openings and closings for grey-scale images," in *Shape in Picture*: Springer, 1994, pp. 197-208.

[55] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern recognition,* vol. 29, no. 1, pp. 51-59, 1996.

[56] M. Pietikäinen, T. Ojala, and Z. Xu, "Rotation-invariant texture classification using feature distributions," *Pattern Recognition,* vol. 33, no. 1, pp. 43-52, 2000.

[57] G. Guo, S. Z. Li, and K. L. Chan, "Support vector machines for face recognition," *Image and Vision Computing,* vol. 19, no. 9-10, pp. 631-638, 2001.

[58] P. J. Phillips, "Support vector machines applied to face recognition," in *Advances in Neural Information Processing Systems,* 1999, pp. 803-809.

[59] A. Tellaeche, G. Pajares, X. P. Burgos-Artizzu, and A. Ribeiro, "A computer vision approach for weeds identification through Support Vector Machines," *Applied Soft Computing,* vol. 11, no. 1, pp. 908-915, 2011.

[60] J. M. Guerrero, G. Pajares, M. Montalvo, J. Romeo, and M. Guijarro, "Support vector machines for crop/weeds identification in maize fields," *Expert Systems with Applications,* vol. 39, no. 12, pp. 11149-11155, 2012.

[61] T. Rumpf, A.-K. Mahlein, U. Steiner, E.-C. Oerke, H.-W. Dehne, and L. Plümer, "Early detection and classification of plant diseases with support vector machines based on hyperspectral reflectance," *Computers and electronics in agriculture,* vol. 74, no. 1, pp. 91-99, 2010.

[62] J. K. Patil, R. Kumar, and Research, "Advances in image processing for detection of plant diseases," *Journal of Advanced Bioinformatics Applications,* vol. 2, no. 2, pp. 135-141, 2011.

[63] L. Wu and Y. Wen, "Weed/corn seedling recognition by support vector machine using texture features," *African Journal of Agricultural Research,* vol. 4, no. 9, pp. 840-846, 2009.

[64] F. Ahmed, H. A. Al-Mamun, A. H. Bari, E. Hossain, and P. Kwan, "Classification of crops and weeds from digital images: A support vector machine approach," *Crop Protection,* vol. 40, pp. 98-104, 2012.

[65] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proceedings of the fifth annual workshop on Computational learning theory,* 1992: ACM, pp. 144-152.

[66] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning,* vol. 20, no. 3, pp. 273-297, 1995.

[67] C.-C. Chang and C.-J. Lin, "LIBSVM: a library for support vector machines," *ACM transactions on intelligent systems,* vol. 2, no. 3, p. 27, 2011.

[68] S. R. Gunn, "Support vector machines for classification and regression," *ISIS technical report,* vol. 14, no. 1, pp. 5-16, 1998.

[69] A. E. Radford, "Vascular plant systematics," 1974.

[70] C.-W. Hsu, C.-C. Chang, and C.-J. Lin, "A practical guide to support vector classification," ed: Taipei, 2003.

[71] A. A. Kalinin *et al.*, "3D shape modeling for cell nuclear morphological analysis and classification," *Scientific reports,* vol. 8, no. 1, pp. 1-14, 2018.

[72] S. Chen, M. Zhao, G. Wu, C. Yao, and J. Zhang, "Recent advances in morphological cell image analysis," *Computational and mathematical methods in medicine,* vol. 2012, 2012.

Chapter 4 – Performances of the LBP based algorithm over CNN models for detecting crops and weeds with similar morphologies

This chapter has been published in the Sensors Journal,vol. 20, no. 8, p. 2193, 2020 (DOI: https://doi.org/10.3390/s20082193). The paper has been changed to the layout, number formats, font size and font style, which was implemented to maintain consistency in the formatting of this thesis.

4.1 Abstract

Weed invasions pose a threat to agricultural productivity. Weed recognition and detection play an important role in controlling weeds. The challenging problem of weed detection is to discriminate between crops and weeds with similar morphology under natural field conditions such as occlusion, varying lighting conditions, and different growth stages. In this paper, we evaluate a novel algorithm, k-FLBPCM (filtered Local Binary Patterns with contour masks and coefficient k), for discriminating between morphologically similar crops and weeds that shows significant advantages, in both model size and accuracy, over state-of-the-art deep convolutional neural network (CNN) models such as VGG-16, VGG-19, ResNet-50 and InceptionV3. The experimental results on the "bccr-segset" dataset in the laboratory testbed setting show that the accuracy of CNN models with fine-tuned hyper-parameters is slightly higher than the k-FLBPCM method, while the accuracy of the k-FLBPCM algorithm is higher than the CNN models (except for VGG-16) for the more realistic "fieldtrip_can_weeds" dataset collected from real-world agricultural fields. However, the CNN models require a large amount of labelled samples for the training process. We conducted another experiment based on training with crop images at mature stages and testing at early stages. The k-FLBPCM method outperformed the state-of-the-art CNN models in recognizing small leaf shapes at early growth stages, with error rates an order of magnitude lower than CNN models for canola-radish (crop-weed) discrimination using a subset extracted from the "bccr-segset" dataset, and for the "mixed-plants" dataset. Moreover, the real-time weed-plant discrimination time attained with the k-FLBPCM algorithm is approximately 0.223ms per image for the laboratory dataset and 0.346ms per image for the field dataset, and this is an order of magnitude faster than that of CNN models.

Keywords: Local Binary Pattern (LBP); k-FLBPCM; Deep Convolutional Neural Networks; Precision Agriculture; Crop/Weed classification and detection.

4.2 Introduction

Precision agriculture plays an indispensable role in increasing the productivity of agriculture, food security and sustainability, and reducing the detrimental impacts on the environment. Amongst the major threats to agricultural production are weed infestation, plant diseases and herbicide resistance.

Identifying weeds at early crop growth stages brings many benefits for weed management prior to crop damage. This results in the reduction of herbicide usage, minimizes the negative impacts on the environment, improves grower profitability and maintains high crop quality [1]. Variable herbicide application systems, based on weed identification algorithms, have shown great promise in experimental results. A study on the effectiveness of the sensor-based precision herbicide application is described in [2]. The average herbicide savings in 13 field trials was 24.6%, using sensors for detecting weeds [2]. In another four-year study, average herbicide savings for controlling grass weeds were 78% in maize and 36% in sugar beet crops [3]. Furthermore, the amount of herbicide used for controlling broad-leaved weeds were saved up to 11% in maize and 41% in sugar beet crops [3].

With the technological advances in precision agriculture, a substantial number of studies have been developed to discriminate crops from weeds [4-7]. One of the most popular and effective methods is plant image analysis [4, 8]. There have been many techniques used for analysing images in the stages of pre-processing, segmentation, feature extraction and classification. Each stage plays an indispensable role in weed detection. However, the performance of computer vision algorithms is greatly dependent on selecting an appropriate set of features [9]. Particularly, the key characteristics of vegetation (crops and weeds), which comprise biological morphology [10-12], spectral features [13-15], spatial contexts [16-18] and visual textures [19-21] can be extracted by applying different characterization methods. Each of these characteristics has its own advantages, and depending on the complexity of the generated datasets for plant species. Machine learning techniques, such as Support Vector Machines (SVM), K-means and Artificial Neural Networks (ANNs) can be applied to classify these species [5, 22].

One of the most competitive and widespread broadleaf weeds in Australia is Wild radish (Raphanus raphanistrum) [23]. Wild radish has a devastating impact on canola crops and farmers have been struggling to effectively eradicate it and minimize its threats to canola crop fields [24]. When the leaf shape of crops and weeds have different morphology, for example broad leaves and narrow leaves, they can be easily distinguished. However, canola and wild radish broadleaf plants have very similar colour and shape. Datasets collected at different growth stages, rotations, and illuminations for canola, corn, wild radish and soil background have been generated to investigate effective plant discrimination based on the combination of local binary pattern operators (LBP) and multiclass support vector machine methods. However, due to their similar leaf shapes, the classification accuracy was considerably reduced [25]. While LBP is one of the most robust and effective methods for plant classification based on morphology [26-31], to overcome the classification limitation when plant species have similar shape and colour, additional features must be combined with LBP features. In this paper, we demonstrate the performance of a novel plant classification technique, entitled k-FLBPCM, which is based on the use of plant contour features and filtered LBP features with a coefficient k to improve the accuracy rate of broadleaf plants of close colours and shapes [32]. We also compare our method with other methods that have recently been reported.

Despite many efforts to extract leaf features and classify plants using complex computer vision algorithms [33-37], plant recognition is still regarded as a challenging problem [38]. For the machine-vision-based classification of plant leaves in real field conditions, many challenging problems arise, including lighting conditions, overlapping, occlusion, and damaged leaves. Recently, studies on deep learning (DL) have produced extremely promising classification results for various applications, such as image recognition, natural language processing and speech recognition [39, 40]. Within the realm of precision agriculture, a variety of agricultural challenges have been solved by using DL [4]. It is also important to note that DL tools represent a subfield of machine learning, enabling artificial neural networks to automatically extract abstract and robust features that are invariant to illumination and distortions from raw data [4]. Particularly, DL extracts the high level features from the hierarchical layers of data representation by composing lower level features [39]. With the availability of high computing capacity and data, DL techniques combining feature extraction and classification stages can potentially reduce manual and expensive engineering processing, thus making accurate real-time plant classification viable and cost-effective [4].

Convolutional Neural Networks (CNNs) are particularly based on deep learning models, and have been widely used for image-based classification of plants. CNNs have exhibited high classification accuracies because of the use of spatial information and correlation filters between layers [41-47]. CNNs typically comprise several layers, namely, convolutional layers, pooling layers and fully connected layers, in addition to activation functions. The convolutional layers are regarded as feature extractors. The role of the pooling layers is to reduce the dimensionality of images, while the fully connected layers are used for classification [48]. CNN architectures have been finely tuned and developed in recent years to allow the reuse of transfer learning. Amongst the popular and successful CNN architectures are AlexNet [49], Visual Geometry Group (VGG) [50], GoogleNet [51], Inception [52] and ResNet [53]. Based on the evolution of the CNN architectures, it is generally observed that the more accurate CNNs tend to have deeper learning. In this paper, we choose VGG-16, VGG-19, ResNet-50 and Inception-V3 architectures that have demonstrated strong performances on various datasets and state-of-the-art results in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) [50, 54], and compare their performances with the k-FLBPCM method.

There have been several comparative studies of CNNs and LBP for image classification [55-57], with datasets captured by various devices in different conditions. While the CNNs and LBP performances have been extensively investigated for proof-of-concept classification demonstration, the computation time for both deep learning and machine learning methods was mentioned limitedly. Despite the attractive classification capabilities of CNNs, some limitations still exist, such as the need for huge datasets for the training process, overfitting problems and time execution [58]. In this paper, four well-

known DL architectures comprising VGG-16, VGG-19, ResNet-50 and Inception-V3 are used to extract relevant features for the identification of crop and weed with similar morphological characteristics. The performances of the DL architectures and the proposed machine learning method (k-FLBPCM + SVM) are compared for the detection of crops and weeds of similar morphologies using in the "bccr-segset" dataset, collected in a laboratory setting (published online) and in the "fieldtrip_can_weeds" field dataset (published online now for this paper). The laboratory dataset, which contains 30,000 plant images, was captured at four different growth stages and has four classes including soil background, canola, corn and wild radish. The field dataset, on the other hand, comprises of 4,914 field images and was gathered under complex field environments and illumination variations (morning and afternoon light). Further, we measure the time typically spent in training and testing of deep neural networks and compare it with that for the k-FLBPCM feature extractor with an SVM classifier.

4.3 Materials

K-FLBPCM method

The LBP method, which was introduced by Ojala et al. in 1996 [59], has long been the most effective and robust texture descriptor in many areas [60-62]. The use of the LBP algorithm has many advantages, such as computational, rotation and illumination invariance. LBP was developed to extract dominant features with the aim of enhancing the effectiveness of classification accuracy, and may be combined with other feature extraction methods to improve classification accuracy in various applications [62-64].

Specifically, for weed and crop classification using machine vision, the recognition of leaves is based on morphological features, such as texture and shape. Due to the similarity in color for canola and wild radish species, color features cannot be considered in the context of identifying green plants. According to the "bccr-segset" and "fieldtrip_can_weeds" dataset, broadleaf canola and wild radish plants pose challenges for identifying their similar morphology at every growth stage. Therefore, we developed a novel LBP based algorithm to solve this problem. To be more specific, texture features were extracted by the combination of LBP operators and morphological features were extracted by applying contour masks on plant images. This method is based on combining contour mask features and filtered LBP features with a coefficient k and is called k-FLBPCM [32]. Due to the independence of morphological features with rotation, different growth stages and geometric translation, the combination of these features enhances the crop/weed classification and discrimination accuracy.

The detailed flowchart of the new LBP method is presented in Figure 1. All plant images in datasets were divided into two branches. For the left branch, all images went through the feature extraction stage without applying morphological operators. At the feature extraction process, different LBP operators were combined. In each bin of the generated LBP histogram, the dominant bin value was removed in

order to allow a better distribution of features. Hence, the bins with the highest value were removed. From the left branch, LBP features (pass_features) were extracted. For the right branch, opening and closing morphological operators were applied to all images using a 5×5 structure element. Before these images were processed by the feature extraction stage, contouring masks were generated from morphological image processing filters with different thicknesses. Then, these masks were processed by using the combination of LBP operators and removing the bins with the highest values as done in the left branch. From the right branch, LBP features with contour masks (cmask_features) were extracted. The combined features were calculated by multiplying pass_features with a coefficient k and summing with cmask_features. At the classification stage, 5-fold cross validation method was applied to prevent overfitting. Then, the SVM classifier with an RBF kernel was used. To achieve higher classification accuracies, hyper-parameters (C and Gamma) were set before the training model and appropriately tuned to attain the maximum accuracy. While C is used to control error, Gamma is used to give curvature weight of the decision boundary. The source code of this method was presented at the link (https://github.com/vinguyenle/k-FLBPCM-method).

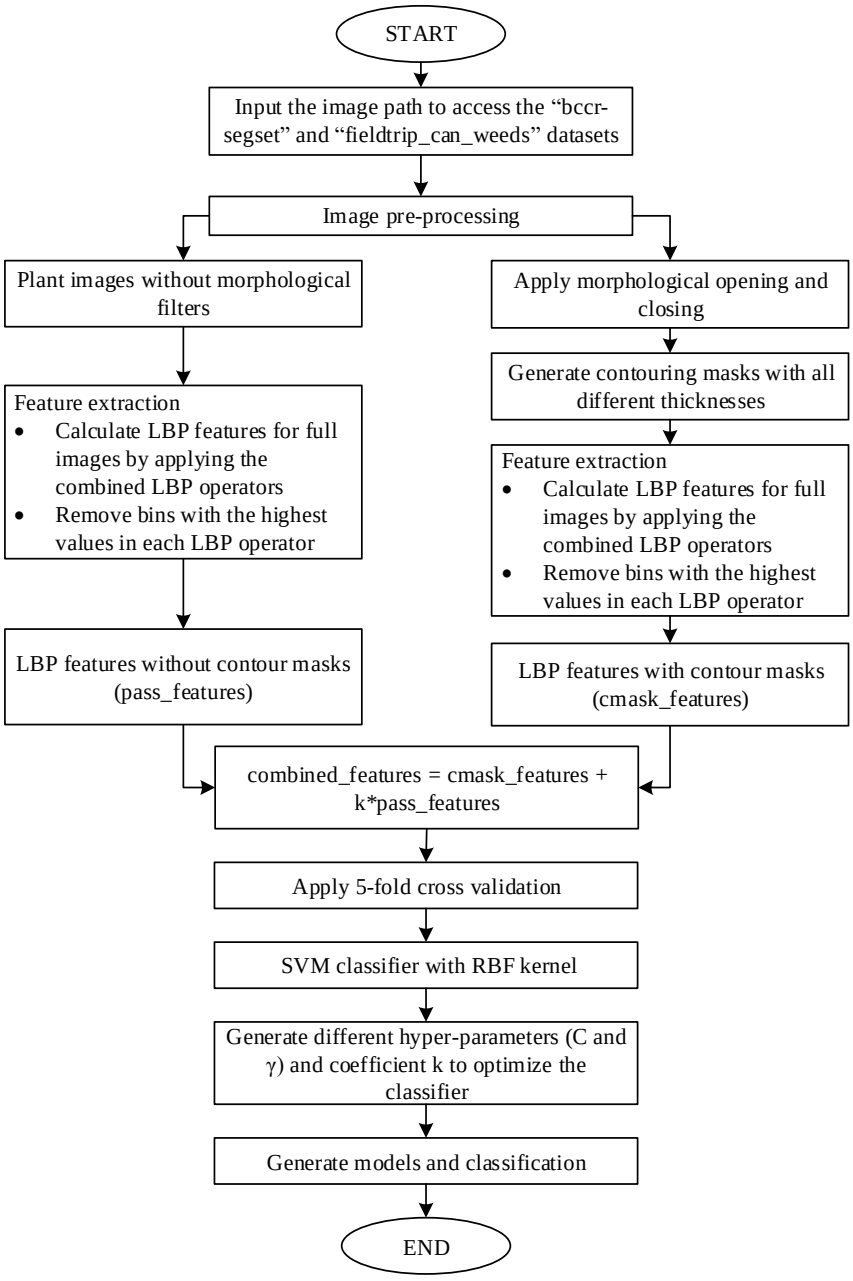


Figure 1. The flowchart presents how the k-FLBPCM algorithm works [32].

