

# Broad-leaf weed detection in pasture

Wenhao Zhang, Mark F. Hansen, Timothy N. Volonakis, Melvyn Smith, Lyndon Smith

Centre for Machine Vision,  
Bristol Robotics Laboratory  
University of the West of England, Bristol  
Bristol, UK  
wenhao.zhang@brl.ac.uk

Jim Wilson, Graham Ralston  
SoilEssentials  
Angus, UK  
jim@soilessentials.com

Laurence Broadbent, Glynn Wright  
Aralia Systems Ltd  
Horsham, UK  
laurence.broadbent@aralia.co.uk

**Abstract**—Weed control in pasture is a challenging problem that can be expensive and environmentally unfriendly. This paper proposes a novel method for recognition of broad-leaf weeds in pasture such that precision weed control can be achieved with reduced herbicide use. Both conventional machine learning algorithms and deep learning methods have been explored and compared to achieve high detection accuracy and robustness in real-world environments. In-pasture grass/weed image data have been captured for classifier training and algorithm validation. The proposed deep learning method has achieved 96.88% accuracy and is capable of detecting weeds in different pastures under various representative outdoor lighting conditions.

**Keywords**—weed detection; machine learning; transfer learning; support vector machine

## I. INTRODUCTION

Pasture is increasingly seen as crop, which when managed effectively can provide a healthy diet for livestock throughout the year. The presence of weeds lessens the nutrition within the crop, and if left unchecked will out-compete the grass. One of the most common and invasive weeds is the broad-leaf weed known as Dockleaf (*Rumex Obtusifolius*). Current methods of controlling weeds tend to rely on blanket spraying of the field using herbicides such as glyphosphate. This is both costly for farmers and can have a negative impact on the environment, e.g. by contaminating nearby watercourses. Precision spraying has been developed with some success on "in-crop" weeds where the weed is more readily differentiated from the crop, either due to the high colour contrast between vegetation and soil, or by exploiting the regular geometry of the crop rows themselves. However, "in-pasture" weed detection is considerably harder as the weeds and grass are both predominantly green, and the grass can often obscure the weed. [1] provided an excellent comparison of machine vision techniques for detecting Dockleaf in pasture, concluding that the best performing was an approach that used Local Binary Patterns (LBP) which describe local textures as a feature extractor and a Support Vector Machine (SVM) as the classifier. This approach achieved an accuracy of just over 80%. In this paper, we replicate this finding on a large dataset and compare this method with other conventional machine learning methods, as well as providing improved performance through the use of deep learning.

## II. LITERATURE REVIEW

Weeds in pastures pose a threat to livestock health and their control using broadcast spraying can have a detrimental impact on the natural environment. Some weed species can result in animal illness, even death, particularly in equine pastures. While weed control in arable crops is a common practice, in grassland it is far less usual. A large amount of herbicide input is not only unfriendly to environment, but incurs great economic costs. This negative impact might be dramatically reduced if the weed control could be more targeted by application in a way such that the amount of herbicide usage is reduced to a minimum and its application more targeted by spot spraying only where weeds exist. In view of this, a few research works have investigated weed detection/recognition in grasslands by using conventional machine learning methods [2, 3, 4].

Machine learning has been providing effective solutions to a wide range of object detection tasks. Generally, desirable features can be readily extracted from indoor and structured environments. However, when it comes to outdoor environments where there exist variations that cannot be modelled explicitly, conventional machine learning methods can struggle to maintain good performance. This calls upon investigations and comparisons of different methods that can obtain reliable features (hand crafted or automatically learned) regardless of the complexity of data acquired.