4.4 Methodology

4.4.1 Data collection in the laboratory

Plant images were adopted from the "bccr-segset" dataset (published online) [25]. All data was captured on a custom-built testing facility at ESRI (Electron Science Research Institute), Edith Cowan University, Australia. The size of all images was 228×228 pixels. As can be seen in Figure 2, the dataset comprises 30,000 plant images partitioned into four classes (canola, corn, wild radish and background) under different rotations, scales and illumination conditions. Images were collected by applying

different rotation angles (45°, 90°, 135°, 180°, 225°, 270°, 315°, 360°), lighting conditions (sunlight and fluorescent light), sizes and morphologies of plants through four growth stages [25].

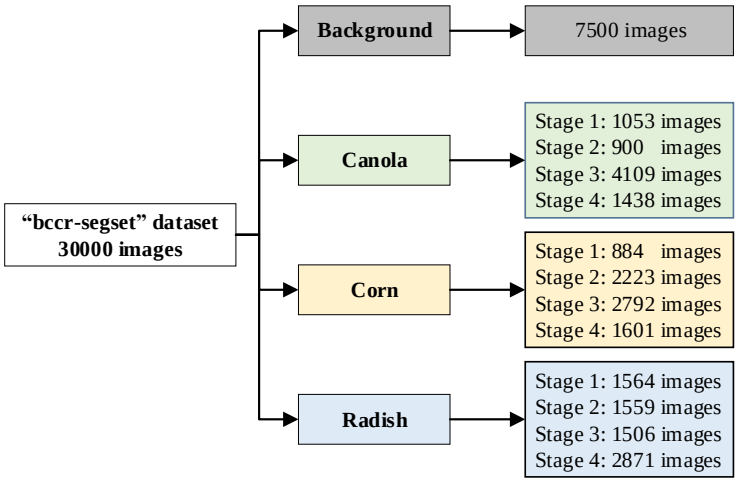


Figure 2. The "bccr-segset" dataset and its four-growth stages.

The 5-fold cross validation was used in the "bccr-segset" dataset. This dataset was randomly shuffled and divided into five equal subsets with 6,000 plant images in each subset. Then, we used 4 folds (24,000 plant images) for training and a fold (6,000 images) for testing. Each testing set was generated in each iteration, until each fold in 5 folds has been used as the testing set. In addition, as for deep neural networks, 24,000 images were divided into two datasets including 21,000 images for training 3,000 images for validation.

Based on the "bccr-segset" dataset collected at different growth stages, we tested the performance of the k-FLBPCM method and CNNs when training and testing sets had the same growth stage and different growth stages. As can be seen in Figure 3, the size of both canola and radish plants at the fourth stage is larger than the image frame, while the full size of canola and radish plants at the second and third stage can be observed. We used the training set at the fourth stage and testing set at the second stage. However, since the performance of both the k-FLBPCM method and CNNs was unsatisfactory, we selected the canola and radish images at the second and third growth stages extracted from the "bccr-segset" dataset (with the name "can_rad_stage2_stage3" online) as follows:

- 1600 images (800 canola images and 800 radish images) at stage 3 for training
- 400 images (200 canola image and 200 radish images) at stage 3 for testing
- 400 images (200 canola image and 200 radish images) at stage 2 for testing

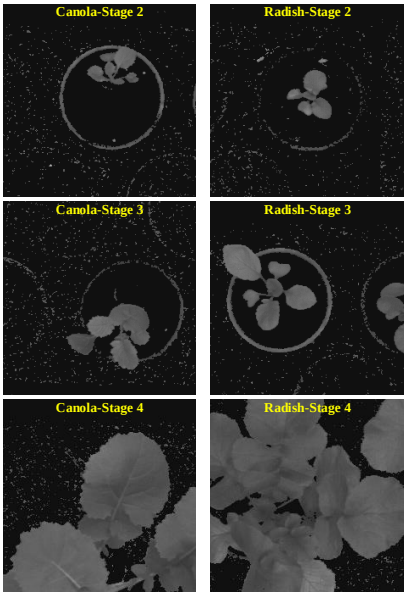


Figure 3. Canola and radish plants at different growth stages in the "bccr-segset" dataset.

We also collected another dataset, called "mixed-plants" dataset (online) to validate the performance of the k-FLBPCM method and CNN models. For this dataset, barley, canola and wild radish were mixed together and then grown in plant pots. There were two groups including a 50:50% barley: canola mixture, and a 50:50% barley: wild radish mixture. These groups were captured at different growth stages as can be seen in Figure 4. The corresponding dataset comprised:

- 3,000 images (1,500 mixed barley-canola images and 1,500 mixed barley-radish images) at stage 4 for training,
- 750 images (375 mixed barley-canola images and 375 mixed barley-radish images) at stage 4 for testing,
- 750 images at both stage 2 and stage 3 (375 mixed barley-canola images at stage 2 and 375 mixed barley-radish images at stage 3) for testing.

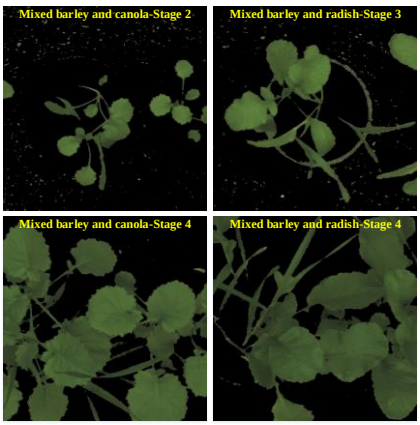


Figure 4. Barley was mixed with canola and wild radish at different growth stages.

4.4.2 Field data collection

Field images were captured by an integrated weed sensing system with the combination of multispectral and spatial sensors at a commercial farm in Cunderdin, Western Australia, shown in Figure 5. This hardware system, which is housed at the Electron Science Research Institute (ESRI), Edith Cowan University, Australia, consists of two components (i) a Xilinx Zynq ZC702 development board with a VITA 2000 camera sensor and (ii) a Plan Discrimination Unit (PDU) [15] based on spectral reflectance measurements.

We collected a "fieldtrip_can_weeds" dataset (published online) under different weather conditions (cloudy, windy, and sunny) and illumination variations (morning and afternoon light). There are 4,914 field images with 3 classes, including background (1,638 images), canola (1,638 images), wild radish (1638 images). When all field images were segmented by using Excess Green minus Excess Red Indices (ExG-ExR) method, the segmented plants were presented in Figure 6. It is worth noting that mixing wild radish and barley in the wild radish class under practical field conditions is to challenge our algorithm and DCNN models.

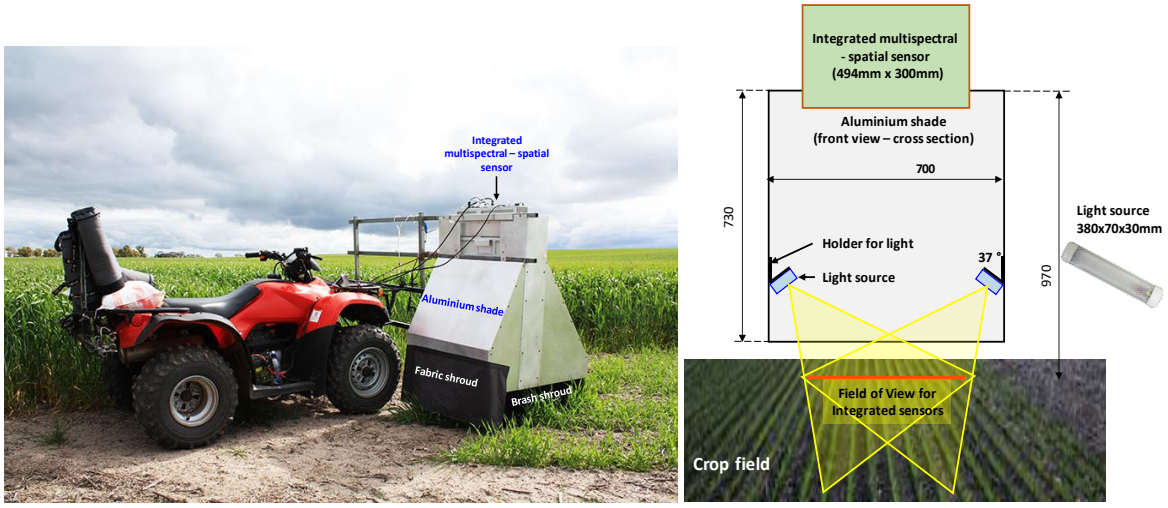


Figure 5. An integrated weed sensing system to collect plant images in the field.

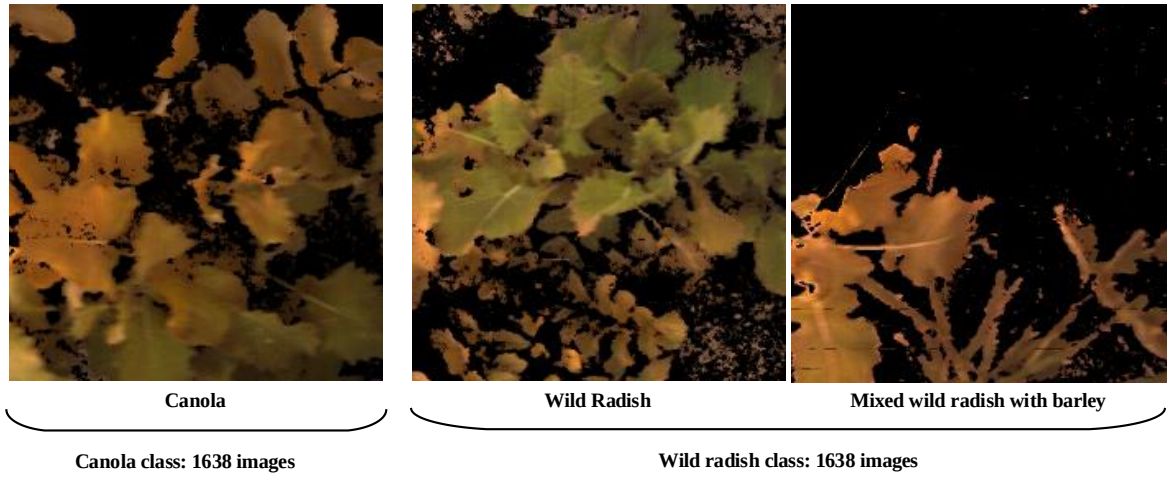


Figure 6. Segmented canola and wild radish classes under complex field environments.

With the aim of comparing the various weed detection methods, we set up experiments with similar conditions. Therefore, the dataset was divided into 5 roughly equal parts. In particular, with the 5-fold cross validation method, a model was trained 5 times, each time a different single part was used as a testing set with 982 field images, while the remaining 4 parts with 3,928 field images were used for training. Then, the cross-validation process was repeated 5 times, with each testing set used only once. As for deep neural networks, it is important to note that 3,928 field images were divided into two datasets including 3,437 images for training and 491 images for validation.

4.4.3 Training k-FLBPCM and CNNs models

Before training, to meet the input dimension requirement of deep neural networks, all plant images were resized. In this paper, the input shape of the VGG-16, VGG-19 or Resnet-50 networks was 224 × 224 × 3 pixels, while the input shape of the Inception-V3 network was 299 × 299 × 3 pixels. It is important to note that all generated models, training and testing processes were implemented using the deep learning framework, Keras (with TensorFlow 2.0 backend). The Ubuntu 18.04 LTS operating system and Python 3.7 were used in this paper. The training and testing were performed on a workstation with an Intel Core i7-7820X CPU, a GeForce GTX1080Ti Graphic Processing Unit (GPU) with 11GB of memory. Additionally, the k-FLBPCM method was also implemented on the same machine to compare with CNN models.

When deep neural networks were trained on natural images, the features learned in the first layers tended to be general and then had transitions to be more specific by the last layers of the network. Thus, transfer learning was regarded as an efficient technique to transfer features learned in one or more datasets and reuse these features to improve learning in other datasets [65]. In other words, the transfer learning method was responsible for keeping the parameters of the previous layers, then removing the last layer of CNN models, and then retraining the last layer. In this paper, we chose the VGG-16, VGG-19,

ResNet-50 and Inception V3 models, which were fine-tuned by using neural networks pre-trained on the ILSVRC versions of ImageNet dataset. Such CNNs are suitable for transfer learning in network-based deep transfer learning [65]. The public ImageNet dataset consists of 1.28 million natural images and 1000 classes corresponding to 1,000 categories. Therefore, the last layer in this dataset has 1,000 output nodes. To apply for the "bccr-segset" dataset collected from the laboratory, we changed the output to 4 output nodes corresponding to 4 categories (background, canola, corn and wild radish). For the "fieldtrip_can_weeds" dataset collected from the field, the output was changed to 3 output nodes corresponding to 3 categories (background, canola, and wild radish).

As for the aforementioned CNN models, each model was loaded with the corresponding weights pre-trained on the ImageNet dataset and resized plant images to the standard image size, before the training process, as shown in Table 1. Then, we used stochastic gradient descent (SGD) to optimize parameters over a training set using mini-batches of 32 and 64 images, and selected dropout rates of 50% and 20% in the training stage for regularization. After some preliminary training experiments, the learning rate was adjusted to 0.001 and the number of epochs was set to 10 for the laboratory and field datasets. VGG-16 and VGG-19 models were kept as the original models and changed to 4 outputs in the last layer, while ResNet-50 and Inception-V3 models were fine-tuned by adding a max pooling layer with a pool size 5×5, a flatten layer, a fully connected layer with a dense 1024 and ReLU activation and the last layer with 4 outputs and softmax activation. It is observed that these added layers show good performances with our dataset.

Table 1. Input image sizes used for the CNNs and k-FLBPCM models, for the laboratory and field datasets.

Methods	Image size
k-FLBPCM	228×228
VGG-16	224×224
VGG-19	224×224
ResNet-50	224×224
Inception-V3	299×299

4.5 Results and Discussions

We conducted three comparison experiments to investigate the performances of the k-FLBPCM method and CNNs on the laboratory and field datasets.

4.5.1 Comparison of the classification accuracies of the k-LBPCM method and DCNNs in the "bccr-segset" dataset

The "bccr-segset" dataset (comprising 30,000 plant images) was equally divided into 4 classes (background, canola, corn and wild radish). We applied 5-fold cross validation on each class to prevent

overfitting. Each class was shuffled randomly, and one of the five folds was taken as the test set, while the remaining folds were considered as the training set. The random splits for each fold were performed using random seeds.

For deep neural networks (VGG-16, VGG-19, ResNet-50 and Inception-V3), 3,000 images in the training set were used for validation. Before training, each model was loaded with the corresponding weights that were pre-trained on ImageNet. Then, we used the transfer learning technique to fine tune models as described in Section 4.3 and 4.4 (Materials and Method). The standard *sparse_categorial_crossentropy* loss function was used for training. After the trial with using optimizers, the SGD optimizer was selected due to its superior performance. The momentum was 0.9 and the learning rate was 0.001. The optimal batch size and dropout of the training set were selected experimentally. The observed average classification accuracies varied across the different models.

The accuracy obtained for the k-FLBPCM method was 98.60% with *C=30, γ=1e-5, thickness=2* and coefficient *k =0.2*. The classification accuracies of deep neural networks were slightly higher than the k-FLBPCM method, as shown in Table 2 and Table 4. Furthermore, VGG-16, VGG-19 and Inception-V3 models attained higher average accuracies than that of the ResNet-50 model. The performance of the VGG-16 model was the highest achieving 99.87% with a batch size of 32 and dropout 0.2 in the "bccr-segset" dataset. As can be seen in Table 3, confusion matrices of the test set for individual classes in the "bccr-segset" dataset were presented to compare the performance of selected methods in distinguishing cultivated plants from weeds with a similar appearance.

Table 2. Classification accuracies of the test set, in the "bccr-segset" dataset, for different methods, for a batch size of 32 and dropout 0.5.

Methods	**Accuracy of the test set**					
	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Average accuracy
k-FLBPCM	98.67%	98.75%	98.60%	98.56%	98.41%	98.60%
VGG-16	99.83%	99.73%	99.75%	99.90%	99.85%	99.81%
VGG-19	99.82%	99.22%	99.82%	99.52%	99.85%	99.65%
ResNet-50	99.48%	99.58%	99.62%	99.72%	99.67%	99.61%
Inception-V3	99.83%	99.75%	99.55%	99.85%	99.92%	99.78%

Table 3. Confusion matrices of the test set, in the "bccr-segset" dataset, for different methods, for a batch size of 32 and dropout 0.5.

Methods	Classes	Background	Canola	Corn	Radish
k-FLBPCM	**Background**	1497	1	0	0
	Canola	0	1457	4	38
	Corn	0	3	1495	0
	Radish	0	37	0	1461
VGG-16	**Background**	1484	0	0	0
	Canola	0	1491	0	1
	Corn	0	0	1494	0
	Radish	0	5	0	1525
VGG-19	**Background**	1483	0	0	0
	Canola	2	1519	0	5
	Corn	2	0	1495	0
	Radish	0	0	0	1494
ResNet-50	**Background**	1483	1	0	0
	Canola	0	1490	0	2
	Corn	1	2	1491	0
	Radish	0	11	0	1519
Inception-V3	**Background**	1483	0	0	0
	Canola	0	1524	1	1
	Corn	2	0	1495	0
	Radish	0	1	0	1493

Table 4. Classification accuracies of the test set, in the "bccr-segset" dataset, for different methods, for a batch size of 32 and dropout 0.2.

Methods	Accuracy of the test set					
	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Average accuracy
k-FLBPCM	98.67%	98.75%	98.60%	98.56%	98.41%	98.60%
VGG-16	99.80%	99.85%	99.87%	99.93%	99.92%	99.87%
VGG-19	99.80%	99.83%	99.85%	99.85%	99.90%	99.85%
ResNet-50	99.82%	99.82%	99.22%	98.92%	99.25%	99.41%
Inception-V3	99.65%	99.72%	99.62%	99.65%	99.60%	99.65%

In order to explore the influence of batch size on the stability of the learning process, the next experiment kept all parameters and changed the batch size from 32 images to 64 images. The average classification accuracies of the VGG-16 and Inception-V3 models were higher than other selected neural networks and the k-FLBPCM method, as shown in Table 5. It is clear from Table 2, Table 4,

Table 5 that the changes in the average classification accuracy of the methods were insignificant when the batch size was increased from 32 to 64.

Table 5. Classification accuracies of the test set among different methods in the "bccr-segset" dataset with the batch size of 64 and dropout 0.2.

Methods	Accuracy of the testing set					
	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Average accuracy
k-LBPCM	98.67%	98.75%	98.60%	98.56%	98.41%	98.60%
VGG-16	99.82%	99.78%	99.90%	99.63%	99.85%	99.80%
VGG-19	99.73%	99.78%	99.83%	99.53%	98.82%	99.54%
ResNet-50	99.65%	99.52%	99.10%	99.70%	99.70%	99.53%
Inception-V3	99.82%	99.68%	99.83%	99.82%	99.85%	99.80%

4.5.2 Comparison of the classification accuracies of the k-LBPCM method and DCNNs in the training and test sets of different growth stages.

In the previous experiments, plant images with different growth stages were shuffled randomly in 5 folds. This means that the features were learned through the training process. As for the k-FLBPCM method, it learned the features of leaf shapes, especially the morphology of canola and radish plants. For the deep neural networks, we suspected that the learned features might not be extracted from the edges of the canola and radish broadleaves. So, an experiment was conducted to compare the performance of these methods with the training and testing sets coming from different growth stages.

Due to the superior performance of VGG-16 and Inception-V3 in the previous experiments, these models were selected to compare with the k-FLBPCM method in this experiment. The transfer learning technique was applied again to reduce the training time and effort required to recognize weeds and crops, and efficiently reuse the generated general features. The layers of VGG-16 model remain unchanged and the last layer was changed from 1,000 outputs (ImageNet) to 2 outputs (canola and radish plants in the "bccr-segset" dataset). Next, the Inception-V3 model was fine-tuned by adding some custom layers, including, a max pooling layer with a pool size 5×5, a flatten layer, a fully connected layer with a dense 1,024 nodes and ReLU activation and a last layer having 2 outputs and softmax activation. Dropout was set to 0.5 for both VGG-16 and Inception-V3 models.

As discussed in Section 4.4, the first experiment used 1,600 images (800 canola images and 800 radish images) collected at stage 3 for training and 400 images (200 canola image and 200 radish images)

collected at stage 3 for testing. In the second experiment we reused the 1,600 images (800 canola images and 800 radish images) collected at stage 3 for training and used 400 images (200 canola image and 200 radish images) collected at stage 2 for testing. We used the ratio 80:20 for the sizes of the training and test sets. The remaining 20% of plant images were reserved for testing and not used in the training process. A SGD optimizer with a learning rate of 0.001, a learning decay of 0.001 and a momentum of 0.9, was used in the "bccr-segset" dataset of canola and radish plants. The impact of the different batch sizes (32 and 64 images) on the accuracy of the networks during training in the "bccr-segset" dataset was not substantial, as discussed in Section 4.5.1. Hence, a batch size of 32 images was adequate, and used in this experiment with a number of epochs of 30 and a dropout of 0.5.

The experimental results shown in Table 6 illustrate the recognition performances of the k-FLBPCM method, VGG-16 and Inception-V3 using training and testing sets, for similar and different growth stages. As for the training and testing sets for similar growth stages (stage 3), the accuracies of VGG-16, Inception V3 and k-FLBPCM methods were relatively similar. It is worth noting that the parameters of the k-FLBPCM method including *C=100*, *γ =1e-7*, *thickness=2*, and *k=0.2* achieved an accuracy of 97.25%. Using stage 3 in the training set and stage 2 in the testing set, the accuracy of the k-FLBPCM method was 96.75% with parameters *C=100*, *γ =1e-6*, *thickness=2*, and *k=0.2*, while the optimal accuracies of VGG-16 and Inception-V3 models dropped to 62.5% (at epoch 18) and 63.8% (at epoch 16), respectively. When the training and testing sets were assigned to different growth stages, the capability of the k-FLBPCM method in recognizing canola and radish plants was significantly higher than those of the VGG-16 and Inception-V3 models. It can be explained that the k-FLBPCM method concentrates on extracting unique features of leaf shapes to train with, whereas VGG-16 and Inception-V3 architectures focus on filtering a wide range of features in plant images through many convolutional layers. Thus, the k-FLBPCM method can identify canola and wild radish plants much more generally than the widespread CNN methods.

Table 6. Comparison of the classification accuracies of methods in the use of canola and radish plants at different growth stages in the "bccr-segset" dataset.

Methods	**Canola and radish in the "bccr-segset" dataset**	
	Train-Stage3 and Test-Stage3	**Train-Stage3 and Test-Stage2**
	Test accuracy	**Test accuracy**
k-FLBPCM	97.25%	**96.75%**
VGG-16	98.96%	62.50%
Inception-V3	97.92%	63.80%

To confirm the ability of the k-FLBPCM method to recognize canola and radish plants with high accuracy using the "mixed-plants" dataset, close to that attained using the "bccr-segset" dataset. We conducted another experiment, where barley plants were mixed with canola and radish plants, thus making plant discrimination more challenging. The training and testing data division of the mixed-plants dataset is described in Section 4.4. 3,000 mixed-plant images, collected at the fourth stage, were used for training. Then 750 mixed-plant images, collected at the fourth stage, were assigned for the test set and another 750 mixed-plant images, collected at the second and third stages, were used for another test set. It is important to note that all images in the test sets were not used for training. However, the training set was combined with each test set, in order to compare the plant discrimination performance of the selected methods.

Table 7 shows the plant classification accuracies for the selected methods. The performance of the VGG-16, Inception-V3 and k-FLBPCM methods for the training and test sets collected at the fourth stage had approximately similar accuracies, namely 100% (at epoch 15), 99.05% (at epoch 30) and 99.73%, respectively. The optimum parameters of the k-FLBPCM method were $C=30$, $\gamma=1e\text{-}7$, $thickness=2$, and coefficient $k=0.2$. However, as observed from Table 7, for the training set using mixed-plant images collected at the fourth stage and the testing set using images collected at the second and third stages, the k-FLBPCM method again outperformed the CNN methods. The accuracies of VGG-16 and Inception-V3 models were 94.70% (at epoch 13) and 87.36% (at epoch 30), respectively, which are considerably lower than the 99.33% accuracy attained by the k-FLBPCM method.

After these two experiments, it was concluded that the k-FLBPCM method maintains high accuracy of recognizing single plants or mixed plants when the size of plant images in the training set are bigger than the ones in the test set even when plant images are collected at different growth stages, whereas, for the same conditions, the accuracies attained by deep neural networks drop to impractical levels. The effectiveness of the k-FLBPCM method is its ability to identify plant species using images collected at earlier growth stages, even if the available data is insufficient for training.

Table 7. Comparison of the classification accuracies of the VGG-16, Inception-V3 and k-FLBPCM methods when mixed-barley-canola and mixed-barley-radish images collected at different growth stages are used for the dataset.

Methods	**"Mixed-plants" dataset**	
	Train-Stage4 and Test-Stage4	**Train-Stage4 and Test-Stage2&Stage3**
	Test accuracy	**Test accuracy**
k-FLBPCM	99.73%	**99.33%**
VGG-16	100%	94.70%

Inception-V3	99.05%	87.36%

The k-FLBPCM algorithm provides better recognition accuracy with both the canola-radish subset, from the "bccr-segset" dataset, and the "mixed-plants" dataset. While the selected CNN models were applied to learn features of plants at the fourth growth stage and then identify plants at smaller growth stages (the second and third stages), their classification accuracy was lower than that of the k-FLBPCM algorithm. This is because the combination of extractors, including LBP features and contouring mask features, in the k-FLBPCM algorithm was able to accurately extract the edges of canola and radish leaves, and this is the key advantage of the k-FLBPCM method, especially with datasets comprising insufficient plant images.

4.5.3 Comparison of the classification accuracies of the k-LBPCM method and DCNNs in the dataset under complex field conditions

The experiments on the "fieldtrip_can_weeds" dataset were similar to those conducted on the "bccr-segset" dataset. The learning rate was 0.001, the dropout was 0.5, and the output of the CNN models was 3 output nodes corresponding to 3 classes (background, canola and wild radish). As can be seen from Table 8, the classification accuracy obtained for the k-FLBPCM method was 90.94% with *C=1000, γ=1e-8, thickness=2* and coefficient *k =0.5*. The accuracies of CNN models (except for VGG-16) were slightly lower than the k-FLBPCM method, indicating the efficacy of the novel algorithm. The experimental results demonstrate the ability of our algorithm to detect canola (crop) and mixed wild radish-barley (weed) with similar morphology under practical field conditions, compared to the CNN models. However, we expect that the CNNs may achieve higher accuracies when big data is input in the networks.

Table 8. Classification accuracies of the test set, in the "fieldtrip_can_weeds" dataset, for different methods, for a batch size of 32 and dropout 0.5.

Methods	**Accuracy of the testing set**					
	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Average accuracy
k-FLBPCM	92.33%	91.33%	90.18%	90.54%	90.34%	90.94%
VGG16	91.34%	91.55%	91.55%	91.75%	91.55%	91.55%
VGG19	90.12%	91.04%	89.41%	89.71%	87.47%	89.55%
Resnet50	88.59%	90.53%	90.94%	89.10%	89.51%	89.73%
Inceptionv3	91.75%	90.73%	91.04%	89.10%	91.75%	90.87%

4.5.4 Comparison of execution times

In precision agriculture, the processing time is an important aspect for real-time operation at practical farming speeds. In addition to the measured accuracies of the VGG-16, Inception-V3 and k-FLBPCM methods reported in the sections mentioned above, both the model training and testing times were measured.

4.5.4.1 Training time

The VGG-16, VGG-19, ResNet-50, Inception-V3 and k-FLBPCM models were implemented on the GPU GTX1080Ti in order to compare their processing times. Table 9 shows the measured total training time periods for all models. The number of epochs was set to 10 and the batch size was 32 images. The training time consumed in each epoch was accumulated from the five folds for all models in the "bccr-segset" and "fieldtrip_can_weeds" datasets. With both datasets, the total training time of the k-FLBPCM model was shorter than that of the VGG-16, VGG-19, ResNet-50 and Inception-V3 models. Note that the time taken to perform the required pre-processing steps was also included in the total training time periods shown in Table 9. Particularly, these steps consist of loading plant images, properly resizing them for input to deep neural networks, and applying morphological operators for the k-FLBPCM method.

Table 9. Total training time of the k-FLBPCM and the VGG-16, VGG-19, ResNet-50 and Inception-V3 models for datasets in the laboratory and in the field.

	Bccr-segset dataset	**Fieldtrip_can_weeds dataset**
Methods	Total training time (second)	Total training time (second)
k-LBPCM	**901.2**	**165.9**
VGG-16	8692	1394
VGG-19	10003	1563
ResNet-50	7657	1483
Inception-V3	11014	1907

4.5.4.2 Testing time

Table 10 shows the average testing time, which was computed by averaging the testing time periods for the five test folds, and the testing time per image, calculated by dividing the average testing time by the number of images in the "bccr-segset" test set (6,000 plant images) and the "fieldtrip_can_weeds" test set (982 field images). As shown in Table 10, the testing time of the k-FLBPCM method was 0.223ms per image in the "bccr-segset" laboratory dataset, which is more than one order of magnitude shorter than the testing times for the VGG-16, VGG-19, ResNet-50 and InceptionV3, which were 2.667ms, 3.033ms, 2.333ms, and 3.5ms, respectively. Similarly, the high efficiency of the k-FLBPCM algorithm

was also demonstrated in the "fieldtrip_can_weeds" field dataset, where only 0.346ms per image was necessary to run the field test set by applying our algorithm, compared to the testing time of CNN models.

Table 10. Testing time of the k-FLBPCM method and CNNs for the laboratory dataset (6,000 images used for the test set) and the field dataset (982 images used for the field test set)

Bccr-segset dataset (In the laboratory) – Test set		
Methods	**Average testing time (second/test set)**	**Testing time (millisecond/image)**
k-LBPCM	**1.34**	**0.223**
VGG-16	16	2.667
VGG-19	18.2	3.033
ResNet-50	14	2.333
Inception-V3	21	3.500
Fieldtrip_can_weeds dataset (In the field) – Test set		
k-LBPCM	**0.34**	**0.346**
VGG-16	3	3.055
VGG-19	3.2	3.259
ResNet-50	3	3.055
Inception-V3	4.6	4.684

Note that the Inception-V3 model requires a longer time in comparison with the other CNN networks, since its architecture is deeper. On the other hand, the k-FLBPCM algorithm has the ability to rapidly extract dominant features due to its computational efficiency. Although the selected deep neural networks eliminate the manual search for good feature extractors through the automatic learning of relevant features, deep neural networks go through many convolutional layers and are susceptible to high computational complexity.

It is important to note that the deep learning-based approaches typically require a large amount of data to outperform the k-FLBPCM method. This explains why the performance of selected neural networks was slightly better than the k-FLBPCM method in recognizing morphologically similar crops and weeds. When pre-trained CNN models are used to train plant images at four different growth stages in the "bccr-segset" dataset, they learn relevant features at each stage. The ability of CNN models is demonstrated by having high sample counts in the dataset and corresponding ground truth annotations. However, for real-time operation at high vehicular speeds, the long image processing time of these complex models makes them impractical if a high-accuracy performance cannot be compromised.

4.6 Conclusions

In this work, we have compared the performances of selected Convolutional Neural Network (CNN) models (VGG-16, VGG-19, ResNet-50 and Inception-V3 models) with the k-FLBPCM algorithm, specifically in identifying crop and weed species of similar morphologies. Experimental results, using the "bccr-segset" laboratory dataset, have shown that the both the CNN models with fine-tuned hyper-parameters and the k-FLBPCM method can achieve classification accuracies close to 99%. With the "fieldtrip_can_weeds" field dataset under complex field environments, the k-FLBPCM method can attain up to 90.94% classification accuracy, compared to 89.55%, 89.73% and 90.87% accuracies of VGG-19, ResNet-50 and Inception-V3, respectively (except for VGG-16 with 91.55%). However, for effective feature learning, these CNN models require a huge number of plant images to be collected at each of the various growth stages. On the other hand, we have demonstrated that the k-FLBPCM method can identify smaller leaf shapes using images collected at the second and third growth stages, with training using images of large leaves collected at the fourth growth stage. Results have shown that the k-FLBPCM method can achieve a canola-radish discrimination accuracy of 96.75% using the subset generated from the "bccr-segset" dataset, while the accuracies of the VGG-16 and Inception-V3 are 62.50% and 63.80%, respectively. Additional experimental results, using the "mixed-plants" dataset, have demonstrated the effectiveness of the k-FLBPCM method with 99.33% accuracy, whereas the accuracies of the VGG-16 and InceptionV3 are 94.70% and 87.36%, respectively. Furthermore, experimental results have shown that the k-FLBPCM model implemented on the GPU GTX1080Ti requires approximately 0.223ms per image in the "bccr-segset" laboratory dataset and 0.346ms per image in the "fieldtrip_can_weeds" field dataset for weed identification and detection. These results show the effectiveness of this algorithm for real-time precision agricultural applications

It is important to note that choosing an appropriate weed detection method depends on whether real-time operation is required and the detection accuracy can be compromised. The combination of extractors in the k-FLBPCM method can especially work well when the edges of crop and weed leaves can be extracted accurately. On the other hand, CNN models may be a better choice for applications requiring automatic feature extraction through, which can be achieved through the convolutional operators.

4.7 Appendix

Deep convolutional neural networks (DCNNs)

4.7.1 Visual Geometry Group (VGG) architecture

The VGG architecture first proposed by K. Simonyan and A. Zisserman at University of Oxford in 2014 [66], where VGG architectures with 16 layers and 19 layers were particularly presented due to their

major improvements in classification experiments from 11 layers to 19 layers [66]. Particularly, the characteristics in the network design included five blocks of convolutional layers using 3 × 3 filters, to reduce the number of model parameters, and three fully-connected layers. The first two fully-connected layers had 4,096 channels with a rectified linear unit (ReLU) activation function at each layer and third layer consisted of 1,000 channels with a softmax activation function. In addition, maximum pooling layers used 2 × 2 filters placed behind the 2nd, 4th, 7th, 10th, and 13th convolution layers for VGG-16 and 2nd, 4th, 8th, 12th, and 16th for VGG-19 in order to enhance the feature expression [66]. The detailed architectures of these networks are shown in Figure 7. Although the performance of VGG-16 and VGG-19 models were obviously appreciated in the competition ILSVRC, these models have some limitations. They use a considerable amount of memory for the optimization of the learning parameters and a great number of parameters, approximately 138 and 143 million parameters for VGG-16 and VGG-19, respectively.