Neural networks have been an active area of research for many decades due to their theoretical ability to model any relationship between input and output, linear or non-linear, provided with sufficient data from which to generalise. As far back as 1980, Fukushima [5] proposed an architecture based on the human visual receptive fields that he named the Neocognitron. It described alternating layers that convolved and sub-sampled an input image. This architecture inspired the development of LeNet-5 [6], a 7-layer convolutional neural network (CNN) which recognised hand written digits. However, due to the large numbers of trainable parameters needed, it could only operate on 32x32px images. With the growth in low cost, powerful graphics cards, it is now possible to design and train far deeper networks on desktop machines. For example an NVidia Corporation Titan-X card (similar to that used in this study) contains 3072 computing cores and 12GB of on-board RAM. In recent years deep learning and

convolutional networks have been applied in a wide range of both research areas and also in industrial settings [7, 8]. To our knowledge this is the first time a deep network has been used to classify docks in pasture.

### III. METHODOLOGY

While object detection is a well-researched area, challenges posed by broad-leaf weed (referred to as weed in the rest of the paper) detection is different in that a grassland represents a type of dynamic and highly unstructured outdoor environment. More complications are caused by similarity in colour between grass crops and weeds. However, regardless of the similarity in colour, grass crops can commonly be characterised by their thin overlapping leaves as opposed to broad-leaf weeds that have relatively large homogeneous leaf surfaces. Therefore, texture information may be informative at discriminating between weeds and grass.

#### A. Hardware of the vision system

To ensure that abundant and good quality data can be captured, we designed a computer vision system that could run weed detection independently from other hardware components (e.g. a nozzle control system for spot spraying herbicides and a GPS module for recording weed locations), but at the same time, could conveniently communicate with the other hardware modules.

The computer vision system consists of a standard machine vision camera, an appropriate camera lens (providing a field of view of approximately 1m by 1m) and an embedded computing device. It can serve as a data capture system as well as a weed detection system. The camera lens was chosen in a way such that it provided a field-of-view that matched the effective spraying area covered by two spray nozzles. The diagram in Figure 1 shows how the vision system fits in the overall hardware platform.

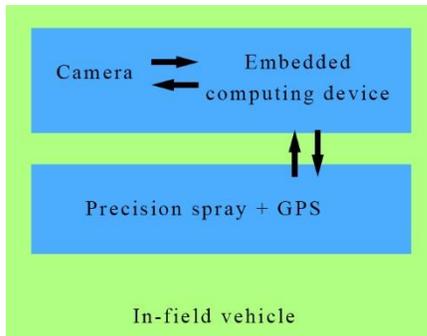


Figure 1. A diagram illustrating the hardware platform

This vision system can be mounted to common in-field vehicles performing weed control by spraying herbicides. For data capture, we mounted this vision system on a quadbike.

#### B. Data acquisition

Pasture fields in Dundee, Scotland, UK were traversed with a quadbike, which travelled in column-wise transects of the fields. The camera was set to continuously capture images at a high frame rate and high resolution (2048\*2048 pixels),

where redundant images containing repeated regions of a field were removed. Each remaining image was then divided into two columns and three rows, resulting in image sizes of 1024 × 682 pixels. Each spray nozzle will therefore cover one of the two columns, i.e. three image patches per image frame.

A MATLAB script was implemented so that individuals could rapidly label the image data using a keyboard arrow key press. This allowed the data to be efficiently and manually labelled in three categories - grass, weed and unsure. Data labelled as ‘unsure’ were discarded from the experiments as ground truth for them was not reliable. All the weed data were collated and masks for the weed regions were then manually labelled as well. Apparently, weeds can be almost infinitely small when they just start to grow, and it is unrealistic to detect them at these sizes. Therefore it was decided to only attempt to detect weeds that have a coverage larger than 5% of the entire image patch.

A representative example of a labelled weed image patch is shown in Figure 2.