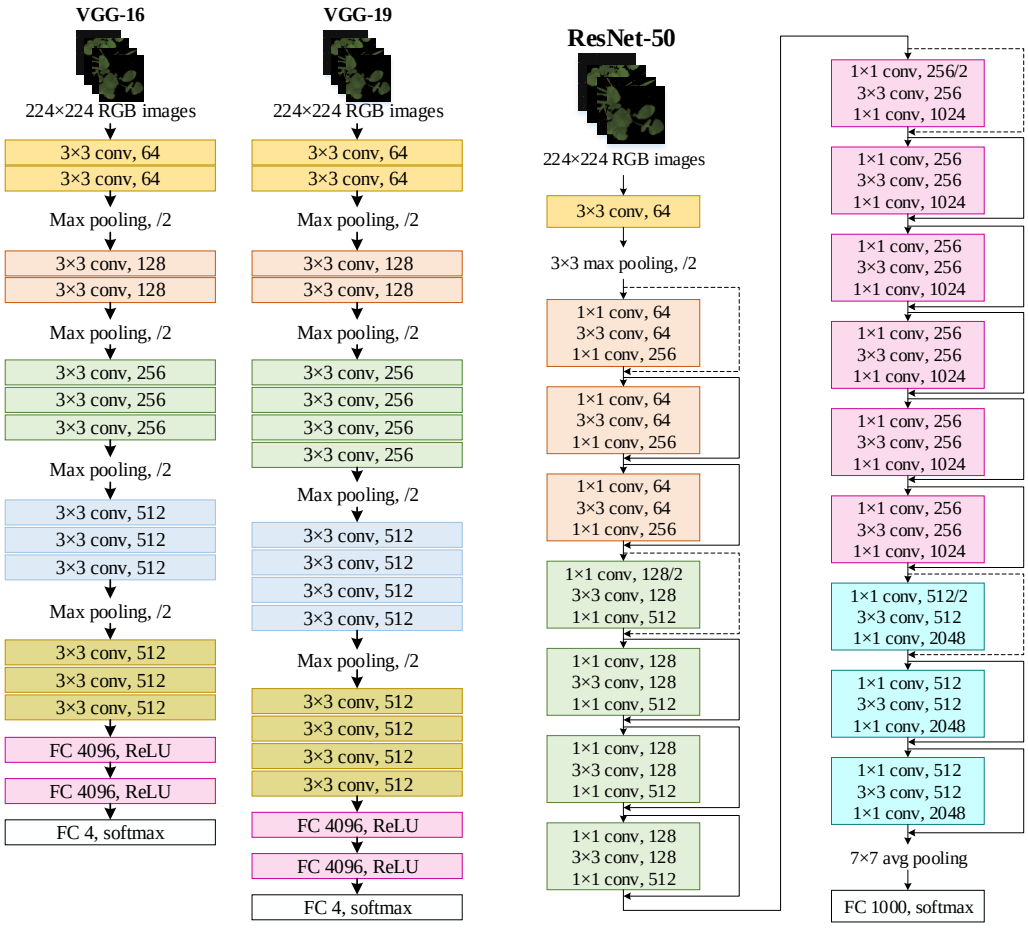


Figure 7. VGG-16, VGG-19 and ResNet architectures.

4.7.2 ResNet architecture

When the neural network layers were increased, researchers found that the challenging problems of training deeper neural networks were the vanishing gradient and accuracy degradation [67, 68].

Consequently, a new network, namely the deep residual network (ResNet) has been proposed, based on using shortcut connections in order to skip blocks of convolutional layers and form residual blocks. Although skipping blocks were implemented, the information integrity was protected. Hence, the next layer only learns a different part between the input and output of the previous layer. This relieved the difficulty of convergence and simplified the learning process. As for the ResNet-50 model, it had 50 weight layers and approximately 23.5 million trainable parameters [53]. The building blocks used a stack of 3 layers including 1×1, 3×3, and 1×1 convolutional layers. The 1×1 layers placed in each block played an important role in reducing and then restoring dimensions. In addition, the down-sampling was conducted directly by convolutional layers that have a stride of 2, and batch normalization was applied right before ReLU activation function and after each convolutional layer. When the dimensions of the input and output were similar, the identity shortcut was used. As for the increased dimensions, the projection shortcut was used to match dimensions through 1×1 convolutional layers. In both cases, when the shortcuts went across feature maps of two sizes, they were performed with a stride of 2. The ResNet architecture, shown in Figure 7, ended with a 1,000 fully-connected layers using softmax activation function.

4.7.3 Inception-V3 architecture

The next architecture used in this paper was Inception-V3 [52], which is an improved version of the GoogleNet architecture, especially Inception-V1 [51] and InceptionV2 [69]. The flowchart of the state-of-the-art Inception-V3 architecture in image classification is illustrated in detail in Figure 8. This architecture comprises approximately 23 million parameters. The network factorized convolutions in a computationally efficient manner including the use of larger filter sizes from 1 × 1, 3 × 3, 5 × 5, or 1 × 7 convolutions followed by 7 × 1 convolutions. These filters produced very good results. In addition, in order to accelerate the convergence of the network, auxiliary classifiers with a regularization effect were introduced. This parallel structures and dimensional reduction of the Inception modules were responsible for mitigating the effect of structural changes on nearby components. Next, pooling layers were added to achieve more efficient grid size reduction [52].

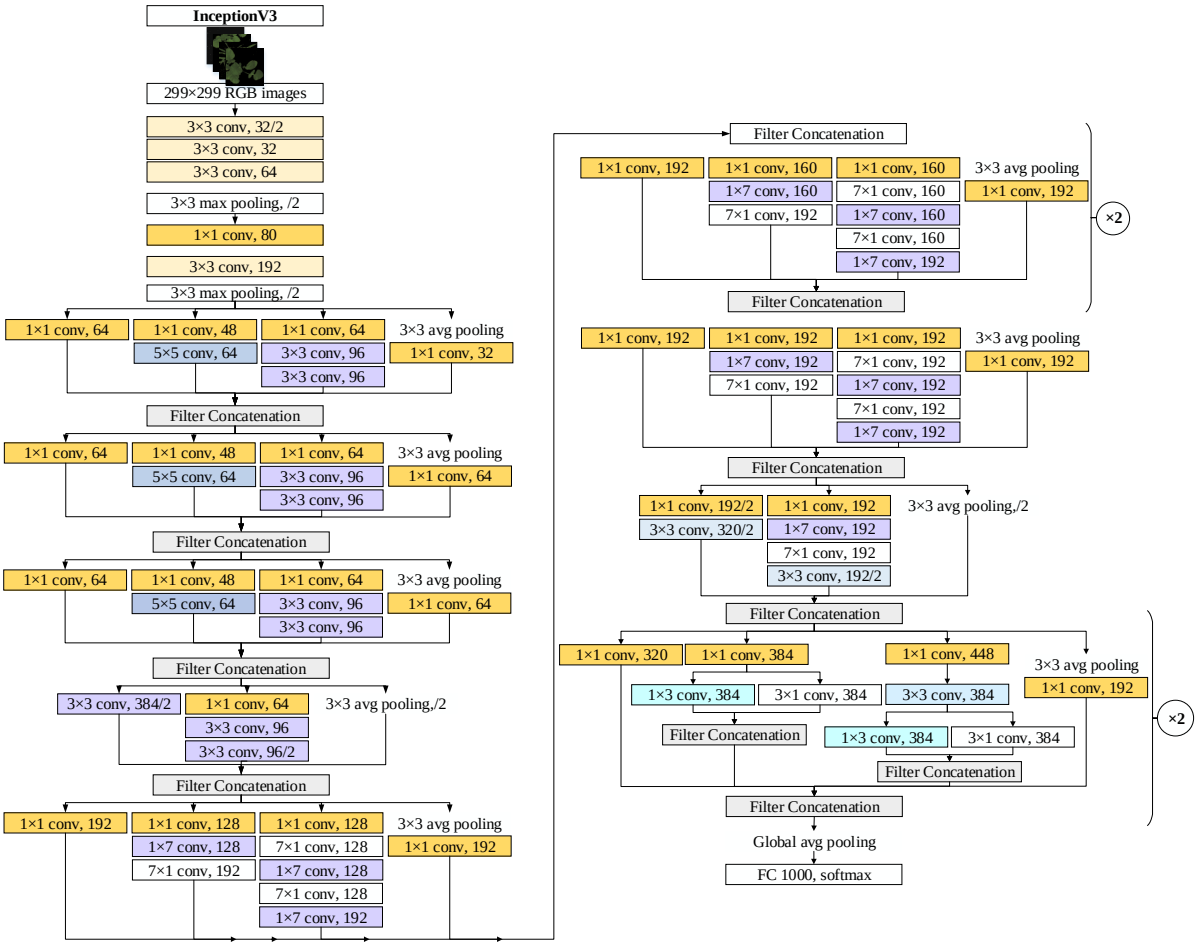


Figure 8. Flowchart of the Inception-V3 architecture.

Despite the progress attained in agriculture in recent years [4, 70], there are still gaps to be explored to better understand the performance of CNNs and conventional ML methods using the crop and weed dataset “bccr-segset”. In this work, we investigated and compared the use of transfer learning for VGG, Inception and ResNet architectures against the k-FLBPCM method for the detection of broadleaf crops and weeds.

4.8 References

[1] M. Aitkenhead, I. Dalgetty, C. Mullins, A. J. S. McDonald, and N. J. C. Strachan, "Weed and crop discrimination using image analysis and artificial intelligence methods," *Computers and electronics in Agriculture*, vol. 39, no. 3, pp. 157-171, 2003.

[2] K.-H. Dammer and G. Wartenberg, "Sensor-based weed detection and application of variable herbicide rates in real time," *Crop protection*, vol. 26, no. 3, pp. 270-277, 2007.

[3] R. Gerhards and S. Christensen, "Real-time weed detection, decision making and patch spraying in maize, sugarbeet, winter wheat and winter barley," *Weed research,* vol. 43, no. 6, pp. 385-392, 2003.

[4] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Computers and electronics in agriculture,* vol. 147, pp. 70-90, 2018.

[5] K. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, "Machine learning in agriculture: A review," *Sensors,* vol. 18, no. 8, p. 2674, 2018.

[6] F. Pallottino, M. Biocca, P. Nardi, S. Figorilli, P. Menesatti, and C. Costa, "Science mapping approach to analyze the research evolution on precision agriculture: world, EU and Italian situation," *Precision Agriculture,* vol. 19, no. 6, pp. 1011-1026, 2018.

[7] F. Pallottino *et al.*, "Machine vision retrofit system for mechanical weed control in precision agriculture applications," *Sustainability,* vol. 10, no. 7, p. 2209, 2018.

[8] A. Wang, W. Zhang, and X. Wei, "A review on weed detection using ground-based machine vision and image processing techniques," *Computers and electronics in agriculture,* vol. 158, pp. 226-240, 2019.

[9] A. Torralba and A. A. Efros, "Unbiased look at dataset bias," in *CVPR*, 2011, vol. 1, no. 2: Citeseer, p. 7.

[10] D. Slaughter, D. K. Giles, and D. Downey, "Autonomous robotic weed control systems: A review," *Computers and electronics in agriculture,* vol. 61, no. 1, pp. 63-78, 2008.

[11] R. B. Brown and S. D. Noble, "Site-specific weed management: sensing requirements—what do we need to see?," *Weed Science,* vol. 53, no. 2, pp. 252-258, 2005.

[12] A. Bakhshipour and A. Jafari, "Evaluation of support vector machine and artificial neural networks in weed detection using shape features," *Computers and Electronics in Agriculture,* vol. 145, pp. 153-160, 2018.

[13] A. AlSuwaidi, C. Veys, M. Hussey, B. Grieve, and H. Yin, "Hyperspectral selection based algorithm for plant classification," in *2016 IEEE International Conference on Imaging Systems and Techniques (IST)*, 2016: IEEE, pp. 395-400.

[14] I. Herrmann, U. Shapira, S. Kinast, A. Karnieli, and D. Bonfil, "Ground-level hyperspectral imagery for detecting weeds in wheat fields," *Precision agriculture,* vol. 14, no. 6, pp. 637-659, 2013.

[15] P. Symonds, A. Paap, K. Alameh, J. Rowe, and C. Miller, "A real-time plant discrimination system utilising discrete reflectance spectroscopy," *Computers and Electronics in Agriculture*, vol. 117, pp. 57-69, 2015.

[16] H. S. Midtiby, B. Åstrand, O. Jørgensen, and R. N. Jørgensen, "Upper limit for context–based crop classification in robotic weeding applications," *Biosystems engineering*, vol. 146, pp. 183-192, 2016.

[17] X. Wu, W. Xu, Y. Song, and M. Cai, "A detection method of weed in wheat field on machine vision," *Procedia Engineering*, vol. 15, pp. 1998-2003, 2011.

[18] H. Liu, S. H. Lee, and C. Saunders, "Development of a machine vision system for weed detection during both of off-season and in-season in broadacre no-tillage cropping lands," 2014.

[19] A. Bakhshipour, A. Jafari, S. M. Nassiri, and D. Zare, "Weed segmentation using texture features extracted from wavelet sub-images," *Biosystems engineering*, vol. 157, pp. 1-12, 2017.

[20] M. H. Bharati, J. J. Liu, and J. F. MacGregor, "Image texture analysis: methods and comparisons," *Chemometrics and intelligent laboratory systems*, vol. 72, no. 1, pp. 57-71, 2004.

[21] P. Lottes, M. Hoeferlin, S. Sander, M. Müter, P. Schulze, and L. C. Stachniss, "An effective classification system for separating sugar beets and weeds for precision farming applications," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, 2016: IEEE, pp. 5157-5163.

[22] A. Nieuwenhuizen, L. Tang, J. Hofstee, J. Müller, and E. Van Henten, "Colour based detection of volunteer potatoes as weeds in sugar beet fields using machine vision," *Precision Agriculture*, vol. 8, no. 6, pp. 267-278, 2007.

[23] N. Peter, W. Mike, and C. John, "Wild radish management and strategies to address herbicide resistance," Grains Research Development Corporation, 2014.

[24] D. Lemerla, "Wild radish threatens canola yields: mixed methods the answer by Alec Nicol," in *GroundCover*, ed, 2002.

[25] V. N. T. Le, B. Apopei, and K. Alameh, "Effective plant discrimination based on the combination of local binary pattern operators and multiclass support vector machine methods," *Information processing in agriculture*, vol. 6, no. 1, pp. 116-131, 2019.

[26] Z. Guo, L. Zhang, and D. Zhang, "A completed modeling of local binary pattern operator for texture classification," *IEEE transactions on image processing,* vol. 19, no. 6, pp. 1657-1663, 2010.

[27] M. Heikkilä, M. Pietikäinen, and C. Schmid, "Description of interest regions with local binary patterns," *Pattern recognition,* vol. 42, no. 3, pp. 425-436, 2009.

[28] S. R. Dubey and A. S. Jalal, "Detection and classification of apple fruit diseases using complete local binary patterns," in *2012 Third International Conference on Computer and Communication Technology*, 2012: IEEE, pp. 346-351.

[29] H. Waghmare, R. Kokare, and Y. Dandawate, "Detection and classification of diseases of Grape plant using opposite colour Local Binary Pattern feature and machine learning for automated Decision Support System," in *2016 3rd international conference on signal processing and integrated networks (SPIN)*, 2016: IEEE, pp. 513-518.

[30] L. Nanni, A. Lumini, and S. Brahnam, "Survey on LBP based texture descriptors for image classification," *Expert Systems with Applications,* vol. 39, no. 3, pp. 3634-3641, 2012.

[31] M. Pietikäinen and G. Zhao, "Two decades of local binary patterns: A survey," in *Advances in independent component analysis and learning machines*: Elsevier, 2015, pp. 175-210.

[32] V. N. T. Le, S. Ahderom, B. Apopei, and K. Alameh, "A novel method for detecting morphologically similar crops and weeds based on the combination of contour masks and filtered Local Binary Pattern operators," *GigaScience,* vol. 9, no. 3, 2020, doi: 10.1093/gigascience/giaa017.

[33] J. Charters, Z. Wang, Z. Chi, A. C. Tsoi, and D. D. Feng, "Eagle: a novel descriptor for identifying plant species using leaf lamina vascular features," in *2014 IEEE international conference on multimedia and expo workshops (ICMEW)*, 2014: IEEE, pp. 1-6.

[34] J. S. Cope, P. Remagnino, S. Barman, and P. Wilkin, "The extraction of venation from leaf images by evolved vein classifiers and ant colony algorithms," in *International Conference on Advanced Concepts for Intelligent Vision Systems*, 2010: Springer, pp. 135-144.

[35] A. Kadir, L. E. Nugroho, A. Susanto, and P. I. Santosa, "Leaf classification using shape, color, and texture features," *arXiv preprint arXiv:1401.4447*, 2013.

[36] N. Kumar *et al.*, "Leafsnap: A computer vision system for automatic plant species identification," in *European Conference on Computer Vision*, 2012: Springer, pp. 502-516.

[37] T. Beghin, J. S. Cope, P. Remagnino, and S. Barman, "Shape and texture based plant leaf classification," in *International Conference on Advanced Concepts for Intelligent Vision Systems*, 2010: Springer, pp. 345-353.

[38] S. H. Lee, C. S. Chan, P. Wilkin, and P. Remagnino, "Deep-plant: Plant identification with convolutional neural networks," in *2015 IEEE International Conference on Image Processing (ICIP)*, 2015: IEEE, pp. 452-456.

[39] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, p. 436, 2015.

[40] Y. LeCun and Y. Bengio, "Convolutional networks for images, speech, and time series," *The handbook of brain theory and neural networks*, vol. 3361, no. 10, p. 1995, 1995.

[41] J. Tang, D. Wang, Z. Zhang, L. He, J. Xin, and Y. Xu, "Weed identification based on K-means feature learning combined with convolutional neural network," *Computers and electronics in agriculture*, vol. 135, pp. 63-70, 2017.

[42] I. Sa *et al.*, "weednet: Dense semantic weed classification using multispectral images and mav for smart farming," *IEEE Robotics and Automation Letters*, vol. 3, no. 1, pp. 588-595, 2017.

[43] A. Milioto, P. Lottes, and C. Stachniss, "Real-time semantic segmentation of crop and weed for precision agriculture robots leveraging background knowledge in CNNs," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, 2018: IEEE, pp. 2229-2235.

[44] H. Okamoto, T. Murata, T. Kataoka, and S. I. HATA, "Plant classification for weed detection using hyperspectral imaging with wavelet analysis," *Weed Biology and Management*, vol. 7, no. 1, pp. 31-37, 2007.

[45] C. Yang, S. O. Prasher, J. Landry, and A. DiTommaso, "Application of artificial neural networks in image recognition and classification of crop and weeds," *Canadian agricultural engineering*, vol. 42, no. 3, pp. 147-152, 2000.

[46] F. Ahmed, H. A. Al-Mamun, A. H. Bari, E. Hossain, and P. Kwan, "Classification of crops and weeds from digital images: A support vector machine approach," *Crop Protection*, vol. 40, pp. 98-104, 2012.

[47] S. Haug, A. Michaels, P. Biber, and J. Ostermann, "Plant classification system for crop/weed discrimination without segmentation," in *IEEE winter conference on applications of computer vision*, 2014: IEEE, pp. 1142-1149.

[48] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural networks*, vol. 61, pp. 85-117, 2015.

[49] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097-1105.

[50] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[51] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1-9.

[52] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818-2826.

[53] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770-778.

[54] O. Russakovsky *et al.*, "Imagenet large scale visual recognition challenge," *International journal of computer vision*, vol. 115, no. 3, pp. 211-252, 2015.

[55] S. H. Lee, C. S. Chan, S. J. Mayo, and P. Remagnino, "How deep learning extracts and learns leaf features for plant classification," *Pattern Recognition*, vol. 71, pp. 1-13, 2017.

[56] H. Yalcin and S. Razavi, "Plant classification using convolutional neural networks," in *2016 Fifth International Conference on Agro-Geoinformatics (Agro-Geoinformatics)*, 2016: IEEE, pp. 1-5.

[57] M. A. Hedjazi, I. Kourbane, and Y. Genc, "On identifying leaves: A comparison of CNN with classical ML methods," in *2017 25th Signal Processing and Communications Applications Conference (SIU)*, 2017: IEEE, pp. 1-4.

[58] A. Kamilaris, F. X. J. C. Prenafeta-Boldú, and e. i. agriculture, "Deep learning in agriculture: A survey," vol. 147, pp. 70-90, 2018.

[59] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern recognition*, vol. 29, no. 1, pp. 51-59, 1996.

[60] M. Pietikäinen, A. Hadid, G. Zhao, and T. Ahonen, *Computer vision using local binary patterns*. Springer Science & Business Media, 2011.

[61] C. Shan, S. Gong, and P. W. McOwan, "Facial expression recognition based on local binary patterns: A comprehensive study," *Image and vision Computing,* vol. 27, no. 6, pp. 803-816, 2009.

[62] S. Brahnam, L. C. Jain, L. Nanni, and A. Lumini, *Local binary patterns: new variants and applications*. Springer, 2014.

[63] C. Zhu, C.-E. Bichot, and L. Chen, "Image region description using orthogonal combination of local binary patterns enhanced with color information," *Pattern Recognition,* vol. 46, no. 7, pp. 1949-1963, 2013.

[64] D. Huang, C. Shan, M. Ardabilian, Y. Wang, and L. Chen, "Local binary patterns and its application to facial image analysis: a survey," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews),* vol. 41, no. 6, pp. 765-781, 2011.

[65] C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, and C. Liu, "A survey on deep transfer learning," in *International Conference on Artificial Neural Networks*, 2018: Springer, pp. 270-279.

[66] K. Simonyan and A. J. a. p. a. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014.

[67] A. Veit, M. J. Wilber, and S. Belongie, "Residual networks behave like ensembles of relatively shallow networks," in *Advances in neural information processing systems*, 2016, pp. 550-558.

[68] O. Wichrowska *et al.*, "Learned optimizers that scale and generalize," in *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, 2017: JMLR. org, pp. 3751-3760.

[69] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv preprint arXiv:1502.03167,* 2015.

[70] A. Kamilaris and F. Prenafeta-Boldú, "A review of the use of convolutional neural networks in agriculture," *The Journal of Agricultural Science,* vol. 156, no. 3, pp. 312-322, 2018.

Chapter 5 – Detecting weeds from crops under complex field environments based on Faster RCNN models

This chapter was submitted to the Precision Agriculture Journal. The manuscript has been changed to the layout, number formats, font size and font style, which was implemented to maintain consistency in the formatting of this thesis.

5.1 Abstract

The power of deep learning in object detection has widely been investigated, demonstrating promising results in recent years. In precision agricultural applications, weed detection plays an indispensable part in site-specific weed management. The published resources of crop and weed datasets under complex field environments including lighting conditions, weather conditions, different growth stages, heavy occlusion and overlap, and weeds with similar properties are limited. In this chapter, we provide a FT_BRC image dataset (published online with 3380 images) collected by a camera installed on a portable trolley under practical field environments from a commercial farm in Cunderdin, Western Australia. Based on their harmful effects on the crop yield, Wild radish (Raphanus raphanistrum) and Capeweed (Arctotheca calendula) weed detection in Barley crops (Hordeum vulgare) is investigated. In the context of locating targeted weeds and estimating weed density, we fully annotate a part of the dataset and use the Faster RCNN model with different feature extractors for weed detection in the field. Experimental results show that the mean average precision (mAP) of the Faster RCNN model with Inception-ResNet-V2 network with 0.555 (at IoU =0.5) is higher compared to other networks and the inference time of this model was approximately 0.38 seconds per image. These results can support to further the development of solutions for weed detection in real-time precision agriculture.

Keywords: Weed detection, Hordeum vulgare, Raphanus raphanistrum, Arctotheca calendula, Faster RCNN.

5.2 Introduction

Weeds pose a serious threat to farmers and producers as they have detrimental effects on crops such as competition for nutrients, space, water and light, reduction of agricultural productivity, and crowding out indigenous biodiversity [1]. Wild radish (Raphanus Raphanistrum) weed is one of the prevalent broadleaf weed species, vigorously competitive and difficult to control in Australia because wild radish seedlings establish rapidly and have a faster growth rate than the crop [2]. This can cause significant crop yield losses of up to 90% [3]. Although wild radish seeds germinate during autumn and winter, they can emerge all year round with sufficient soil moisture [4]. Therefore, detecting wild radish at early stages plays an important role in controlling weeds in cereal crops. Another type of weed investigated

in this study is capeweed (Arctotheca Calendula) which also competes with cereal crops, increases nitrate and nitrite toxicity, and poisons ruminants [5, 6]. According to a report [7], there was approximately 7 to 90 capeweed plants/m^2 in Western Australia and 76% of cereal crops had capeweed invasions in New South Wales. Barley (Hordeum vulgare**)** is an important cereal crop, and important contributor to the Australian agricultural market [8].

Generally, the goal of weed detection is to identify where weeds are located in a given image and which their types. However, it is difficult to accurately detect and classify weeds under complex field environments, such as illumination variation, poses, viewpoints, and occlusions. A variety of conventional machine learning and deep convolutional neural network (DCNN)-based methods for object detection have recently been reported [9-13]. Particularly, in the precision agricultural practices, various approaches have been attempted to detect weed species [14]. For example, the development of weed detection in sugar beet fields with occlusion and overlapping problems was conducted by combining local and global texture features [15]. The main drawback of this method is the inability to accurately detect the weed locations in the images. Another method for detecting weeds in soybean crop images captured by Unmanned Aerial Vehicles (UAV) was implemented using Convolutional Neural Networks (CNNs) [16]. A different study on discriminating clover from weeds and grass using CNNs was reported in [17]. These methods were mainly designed for segmenting plant patches in different species, not for detecting individual plants and different weeds in images. A method based on using Single Shot Detection (SSD) was explored for the detection of weeds in cereal fields with leaf occlusion [18]. This method requires a large number of crop images with large-size weeds since large-size weed detection (heavy occlusion) is difficult through SSD. In contrast, another study applied CNNs to detect weeds in close-range imagery of agricultural fields. However, this method has limited capability to small weed detection, and weeds obscured by other plants were not able to be identified under simple field environments [19]. Yet another study focused on the detection of maize seedling under different growth stages and field environments by using Faster RCNN [20]. Its limitation in detecting a single crop makes it unsuitable for operation on practical field conditions. A comparison analysis of two deep learning frameworks including Faster RCNN and YOLOv3 [21] with ResNet-50, ResNet-101 and Darknet53 was conducted to explore the development of a smart sprayer system that can achieve real-time weed detection with high accuracy [22]. Another study on detecting cotton plant seedlings and weeds by using the Faster RCNN model with Inception-ResNet-V2 [23]. However, the accuracies of these methods are low because the datasets used to train the algorithms were for small plants at just one growth stage, namely seedlings, where no difficulties are encountered in terms of occlusion or overlap between crops and weeds. In the field, weeds typically appear in varied patch sizes. Therefore, to develop an accurate weed detection method, it is necessary to detect the precise positions of the target weeds at different growth stages, under occlusion, overlapping and weather conditions.

In recent years, the performances of object detection techniques have been improved by taking advantages of DCNNs in learning robust and high-level features of images. Particularly, in 2014, Girshick et al. proposed the application of convolutional neural network features (RCNN) to regions of images for detecting objects [24]. Another method, called Spatial Pyramid Pooling Networks (SPPNet) [25], was also published in 2014 overcoming the issue of repetitive computation for convolutional features and demonstrating 20 times faster computation speed than their RCNN counterparts on Pascal VOC07 dataset. However, the limitation of this method was fine tuning of the fully connected layers. In 2015, a fast-RCNN was proposed to train a detector and a bounding box regressor with the same network configurations and improve its mean average precision (mAP) and detection speed by around 200 times compared to RCNN [26]. Simultaneously, faster-RCNN was proposed by S. Ren et al., and integrated into an end-to-end learning framework [27]. The main idea of this approach is to use Region Proposal Network (RPN) and improve the speed bottleneck of Fast-RCNN. This made the mAP of Faster-RCNN detector higher than other detectors on Pascal VOC07, Pascal VOC12, COCO [28].

The main contributions of this Chapter are:
(i) Collecting a dataset of plants under complex environment (significant occlusion) and complicated weather conditions (sunny, windy, cloudy, etc.) and annotating weeds.
(ii) Comparing Faster RCNN models and finding an appropriate model for Raphanus raphanistrum and Arctotheca calendula detection.
(iii) Generating Faster RCNN models and evaluating the performance of a weed detector in field conditions.

The Chapter is organised as follows. The background of deep neural networks, especially R-CNN networks, is described in Section 5.3. Section 5.4 describes the methodology of the research work, Section 5.5 discusses the results, and finally, Section 5.6 draws a conclusion and provides future research directions.

5.3 Materials

5.3.1 Faster RCNN Overview

Recent object detection neural networks have utilized “region proposal algorithms” to generate object locations. However, the computational cost of traditional region proposal algorithms is still high, making their slow run-time impractical for real-time applications. In our work, CNN based Faster RCNN models were applied to detect different weed species in barley crops. Figure shows our pipeline that uses the Faster RCNN for real-time weed detection.

In general, the Faster RCNN consists of two main parts: RPN to generate proposal regions and Fast RCNN detector to classify the regions [27]. Faster RCNN uses the region proposal network (RPN) to

extract object locations in 2D images. RPN shares some convolutional layers with recent object detection networks. Particularly, to generate region proposals, a small network is added on top of shared convolutional feature map. This small network helps extract lower-dimensional features from feature map outputs. Then, a box-classification layer is used to estimate whether there is an object in the proposal while a box-regression layer outputs coordinates the boxes. RPN plays a major role in the Faster RCNN pipeline as it makes a significant impact on the accuracy of the classifier [27].

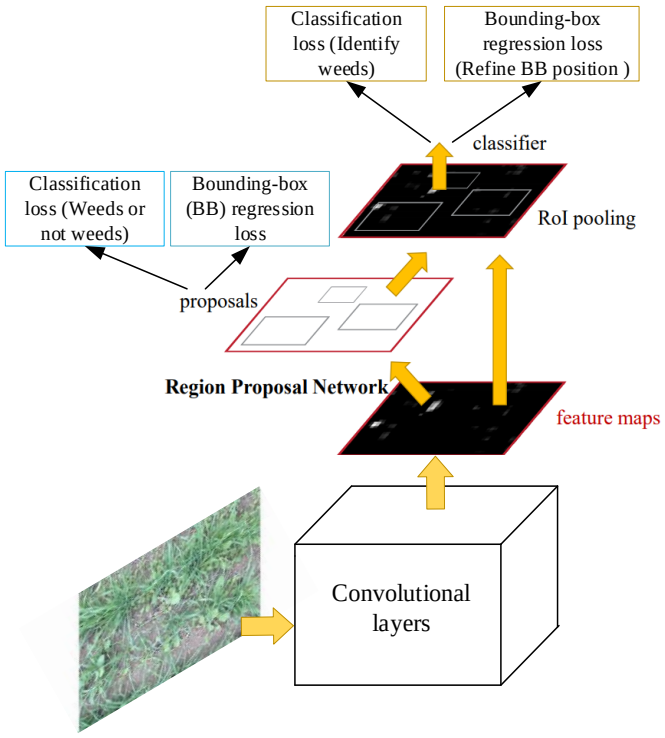


Figure 1. Illustration of the Faster R-CNN model for weed detection [27]

5.4 Methodology

5.4.1 Data collection

We collected the FT_BRC (online) with 3380 images under complex weather conditions from a commercial farm at Cunderdin, Western Australia. All images were captured by a portable trolley equipped with a Xilinx Zynq ZC702 development board and an On-Semi VITA 2000 camera as shown in Figure 2 (designed by and installed at the Electron Science Research Institute, Edith Cowan University) [29, 30]. The On-Semi VITA 2000 camera sensor was installed on the board to capture 1920 ×1080 images at 60 frames per second and a spatial resolution of ≈1mm/pixel. Barley crops were collected under various weather conditions (cloudy, sunny, and windy), heavy occlusion and overlap, and the different growth stages of weeds (wild radish and capeweed). This dataset emulates challenging scenarios encountered in real agricultural fields.

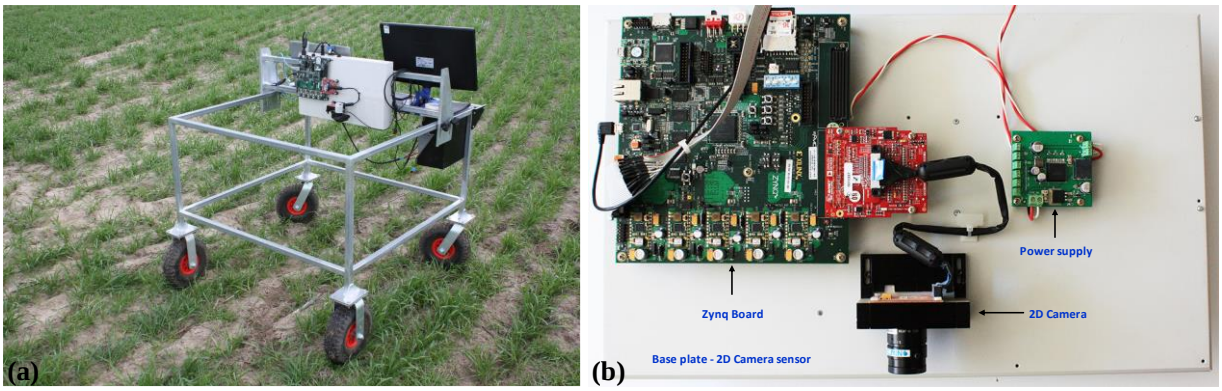


Figure 2. (a) Portable trolley equipped with a Xilinx Zynq ZC702 development board and an On-Semi VITA 2000 camera for capturing plant images in complex field environments. (b) close-up photo of the weed detection sensor.

We detected two different types of weeds including wild radish and capeweed weeds, which significantly contribute to the reduction of barley crop yield in farms. During the growing stages, wild radish may be confused with capeweed due to their highly similar visual appearance and their germination during autumn and winter during results in similar growth patterns, as shown in Figure 3. When these weeds are growing to a certain stage, their subsequent leaves grow singly and are deeply lobed with a rounded apex [7]. This creates difficulties in detecting these weeds under practical field conditions. However, it is important to note that the main differences of these weeds are (i) the underside of wild radish leaves having a darker shade of green while capeweed leaves having a lighter green colour [7], and (ii) the terminal lobes of wild radish leaves growing relatively larger than capeweed leaves. With the images captured on the upper side of the plants, obviously, the first difference cannot be considered in this situation. Therefore, the weeds are identified in the field by relying on the terminal lobes of leaves.

Figure 3. Appearance of (a) Capeweed and (b) Wild radish under complex field environments.

5.4.2 Data labelling

After collecting plant images in the field, a “LabelImg” software was chosen to draw bounding boxes around various weeds at different growth stages. This graphical image annotation tool is an open-source

widget toolkit for generating a graphical user interface. As illustrated in Figure 4, the interface of the LabelImg tool presents an image of crops and weeds collected under heavy leaf occlusion with green bounding boxes for wild radish and white bounding boxes for capeweed. Additionally, it is important to note that field images were formatted as .xml files with Pascal VOC data format. A part of the FT_BRC dataset was fully annotated with bounding boxes and considered as the ground truth labelled plant images used for further developing weed research purposes. The training set with bounding boxes consisted of 258 images and the testing set contains 65 images. The number of images in the training and testing set was limited due to multiple bounding boxes annotated in each image. Particularly, the numbers of bounding boxes in the training and testing set were 2108 and 460, respectively. In our dataset, images with heavy occlusion and overlap could contain up to 40 bounding boxes per image.