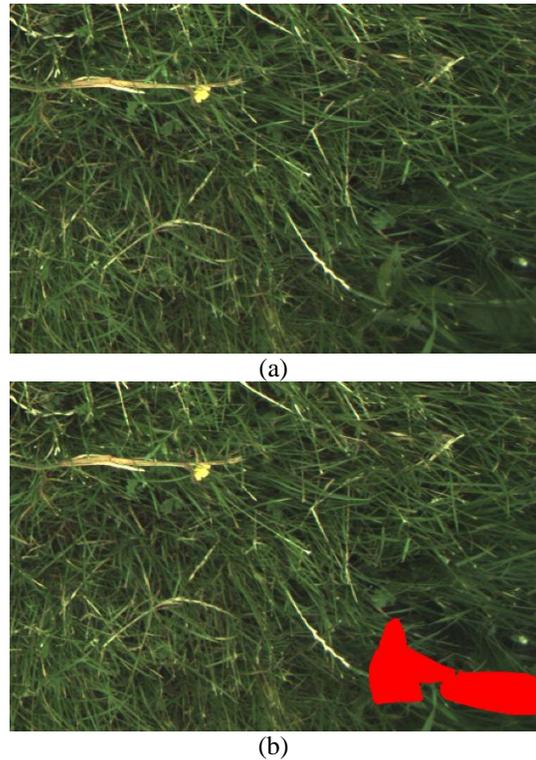


Figure 2. An example of a labelled weed image where the red region is the mask for the weeds. (a) the original image. (b) the image superposed with the mask.

Overall, 6087 image patches were captured and labelled, including 4080 grass image patches and 2007 weed image patches.

#### C. Conventional machine learning

Conventionally, machine learning demands hand crafted features and a model that can perform either regression or