Figure 4. A LabelImg tool to label plant images under practical field conditions with multiple bounding boxes.

5.4.3 Faster RCNN with Inception-Resnet-V2

To improve performance of Faster RCNN pipeline, researchers have proposed several ways that help extract better features for the RPN as well as the classifier. There are two obvious ways to achieve this. One way is trying to increase the width and height of the original backbone network without changing its architecture. The other way is to change the backbone architecture by using state-of-the-art feature extraction networks (e.g. VGG, ResNet, Inception, etc). The former way requires much data to follow as the network becomes deeper and deeper. Otherwise, the network gets overfit easily. The latter way looks more promising because several architectures have been proposed recently, which achieve better performance on ImageNet dataset.

Szegedy et al. [31, 32] have proposed Inception-Resnet-v2 that combines two well-known architectures: Residual connections proposed by He et al. [33] and Inception architecture. Here, they replaced the concatenate stage in the inception network with residual connections. This improvement helps the inception architecture take the advantages of residual connections without losing its computational efficiency. Figure 5 illustrates the architecture of an Inception-Resnet-v2 layer. Compared to some recent deep networks, Inception-Resnet-v2 have achieved state-of-the-art accuracy in ILSVRC-2012 validation set [31]. This combination of features makes Inception-Resnet-v2 architecture a useful addition to our comparison of selected architectures.

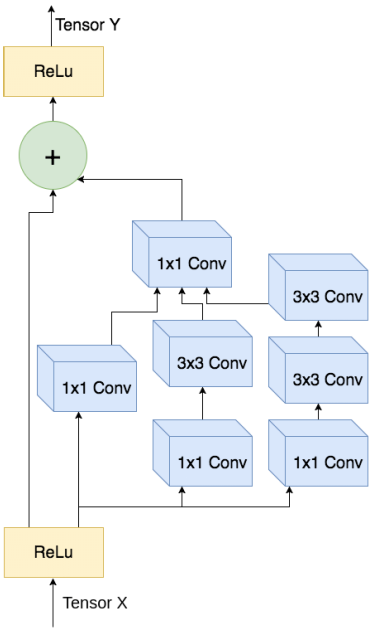


Figure 5. Structure of an Inception-ResNet-V2 layer [34]

5.4.4 Performance metrics

The evaluation metrics including precision and recall were computed in this study. The performance of the neural network at positive defection was measured by the precision value, while the effectiveness of the neural network to recognize the targeted weeds was measured by the recall value. Particularly, the high recall value shows the high successful rate of detecting the target weeds, whereas the high precision value illustrates the high successful rate of detecting areas where weeds (wild radish and capeweed) do not appear. The precision and recall metrics were calculated as follows [35-37]:

$$\text{Recall} = \frac{\text{True Positive (TP)}}{\text{True Positve (TP)} + \text{False Negative (FN)}} \quad (1)$$

$$\text{Precision} = \frac{\text{True Positive (TP)}}{\text{True Positve (TP)} + \text{False Positive (FP)}} \quad (2)$$

For object detection problems, we evaluated our pipeline performance based on common metrics including mean average precision (mAP), average recall (AR) and Intersection over Union (IoU). In our scenario, the IoU metric was used to measure the accuracy of a weed detector on the FT_BRC dataset. This metric was calculated by diving the area of intersection between the bounding boxes of the predictions and ground truth labels by the area of union of the two bounding boxes, where the predicted bounding boxes are generated by our model and ground-truth bounding boxes were manually labelled by the LabelImg image annotation tool. The IoU thresholds should be set in a range from 0.5 to 0.95. In this case, the correct weed detection can be classified as True Positive (TP), if IoU $\geq$ threshold value, the wrong weed detection can be classified as False Positive (FP), if IoU $\leq$ threshold value. False Negative (FN) presents that the ground-truth bounding boxes of weed species are not detected by the Faster RCNN model.

5.4.5 Training Faster RCNN for real-time weed detection

Motivated by [27], we followed a pragmatic 4-step training algorithm to train our chosen Faster RCNN. At the beginning, the RPN was trained independently. We initialized the backbone CNN with an ImageNet pre-trained model and fine-tuned for the region proposal task. Due to the scale difference and aspect ratios of weeds, anchors were used in the RPN and set to be positive if they had the highest IoU and the overlapping rate between the anchor and ground-truth box was higher than 0.7. If the overlapping rate was smaller than 0.3, the anchor was set as a negative sample. To limit the number of negative samples, we balanced the number of positive and negative samples (with ratio 1:1). Then, Fast RCNN was used as a detector for Faster RCNN [27]. The detection model (Fast RCNN) was trained by using region proposals produced from our RPN. Note that the Fast RCNN was also initialized with an ImageNet pre-trained model. In the next step, the RPN training was initialized by the Fast RCNN and only unique layers of RPN were fine-tuned while the shared convolutional layers were fixed. In the last step, we again fixed the shared convolutional layers and fine-tuned only the unique layers of the Faster RCNN. Due to the limitation of data, it was challenging to train whole pipeline from scratch, and we utilized transfer learning by leverage check-point model, which was pre-trained with MS-COCO dataset [38]. MS-COCO is a dataset published in 2015, which consists of 80 object categories. The dataset includes more than 300,000 images with around 2.5 million labeled instances.

5.5 Results and Discussions

All experiments were executed using the Ubuntu 18.04 LTS operating system and GeForce GTX1080Ti card. Note that all of our models were implemented in TensorFlow. We divided a part of dataset with bounding box annotations in xml files into 258 images for the training set and 65 images for the test set. We conducted a comparison among the Faster RCNN models using different CNN architectures, including ResNet-50, ResNet-101, Inception-V2 and Inception-ResNet-V2, in order to find out the best

model that exhibits the best performance with our weed dataset in the field.

For Faster RCNN models used in conjunction with Inception-ResNet-V2 architectures, Figure 6 illustrates all typical training loss graphs with our dataset and all of the losses significantly dropped throughout the training phase. The decrease in training losses means that the model was learning during the training session. Figure 6 was exported from a visualization tool-TensorBoard. In addition, the training time for each step was approximately 1.36 seconds using NVIDIA GeForce GTX1080Ti. During the comparison among Faster RCNN models, we set the learning rate to 0.0003 and trained the pipeline with 200 epochs. The number of classes was changed to 2.

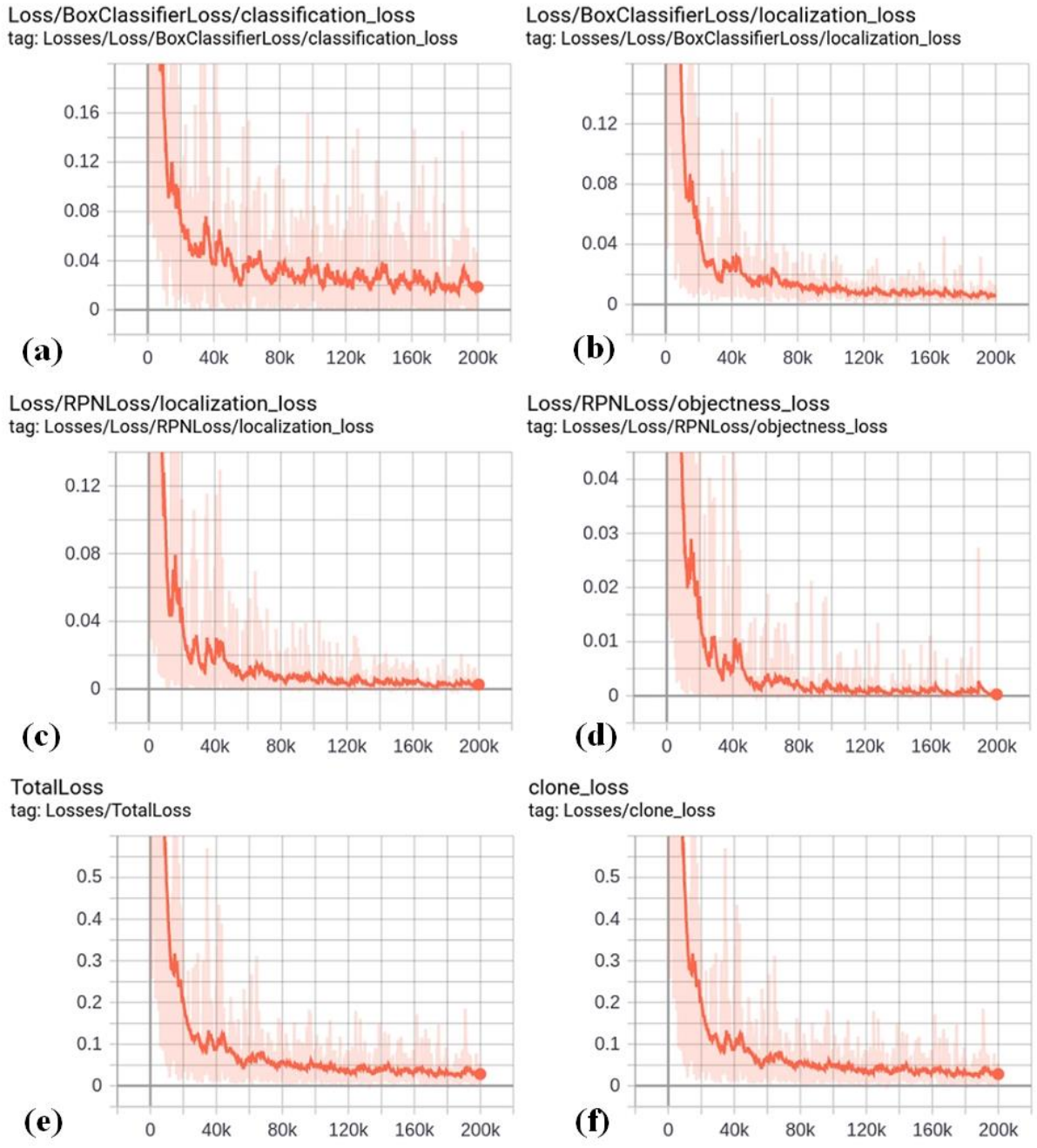


Figure 6. Loss plots for training the Inception-ResNet-V2 based faster RCNN model with the FT_BRC dataset. (a) Classification loss of detected weeds, (b) Localization loss (the loss of the bounding box regressor), (c) RPN localization loss, (d) RPN objectness loss, (e) Total loss and (f) Clone loss.

According to the experimental results on the test set, Table 1 shows the performance of Faster RCNN models with selective backbone architectures including ResNet-50, ResNet-101, Inception-V2 and Inception-ResNet-V2. As can be observed in Table 1, the performance of the Faster RCNN model with Inception-ResNet-V2 was higher than other models. Particularly, the mean average precision (mAP) of Faster RCNN model with Inception-ResNet-V2 was 0.289, while the mAP values with IoU of 0.5 and 0.75 were 0.555 and 0.297, respectively. The average recall (AR) values for the Faster RCNN model with Inception-ResNet-V2 were 0.148, 0.366 and 0.433 for 1, 10 and 100 detections per image, respectively. The performance plots illustrating the mAP and AR values for the Faster RCNN Inception-ResNet-V2 model are shown in Figure 7.

Table 1. Performance of detection model with the test set.

Detection Models	Detection Boxes Precision		
	mAP	mAP@.50IOU	mAP@.75IOU
Faster RCNN ResNet-50	0.227	0.486	0.197
Faster RCNN ResNet-101	0.251	0.515	0.253
Faster RCNN Inception-V2	0.171	0.409	0.128
Faster RCNN Inception_ResNet-V2	**0.289**	**0.555**	**0.297**
	Detection Boxes Recall		
	AR@1	AR@10	AR@100
Faster RCNN ResNet-50	0.116	0.301	0.377
Faster RCNN ResNet-101	0.126	0.317	0.406
Faster RCNN Inception-V2	0.091	0.174	0.292
Faster RCNN Inception_ResNet-V2	**0.148**	**0.366**	**0.433**

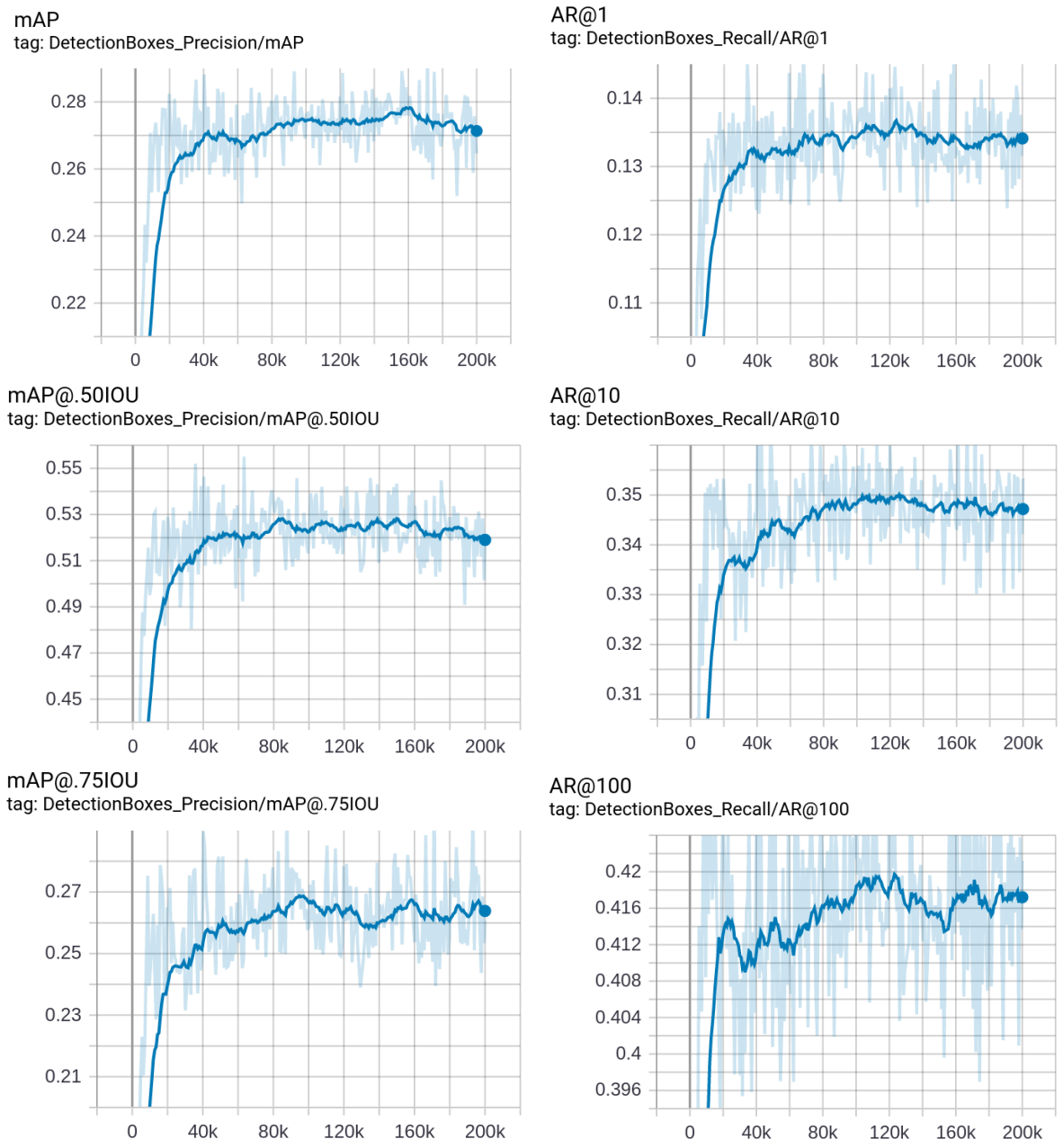


Figure 7. Performance plots of the Faster RCNN Inception-ResNet-V2 consist of Mean Average Precision (mAP) with different IoU threshold values (0.5 and 0.75) and Average Recall (AR) with 1, 10 and 100 detections.

Figure 8 illustrates the experimental results in the test set, showing typical results when the Faster RCNN Inception-ResNet-V2 model was used to deal with plant images under different weather conditions (windy, cloudy, and sunny). Figure 8 (a) presents the ground-truth bounding boxes and predicted bounding boxes of barley crops and weeds in a good weather condition (without windy and shadow). In addition, capeweed in this scenario was not detected in the image with predicted bounding boxes due to its tiny size and the overlapping of barley leaves. Figure 8 (b) illustrates the barley crops

and weeds in a windy day. Although the image was blurry, most of wild radish weeds at different growth stages and under occlusion conditions were detected. In the case where wild radish and capeweed appeared in the image of the barley crop, as shown in Figure 8 (c), despite the similar morphology of these weeds, the Faster RCNN Inception-ResNet-V2 model was able to precisely detect two types of weeds at different sizes. The next scenario illustrated in Figure 8 (d) where varying light conditions in the agricultural field environments are inevitable and common. Due to light reflection and shadows, it was difficult to detect wild radish under significant lighting variation and heavy occlusion. However, the chosen model detected part of wild radish weeds in the picture. Moreover, the inference time was around 0.38 seconds per image. Hence, the Faster RCNN model with Inception-ResNet-V2 is a promising candidate, with some optimization, for real-time applications.

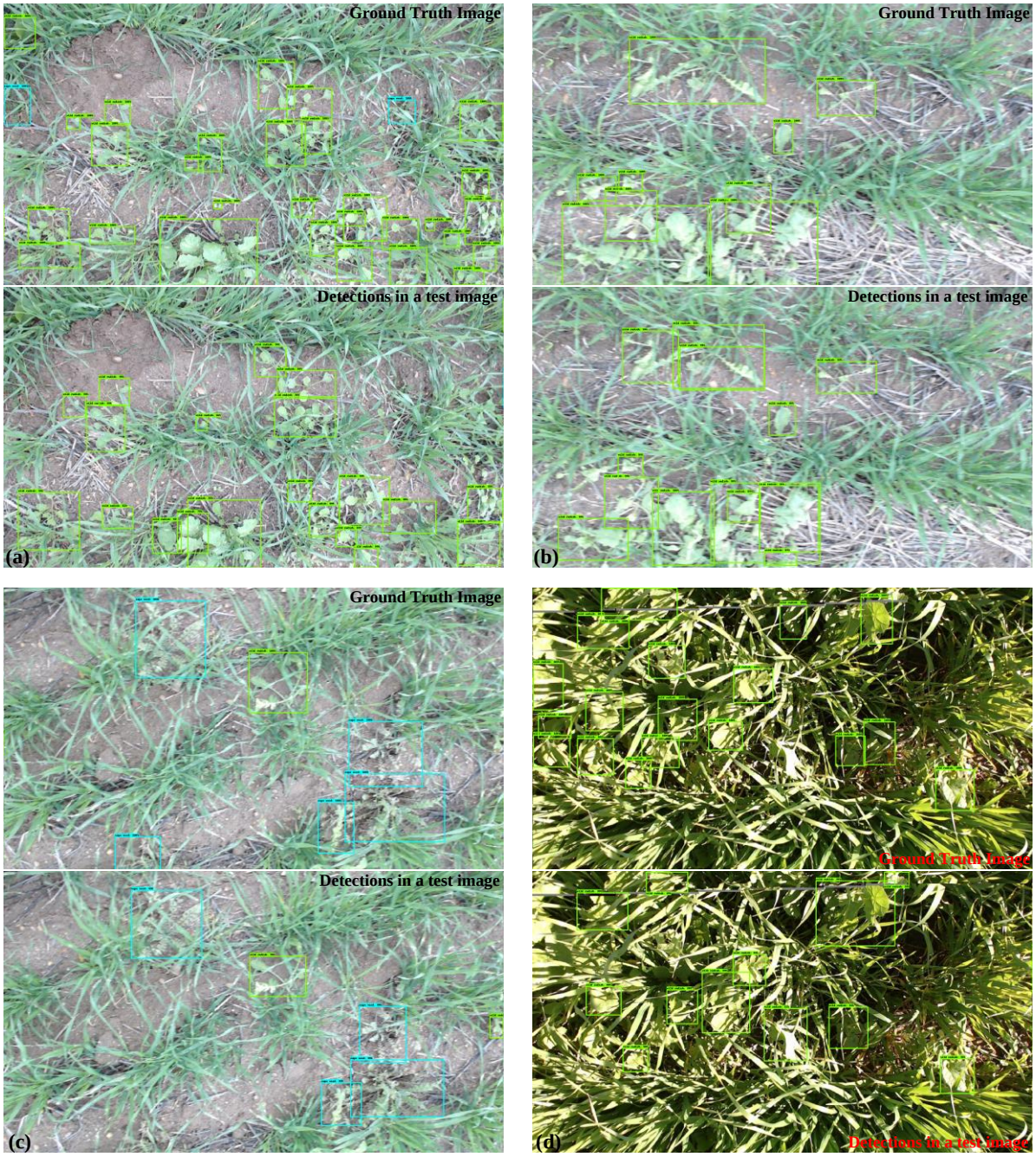


Figure 8. Ground True Images and detections in test images after applying the Fast-RCNN Inception-ResNet-v2 model (a) Weed detections under a good weather condition, (b) Weed detections in a windy day, (c) Detecting wild radish and cape weeds with similar morphology, and (d) Detecting wild radish under a shadow condition.

5.6 Conclusions

We have collected and labelled the FT_BRC dataset of the barley crop and weeds at the commercial farm at Cunderdin, Western Australia. Results have demonstrated the feasibility of using Faster RCNN

methods, especially the Faster RCNN model with Inception-ResNet-V2, for weed detection under complex field environments such as weather conditions, illumination variations, occlusion and overlap and different growth stages. Our chosen pipeline has produced potential and promising results with mAP = 0.555 at IoU =0.5 and mAP = 0.297 at IoU =0.75, which are relatively similar to the results of the MS-COCO dataset. In addition, the inference time was around 0.38 seconds per image enabling the detection weed in real-time.

The performance of this method for accurately detecting weeds in the field can be further improved by increasing the number of images with bounding box annotations in the FT_BRC dataset. Through further research the weed detection accuracy can be increased if a larger dataset is generated. Annotating more bounding boxes in all images of the FT_BRC dataset and collecting more images in different fields and regions to build a standard dataset for agricultural applications can achieve higher accuracies in detecting weeds with similar appearance in the agricultural field. Furthermore, the Faster RCNN method can be combined with other robust techniques to improve or enhance the detection of weed species under practical field conditions.

5.7 References

[1] A. Tellaeche, X. P. BurgosArtizzu, G. Pajares, A. Ribeiro, and C. Fernández-Quintanilla, "A new vision-based approach to differential spraying in precision agriculture," *computers and electronics in agriculture*, vol. 60, no. 2, pp. 144-155, 2008.

[2] G. P. Madafiglio, "Population management of Raphanus raphanistrum L.(wild radish) by regulating seed production," 2002.

[3] R. E. Blackshaw, D. Lemerle, R. Mailer, and K. R. Young, "Influence of wild radish on yield and quality of canola," *Weed Science*, vol. 50, no. 3, pp. 344-349, 2002.

[4] A. H. Cheam, "Proceedings of the Wild Radish and Other Cruciferous Weeds Symposium," in *Wild Radish and Other Cruciferous Weeds Symposium (2006: Dept. of Agriculture and Food, Western Australia)*, 2006: Dept. of Agriculture and Food.

[5] G. Arnold, P. Ozanne, K. Galbraith, and F. Dandridge, "The capeweed content of pastures in south-west Western Australia," *Australian Journal of Experimental Agriculture*, vol. 25, no. 1, pp. 117-123, 1985.

[6] H. Wood, "The introduction and spread of capeweed, Arctotheca calendula (L.) Levyns (Asteraceae) in Australia," *Plant Protection Quarterly*, 1994.

[7] A. H. a. J. Moore, "Profiles of common weeds of cropping-WEED 14: CAPEWEED (Arctotheca calendula)," Grains Research and Development Corporation, 2018. [Online]. Available: https://irec.org.au/wp-content/uploads/2018-GRDC-Common-Cropping-weeds.pdf

[8] (2016). *Western Australian wheat industry*. [Online] Available: https://www.agric.wa.gov.au/grains-research-development/western-australian-wheat-industry

[9] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Computers and electronics in agriculture*, vol. 147, pp. 70-90, 2018.

[10] Z.-Q. Zhao, P. Zheng, S.-t. Xu, and X. Wu, "Object detection with deep learning: A review," *IEEE transactions on neural networks and learning systems*, vol. 30, no. 11, pp. 3212-3232, 2019.

[11] P. Druzhkov and V. Kustikova, "A survey of deep learning methods and software tools for image classification and object detection," *Pattern Recognition and Image Analysis*, vol. 26, no. 1, pp. 9-15, 2016.

[12] K. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, "Machine learning in agriculture: A review," *Sensors*, vol. 18, no. 8, p. 2674, 2018.

[13] R. Schapire, "Machine learning algorithms for classification," *Princeton University*, vol. 10, 2015.

[14] A. Wang, W. Zhang, and X. Wei, "A review on weed detection using ground-based machine vision and image processing techniques," *Computers and electronics in agriculture*, vol. 158, pp. 226-240, 2019.

[15] P. Lottes, M. Hörferlin, S. Sander, and C. Stachniss, "Effective vision-based classification for separating sugar beets and weeds for precision farming," *Journal of Field Robotics*, vol. 34, no. 6, pp. 1160-1178, 2017.

[16] A. dos Santos Ferreira, D. M. Freitas, G. G. da Silva, H. Pistori, and M. T. Folhes, "Weed detection in soybean crops using ConvNets," *Computers and Electronics in Agriculture*, vol. 143, pp. 314-324, 2017.

[17] S. Skovsen *et al.*, "Estimation of the botanical composition of clover-grass leys from RGB images using data simulation and fully convolutional neural networks," *Sensors*, vol. 17, no. 12, p. 2930, 2017.

[18] M. Dyrmann, S. Skovsen, M. S. Laursen, and R. N. Jørgensen, "Using a fully convolutional neural network for detecting locations of weeds in images from cereal fields," in *The 14th International Conference on Precision Agriculture.*, 2018: International Society of Precision Agriculture Montreal, QC, pp. 1-7.

[19] A. Dutta, J. M. Gitahi, P. Ghimire, and R. Mink, "Weed detection in close-range imagery of agricultural fields using neural networks," *Publikationen der DGPF*, vol. 27, pp. 633-645, 2018.

[20] L. Quan *et al.*, "Maize seedling detection under different growth stages and complex field environments based on an improved Faster R–CNN," *Biosystems Engineering*, vol. 184, pp. 1-23, 2019.

[21] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *arXiv preprint arXiv:1804.02767*, 2018.

[22] V. Partel, J. Kim, L. Costa, P. Pardalos, and Y. Ampatzidis, "Smart Sprayer for Precision Weed Control Using Artificial Intelligence: Comparison of Deep Learning Frameworks," 2019.

[23] Y. Jiang, C. Li, A. H. Paterson, and J. S. Robertson, "DeepSeedling: deep convolutional network and Kalman filter for plant seedling detection and counting in the field," *Plant methods*, vol. 15, no. 1, p. 141, 2019.

[24] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 580-587.

[25] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," presented at the European conference on computer vision 2014.

[26] R. Girshick, "Fast r-cnn," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1440-1448.

[27] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in *Advances in neural information processing systems*, 2015, pp. 91-99.

[28] Z. Zou, Z. Shi, Y. Guo, and J. Ye, "Object detection in 20 years: A survey," *arXiv preprint arXiv:1905.05055*, 2019.

[29] V. N. T. Le, B. Apopei, and K. Alameh, "Effective plant discrimination based on the combination of local binary pattern operators and multiclass support vector machine methods," *Information processing in agriculture*, vol. 6, no. 1, pp. 116-131, 2019.

[30] V. N. T. Le, S. Ahderom, B. Apopei, and K. Alameh, "A novel method for detecting morphologically similar crops and weeds based on the combination of contour masks and filtered Local Binary Pattern operators," *GigaScience*, vol. 9, no. 3, 2020, doi: 10.1093/gigascience/giaa017.

[31] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, inception-resnet and the impact of residual connections on learning," in *Thirty-first AAAI conference on artificial intelligence*, 2017.

[32] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1-9.

[33] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770-778.

[34] E. Granger, M. Kiran, and L.-A. Blais-Morin, "A comparison of CNN-based face and head detectors for real-time video surveillance applications," in *2017 Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA)*, 2017: IEEE, pp. 1-7.

[35] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Information Processing & Management*, vol. 45, no. 4, pp. 427-437, 2009.

[36] A. Tao, J. Barker, and S. Sarathy, "Detectnet: Deep neural network for object detection in digits," *Parallel Forall*, vol. 4, 2016.

[37] D. Hoiem, Y. Chodpathumwan, and Q. Dai, "Diagnosing error in object detectors," in *European conference on computer vision*, 2012: Springer, pp. 340-353.

[38] T.-Y. Lin *et al.*, "Microsoft coco: Common objects in context," in *European conference on computer vision*, 2014: Springer, pp. 740-755.

Chapter 6 – Conclusions and Future Work

6.1 Contributions

This thesis has presented a novel method for real-time discrimination and detection of weeds and crops with similar morphologies under simulated and practical field environments. Taking advantage of the computational efficiency and invariance of illumination and rotation features of the conventional LBP method, we have investigated and developed a novel approach for achieving high accuracy in the classification of crops and weeds with a similar appearance, and solving the existing agricultural challenges such as occlusion, overlap of plant leaves, the alternation of broadleaf crops and weeds at different growth stages, and lighting conditions under variable weather conditions.

Due to the limited number of datasets in agriculture, our research has provided useful datasets collected in the laboratory, by a Testbed system, and in the field, using an integrated weed sensing system. We selected typical species represented for crops (canola, barley, and corn) and weeds (wild radish and capeweed). In the laboratory, crops and weeds were captured at different growth stages under simulated field environments. In the field, images of plant species were captured in real and complicated environments at a commercial farm at Cunderdin, Western Australia. These datasets were used, throughout this thesis, to evaluate and validate our advanced algorithms.

To enhance the performance of the original LBP method in weed management, we combined three different LBP operators with SVM to extract dominant features and classify broadleaf canola and wild radish, corn and background in the "bccr-segset" dataset at four growth stages. However, the probability of identifying visually similar crops and weeds was still limited. Continuously, we investigated and developed the advanced LBP-based algorithm by eliminating insignificant features in three LBP operators, then combining dominant features extracted by three combined LBP operators with LBP features extracted by plant-leaf contour masks and a coefficient k. Based on the experimental results conducted in laboratory and field environments, the improvement of our novel algorithm is demonstrated by enhancing the average classification accuracy of crops and weeds collected in the laboratory from 91.85% up to 98.63%, and achieving the average classification accuracy of 90.94% compared to the well-known CNN models, VGG-19, ResNet-50 and Inception-V3. Compared with the DL methods, and using the "Can-rad" and "Mixed-plants" datasets, the k-FLBPCM method has exhibited robustness to plant size variations. Despite the training dataset at the early or late growth stages, results have shown that the k-FLBPCM method has the ability to effectively detect crops and weeds at different plant growth stages as opposed to CNN models. Another advantage of the k-FLBPCM algorithm is that it requires shorter training and testing times than CNN models. The k-FLBPCM algorithm represents a step forward in achieving real-time accurate weed detection, and this

overcomes the excessive application of herbicides issues of blanket spraying practices by applying herbicides only when weeds are detected.

6.2 Future work

Although our proposed methods addressed several important challenges, namely, detecting crops/weeds at different growth stages, visual similarities of crops/weeds, viewpoint/pose variations, vegetation illumination variations, occlusion and overlapping of crops/weeds, further investigations and developments can be conducted in the near future. This thesis focused on the common agricultural problems and data collection in Australia. Generally, plant datasets can be expanded to collect additional data, especially more productive crops and invasive weeds to meet agricultural demands in other areas. In addition, the ability to detect various weed species from cereal crops in the field can be increased by training k-FLBPCM and CNN models with more plant images. It is important to note that while our method was designed only for vegetation discrimination with similar appearance in agriculture, it can also be expanded to other visually similar objects in multiple scenarios and applications. The advanced method can classify and detect objects with scale, rotation and illumination variations, and similar morphologies in real time.

In this thesis, the pre-processing and segmentation steps have been carefully considered by applying different algorithms to remove disturbance factors and unexpected noise in our experiments. However, the vegetation segmentation of images in the field is still limited in representing the whole vegetation patterns. To further improve the effectiveness of weed detection, the patterns of leaves captured in the field should be well segmented before inputting in the feature extraction and the training process. With the optimization of steps in image processing, the performance of our algorithm can be further enhanced to precisely detect weeds and estimate the weed density in the field.

Although our integrated weed sensing system used in the experiments has both spectral reflectance (using Normalized Difference Vegetation Index - NDVI) and texture (spatial imaging) analyses, only image analyses were used in this PhD project. Future work can investigate the combination of spectral reflectance features and texture features to further increase the accuracy and speed of the weed detection system in the field.

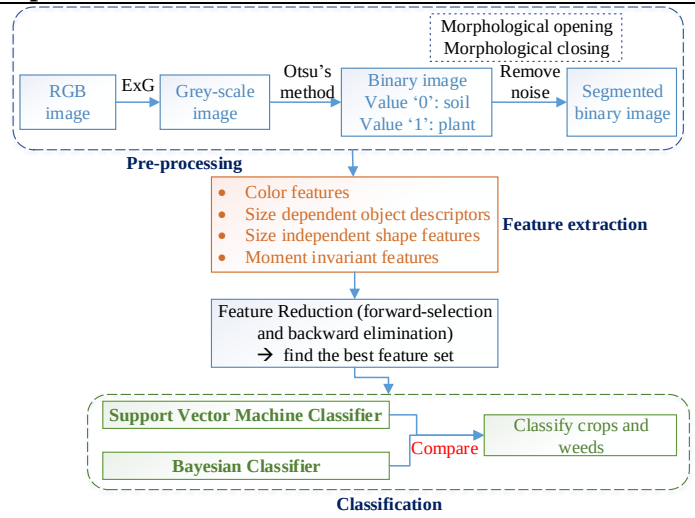
Collecting datasets in the field took a long time due to plant growth over the growing season. Moreover, only a part of plant images in the dataset collected from the field were fully annotated in Chapter 5 (mainly because of the time limitation). Therefore, the mAP results of the faster RCNN models with different convolutional neural network architectures were still limited. Although the results of weed detection under various weather conditions, heavy occlusion and overlap are promising in Chapter 5, the faster RCNN models can further be improved by combining features in the k-FLBPCM method with CNN features, or fully annotating bounding boxes for all images in the field dataset. This enables

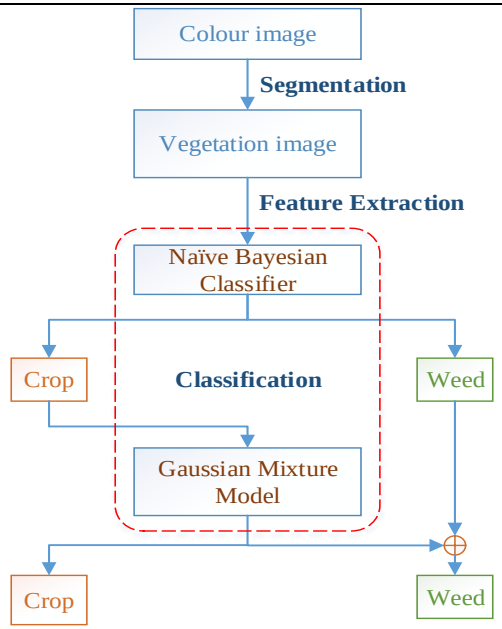
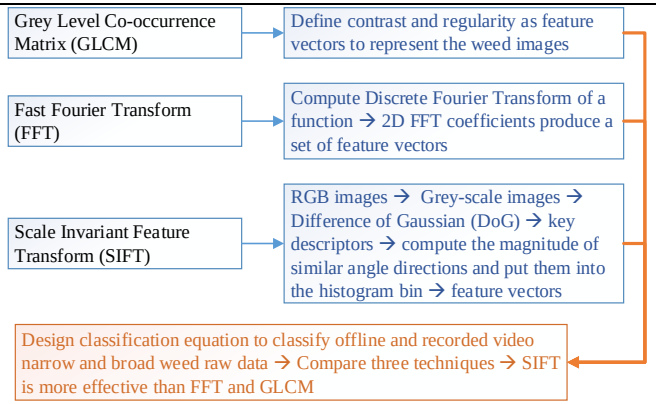
accurate weed detection under complex field environments in real time. This approach has enabled to spray targeted weeds more rapidly and precisely, hence resulting in significant herbicide savings in addition to economic and environment benefits.