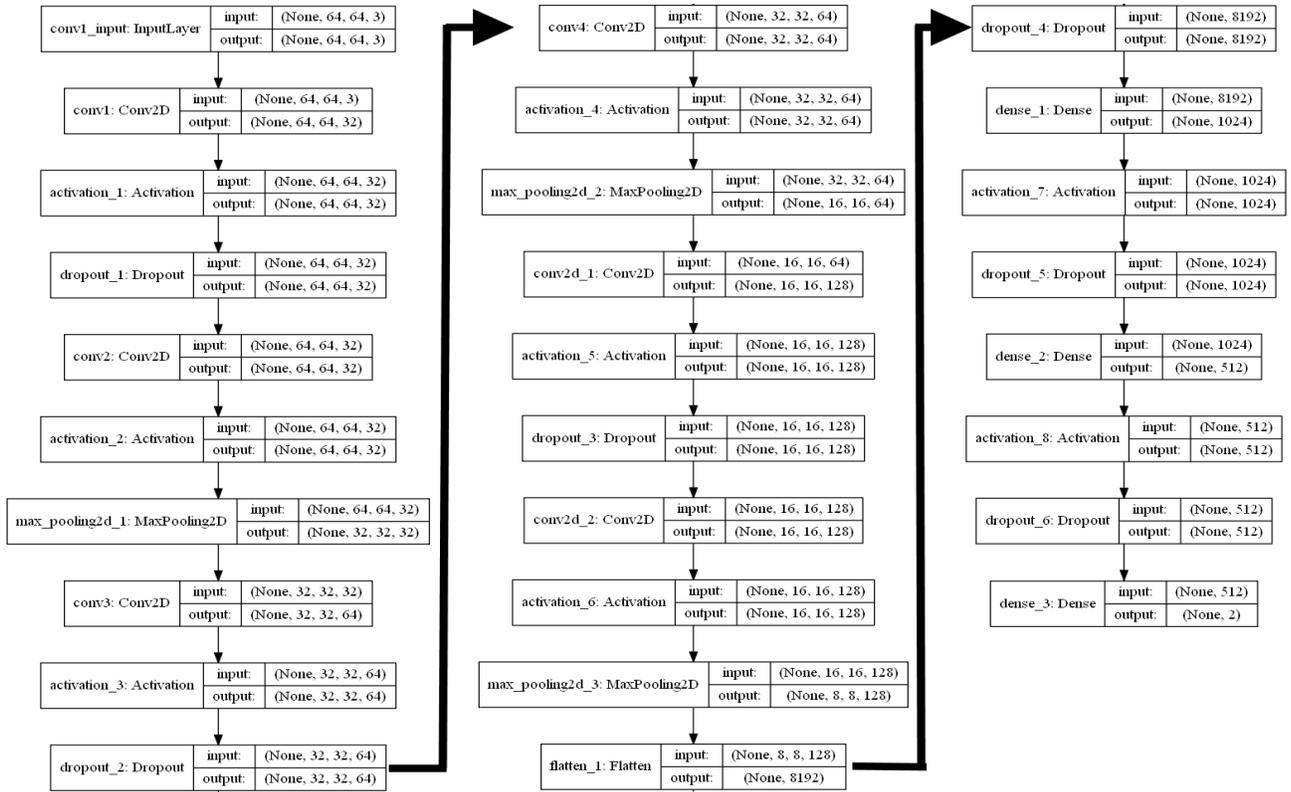


Figure 3. Architecture of our CNN consisting of six convolutional layers with alternating dropout and max-pooling layers. Classification occurs in the final three fully connected layers.

classification tasks such that image samples can be characterised and discriminated.

Local binary patterns [9] have been found to be excellent descriptors of local textures and have been widely used for various applications. Extracting uniform local binary pattern histograms (LBPH) can compute meaningful feature descriptors that are rotation invariant whilst retaining low dimensionality. Both properties are advantageous for real-world and real-time applications. Since we need to capture fine grained texture information, a neighbourhood of 8 and a radius of 1 were chosen for obtaining LBPH descriptors.  $L2$  normalisation was implemented at a post-processing stage. To investigate the impact of utilising colour information, we extracted LBPH features from both greyscale images and colour images (concatenated features from the Red, Green and Blue channels), respectively. A comparison of these results can be found in Section IV.

Weed detection may be considered a two-class problem, i.e. weed and non-weed. This can be effectively solved by a support vector machine (SVM) [10] which was originally formulated as a binary classification method. As we captured a large amount of labelled image data, this supervised method should be able find a linear or a non-linear hyperplane that can separate the two classes well. Other classification and regression method have also been tested, including K-nearest

neighbour (KNN), Ensemble methods, Complex Tree and Logistic Regression.

These methods have been implemented in MATLAB [11] for evaluation. The results in Section IV report the performance of these methods.

#### D. A deep learning based method

CNN approaches in particular, combine both the feature extraction and classification stages of classical pattern recognition, by propagating the training of a fully connected classifier back through the convolutional layers in order to select the best features.

To train a network of this depth not only takes a great deal of computing power, but also a very large amount of training data. In order to boost the amount of data we perform small manipulations to increase the variance and also improve the robustness of the network (shifting the images by up to 64 pixels in the x/y plane, flipping the images horizontally and rotating by up to 30 degrees). The colour images are  $1024 \times 682$  pixels in size and these are resized to  $64 \times 64$  pixels (this was determined empirically to give an acceptable compromise of good performance vs speed of processing), before being fed into the first layer of the network. The target for each image is either Weed, if the image contains weed, or 'Not Weed' if it does not. These are converted to a 'binary' representation of output nodes, where a '1' in the relevant index corresponds to

that class e.g. '10' corresponds to Weed, and '01' corresponds to 'Not Weed'. The network we train consists of six convolutional layers, with alternating dropout and max-pooling layers in between. The purpose of these is to aid convergence by preventing local minima, and providing scale/location invariance. The classification layer consists of three fully connected layers with the final layer containing 2 outputs as described above. The architecture is shown in Figure 3.

The convolutional neural network was written using available libraries in Python3.5 (Keras2.0, scikit-learn0.18.2, tensorow-gpu-1.1.0). All code has been run on a desktop computer with an Intel i5 processor, 16GB RAM using Windows7 64-bit, and an NVidia Titan-X 12GB graphics card with Maxwell Architecture. Training our network over 5000 epochs takes approximately 20 minutes. The results of the model from the best performing epoch on the test data is then used. The results in Section IV report performance on the test partition.