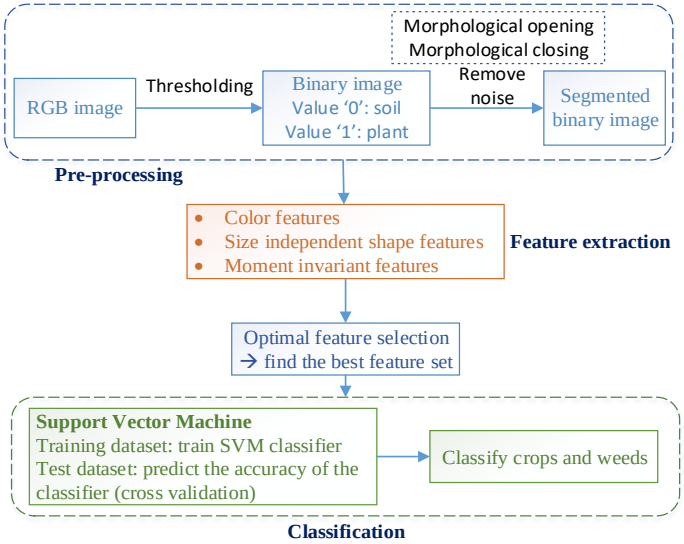
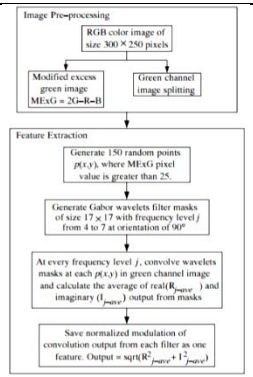
Appendix

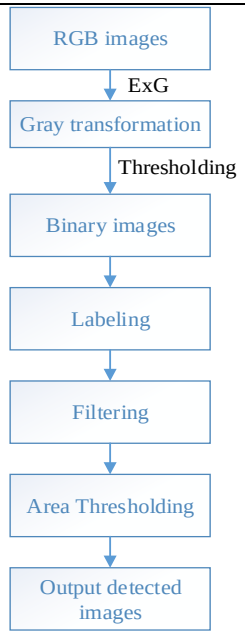
Appendix for Chapter 1

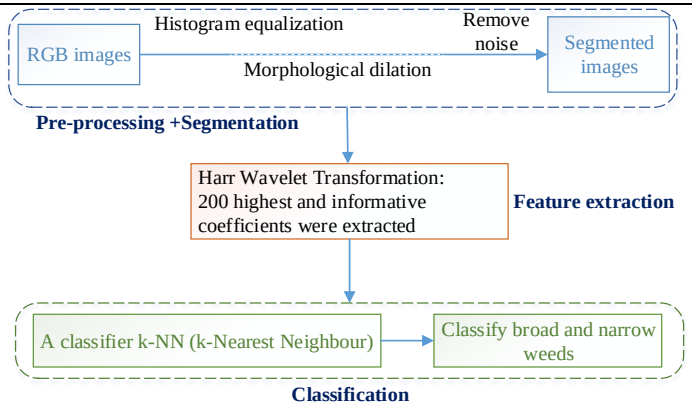
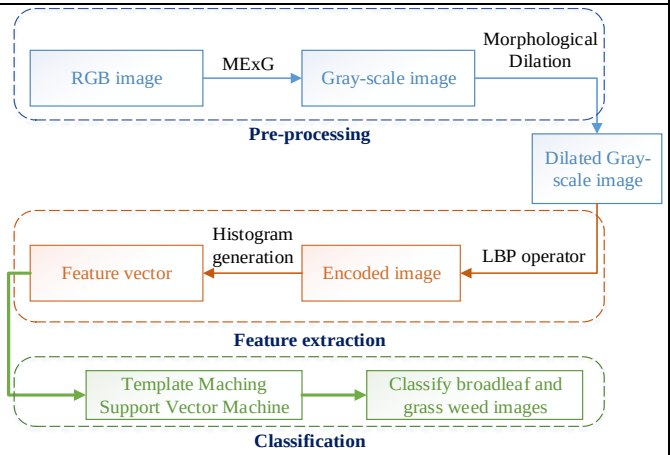
A summary of popular methods for weed detection and classification.

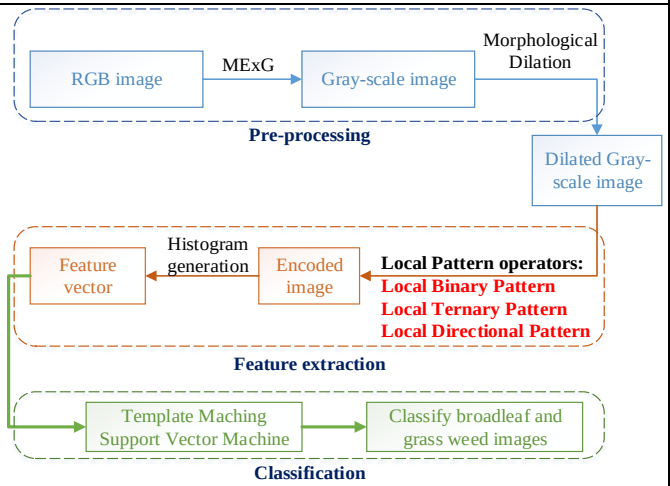
Methods	Dataset	Operation	Benefits	Limitations
"Performance analysis of support vector machine and Bayesian classifier for crop and weed classification from digital images" [1]	- Objects: chilli plants and 5 weed species - Camera: image resolution was 1200 × 768 pixels - Experiment: image resolution was decreased to 448 × 336 pixels - A set of 224 test images		- Reducing features → decreasing the computational complexity and remove noisy features - SVM achieves 98.22% accuracy over a set of 224 test images - Bayesian classifier achieves 95.79% accuracy over the same set of images.	- This study did not cover the mutual overlap of plants - Recognise only one single plant - This method was not mentioned to application in real time conditions

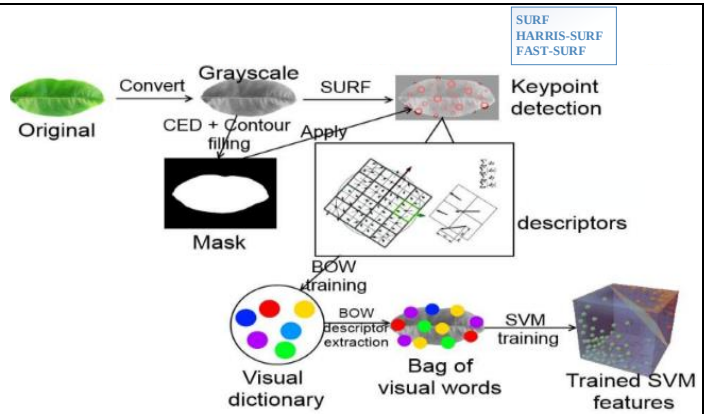
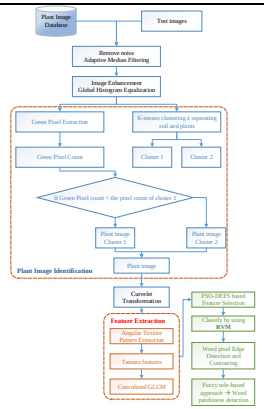
"Bayesian classification and unsupervised learning for isolating weed in row crops" [2]	- A set of 149 images taken from two different crops: corn and soybean plants at the 2-4 stage of corn and the 2-3 stage of soybean		- No need to have prior train and prior knowledge on crops and weeds - Good estimation of the weed coverage on field sections	- Limited in evaluating the performance on plant images in a section of a field
"Machine vision system for automatic weeding strategy using image processing technique"[3]	- Objects: Oil Palm, broad and narrow weeds - A recorded video during 30 minutes under real conditions - Testing 1000 sample of offline images		- Responds positively to real time condition in lab - Corrects classification rate: SIFT (99.5%) for recognizing narrow weeds and SIFT (99.8%) for detecting broad weeds	- System was limited in real time condition - Has not been developed for mix weed images

"Classification of crops and weeds from digital images: A support vector machine approach" [4]	- Objects: chilli crops and selected weed species in mature stage - Images were captured by a camera with the resolution 1200x768 pixels - 224 images with the decreased resolution 448x336 pixels		- Combines different types of features with SVM to classify crops and weeds → 97% accuracy over a set of 224 test images - Uses field images to evaluate the performance - Quick computation time → can apply in real-time systems	- Limited to a single plant without mutual overlapping with other plant leaves - Segmentation errors in the form of plant holes and noise backgrounds
"Classification of broadleaf and grass weeds using Gabor wavelets and an artificial neural network" [5]	- Objects: broadleaf weeds and grasses at a growth stage - 40 samples images with 20 samples from each class - Images were captured by a camera with the resolution 640x480 pixels		- This method classifies quite good.	- Feature extraction algorithm only applied unidirectional wavelet filters - Only one type of weed in an image - Complicated computation → long computation time → needs improvement to apply under real-time constraints

	- 300x250 pixel sub-images were cropped from the centre of images			
"Weed and crop segmentation and classification using area thresholding" [6]	- Objects: crop and some weeds - 41 sample images have been tested - Camera: images with 3648x2736 pixels In the experiment, images with 320x240 pixels → reducing computation time		- This method can classify crop and weed plants, although it still needs to address error rate improvement	- Does not apply in real-world scenario - High weed misclassification rate (~33.3%) - Does not solve the overlap of weed and crop plants

"Weed classification based on Haar Wavelet Transform via k-nearest neighbour (k-NN) for real-time automatic sprayer control system" [7]	- Objects: broad and narrow weeds - Database of 200 samples of each category - Image size in the experiment: 240x320 pixels		- Using Harr Wavelet Transform for discriminate specific weeds in real time with an average accuracy 94% and computation time of 40ms	- Misclassification was found in images - Environmental conditions (Lighting and wind) have not been solved, which affect the performance of algorithm
"A Study on Local Binary Pattern for Automated Weed Classification Using Template Matching and Support Vector Machine" [8]	- 200 colour images of broadleaf and grass weeds with 100 samples from each class - Image size in the experimental analysis: 320x240 pixels		- Converts rotation variant LBP to rotation invariant LBP - A computationally efficient approach using LBP operator to generate a feature vector with only a single scan of the image - Works in natural lighting conditions - LBP based feature representation is robust against monotonic illumination variation →accurate classification in unrestrained environment	- Has not been developed for mix weed images

"Automated Weed Classification with Local Pattern-Based Texture Descriptors" [9]	- Objects: broadleaf and grass weeds - 400 sample field images with 200 samples from each category - Camera: images with 1200x768 pixels - In the experiment, images with 320x240 pixels		- Potential for real-time operations - This algorithm is robust and adapts to practical conditions (illumination variation and noise) in the field. - Converting rotation variant local patterns to rotation invariant patterns - Compared to wavelet based texture analysis methods (Gabor and Harr wavelet transformation), Local Directional Pattern (LDP) operator is computed more efficiently	- Has not been investigated for mix weed images

"Efficient modern description methods by using SURF algorithm for recognition of plant species" [10]	- Using Flavia dataset		- With a lower descriptor dimension, faster computation, good ability to distinguish features → the accuracy of SURF was higher than other methods - Advantage of this BoW model is ability to increase classification accuracy	- Recognise single leaf only - Has not been trialled with a real dataset - SURF algorithm is not good at tackling viewpoint change and illumination change
"A novel wrapping curvelet transformation based angular texture pattern extraction method for weed identification " [11]	- Dataset 1: Brinjal field images → 500 images - Dataset 2: Carrot field images → 60 images - Training: using a set of 560 images - Testing: using a set of 500 images - However, evaluating the performance of this method was		- Addresses complex background, illumination variation, and the overlap of crop and weed in field images - Uses images of mixed crop/weed - The accuracy of this method is 98.3%, i.e., higher than SVM - based method	- Has not been trialled in real-time scenarios due to the computational inefficiency

	based on a set of 150 images			

References

[1] F. Ahmed, A. H. Bari, E. Hossain, H. A. Al-Mamun, and P. Kwan, "Performance analysis of support vector machine and Bayesian classifier for crop and weed classification from digital images," *World Applied Sciences Journal,* vol. 12, no. 4, pp. 432-440, 2011.

[2] F.-M. De Rainville *et al.*, "Bayesian classification and unsupervised learning for isolating weeds in row crops," *Pattern Analysis and Applications,* vol. 17, no. 2, pp. 401-414, 2014.

[3] K. H. Ghazali, M. M. Mustafa, and A. Hussain, "Machine vision system for automatic weeding strategy using image processing technique," *American-Eurasian Journal of Agricultural and Environmental Science,* vol. 3, no. 3, pp. 451-458, 2008.

[4] F. Ahmed, H. A. Al-Mamun, A. H. Bari, E. Hossain, and P. Kwan, "Classification of crops and weeds from digital images: A support vector machine approach," *Crop Protection,* vol. 40, pp. 98-104, 2012.

[5] L. Tang, L. Tian, and B. L. Steward, "Classification of broadleaf and grass weeds using Gabor wavelets and an artificial neural network," *Transactions of the ASAE,* vol. 46, no. 4, p. 1247, 2003.

[6] S. H. Hlaing and A. S. Khaing, "Weed and crop segmentation and classification using area thresholding," *IJRET,* vol. 3, pp. 375-382, 2014.

[7] I. Ahmad, M. H. Siddiqi, I. Fatima, S. Lee, and Y.-K. Lee, "Weed classification based on Haar wavelet transform via k-nearest neighbor (k-NN) for real-time automatic sprayer control system," in *Proceedings of the 5th International Conference on Ubiquitous Information Management and Communication*, 2011: ACM, p. 17.

[8] F. Ahmed, A. H. Bari, A. Shihavuddin, H. A. Al-Mamun, and P. Kwan, "A study on local binary pattern for automated weed classification using template matching and support vector machine," in *Computational Intelligence and Informatics (CINTI), 2011 IEEE 12th International Symposium on*, 2011: IEEE, pp. 329-334.

[9] F. Ahmed, M. H. Kabir, S. Bhuyan, H. Bari, and E. Hossain, "Automated weed classification with local pattern-based texture descriptors," *Int. Arab J. Inf. Technol.*, vol. 11, no. 1, pp. 87-94, 2014.

[10] M. F. Kazerouni, J. Schlemperz, and K.-D. Kuhnert, "Efficient Modern Description Methods by Using SURF Algorithm for Recognition of Plant Species," *Advances in Image and Video Processing,* vol. 3, no. 2, p. 10, 2015.

[11] D. A. Kumar and P. Prema, "A NOVEL WRAPPING CURVELET TRANSFORMATION BASED ANGULAR TEXTURE PATTERN (WCTATP) EXTRACTION METHOD FOR WEED IDENTIFICATION," *ICTACT Journal on Image & Video Processing,* vol. 6, no. 3, 2016.

Statement of Contribution

To Whom It May Concern,

I, Vi Nguyen Thanh Le, conceived, designed and undertook all experiments. Vi Nguyen Thanh Le analysed all data, interpreted results and authored the publications. Prof. Kamal Alameh, Dr. Selam Ahderom interpreted results and edited the publications entitled:

- V. N. T. Le, B. Apopei, and K. Alameh, "Effective plant discrimination based on the combination of local binary pattern operators and multiclass support vector machine methods," *Information processing in agriculture*, vol. 6, no. 1, pp. 116-131, 2019.
- V. N. T. Le, S. Ahderom, B. Apopei, and K. Alameh, "A novel method for detecting morphologically similar crops and weeds based on the combination of contour masks and filtered Local Binary Pattern operators," *GigaScience*, vol. 9, no. 3, 2020, doi: 10.1093/gigascience/giaa017.
- V. N. T. Le, S. Ahderom, and K. Alameh, "Performances of the LBP Based Algorithm over CNN Models for Detecting Crops and Weeds with Similar Morphologies," *Sensors*, vol. 20, no. 8, p. 2193, 2020.
- V. N. T. Le, S. Ahderom, and K. Alameh. (2020). Detecting weeds from crops under complex field environments based on Faster RCNN models. Precision Agriculture Journal (Under revision).

Vi Nguyen Thanh Le

I, as a Co-Author, endorse that this level of contribution by the Candidate indicated above is appropriate.

Prof. Kamal Alameh

Dr. Selam Ahderom

Electron Science Research Institute, Edith Cowan University, 270 Joondalup Drive, WA 6027, Australia