#### IV. EXPERIMENT AND RESULTS

As a single spray nozzle has a coverage of an entire image patch, it is unnecessary to detect the exact coordinates of the weeds within the patch, but instead each image patch needs to be classified as either containing weeds or not containing weeds. Therefore the experiments were carried out for a two-class classification problem.

For all experiments, the data were split 80:20 into training and testing partitions. Table I shows the partitioning of these data.

TABLE I. A DATA PARTITION RATIO OF 80:20 WAS USED TO SPLIT THE TRAINING AND TEST DATASETS FOR TRAINING OUR OWN CNN. RESULTS ARE REPORTED ON THE TEST PARTITION

	Number of training images	Number of testing images	Total
Weed	1591	416	2007
Not Weed	3278	802	4080
Total	4869	1218	6087

Various other state-of-the-art methods were also tested, allowing a comparison to be made amongst several representative classification and regression methods for weed detection. The results are summarised in Figure 4.

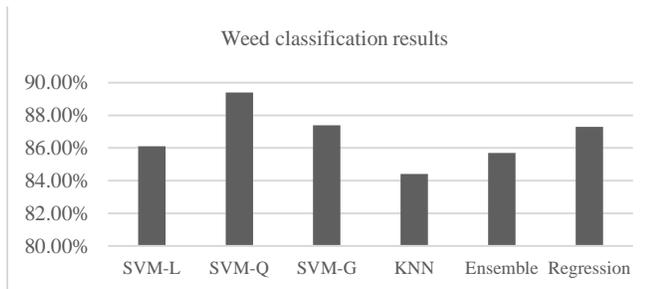


Figure 4. Weed detection accuracy of a linear SVM (SVM-L), quadratic SVM (SVM-Q), medium Gaussian SVM (SVM-G), weighted KNN (KNN), Ensemble subspace discriminant (Ensemble) and logistic regression (regression) [11]

As shown by the results, a quadratic SVM produced the highest accuracy (i.e. 89.40%) amongst a number of conventional machine learning methods.

The confusion matrix for the quadratic SVM can be seen in Figure 5.

True label	Predicted label	
	weed	not weed
weed	0.80	0.20
not weed	0.06	0.94

Figure 5. The normalised confusion matrix of the quadratic SVM. This method correctly identifies weeds with 80% accuracy.

It should be noted that these results are based on LBPH features extracted from RGB colour images. When only greyscale information was used, the performance dropped to 87.7% from 89.4%.

Similarly, we evaluated weed detection accuracy of our convolutional neural network. The results can be found in Table II and Figure 6.

TABLE II. BREAKDOWN OF RESULTS WHERE ACCURACY = NUMBER CORRECT / TOTAL NUMBER OF THAT CLASS, FALSE POSITIVES = SUM OF INCORRECT LABELS FOR A PREDICTED CLASS / TOTAL NUMBER FOR THAT CLASS, AND FALSE NEGATIVE = SUM OF INCORRECT LABELS FOR A TRUE CLASS / TOTAL NUMBER FOR THAT CLASS.

	Accuracy (%)	False Positive Rate (%)	False Negative Rate (%)
Weed	93.15	1.8	6.8
Not Weed	98.97	3.8	1

True label	Predicted label	
	weed	not weed
weed	0.93	0.07
not weed	0.01	0.99

Figure 6. The normalised confusion matrix of our CNN. This method correctly identifies weeds with 93% accuracy.

It can be seen that although LBPH and SVM gave desirable performance given the scene complexity, the CNN achieved an accuracy that is 13.00% higher for weed detection at an overall accuracy of 96.88%. A few visual examples of weed/grass classification results are shown in Figure 7.

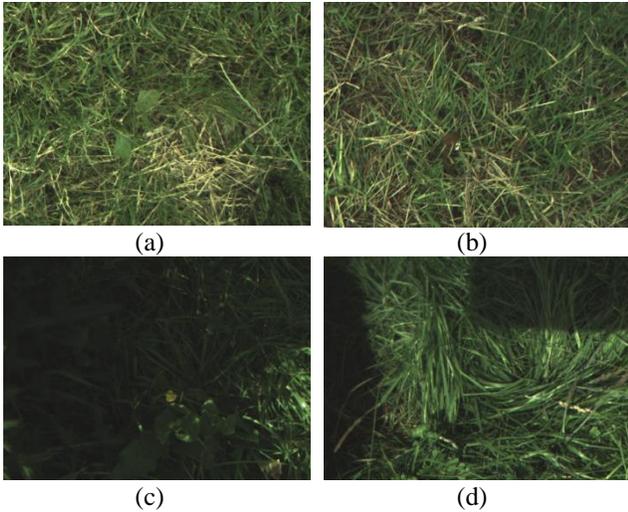


Figure 7. Visual examples of classification results. (a) successful identification of weed, (b) successful identification of grass, (c) weed identified as grass, and (d) grass identified as weed.

### CONCLUSION

Broad-leaf weed detection in pasture is a challenging problem which is faced as it involves uncontrollable lighting as well as a highly complex and unstructured scenes. This paper investigates different methods for achieving high detection accuracy and robustness in a real-world application. A number of methods have been proposed and implemented. Results show that conventional machine learning methods can achieve a desirable high accuracy by exploring image texture information, where local binary patterns and a quadratic support vector machine achieved the highest accuracy of 89.4%. A deep learning method utilising a CNN further increased the accuracy to 96.88%. Although large areas of extreme shadowing can cause false detections, this method is found to be robust to environmental variations in pastures. In our future works, weed species recognition is to be carried out such that weed control can be implemented in a more adaptive manner with enhanced effectiveness.

### ACKNOWLEDGMENT

This research has been funded by InnovateUK, UK (Grant Reference: 132339) as part of the collaboration with SoilEssentials Ltd and Aralia Systems Ltd. We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Maxwell Titan X GPU used for this research.

### REFERENCES

- [1] A. Binch and C. W. Fox. "Controlled comparison of machine vision algorithms for Rumex and Urtica detection in grassland." *Computers and Electronics in Agriculture*, 140:123-138, August 2017.
- [2] T. Kounalakis, G. A. Triantafyllidis, and L. Nalpantidis, "Weed recognition framework for robotic precision farming." In 2016 IEEE International Conference on Imaging Systems and Techniques (IST), pp. 466-471.
- [3] F. K. van Evert, G. Polder, G. W. A. M. Van Der Heijden, C. Kempenaar, and L. A. P. Lotz. "Real-time vision-based detection of Rumex obtusifolius in grassland." *Weed Research* vol. 49, no. 2, 2009, pp.164-174.
- [4] F. K. van Evert, J. Samsom, G Polder, M Vijn, H. J. van Dooren, A Lamaker, G. W. van der Heijden, C Kempenaar, T van der Zalm, and L AP Lotz. "A robot to detect and control broad-leaved dock (Rumex obtusifolius L.) in grassland". *Journal of Field Robotics* vol. 28, no. 2 2011, pp.264-277.
- [5] K. Fukushima. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, vol 36, no. 4, 1980, pp. 193-202
- [6] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. "Gradient-based learning applied to document recognition." In *Proceedings of the IEEE*, vol. 86, no. 11, 1998, pp. 2278-2324
- [7] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf. "Deepface: Closing the gap to human-level performance in face verification". In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 1701-1708.
- [8] M. F. Hansen, M L. Smith, L. N. Smith, M. G. Salter, E.M. Baxter, M. Farish, and B. Grieve. Towards on-farm pig face recognition using convolutional neural networks. *Computers in Industry*, 2018, vol 98, pp. 145-152
- [9] 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, July 2002, pp. 971-987.
- [10] M. M. Adankon and M Cheriet. "Support vector machine." In *Encyclopedia of biometrics*, pp. 1303-1308. Springer US, 2009.
- [11] MATLAB 2017b and Statistics and Machine Learning Toolbox, The MathWorks, Inc., Natick, Massachusetts, United States.