

# Weed Classification in Grasslands using Convolutional Neural Networks

Lyndon N. Smith<sup>1</sup>[0000-0001-5821-0586] Arlo Byrne, Mark F. Hansen, Wenhao Zhang and Melvyn L. Smith

Centre for Machine Vision, Bristol Robotics Laboratory,  
Department of Engineering Design and Mathematics,  
University of the West of England, Bristol, UK

<sup>1</sup>Corresponding author: Lyndon.Smith@uwe.ac.uk

**Abstract.** Automatic identification and selective spraying of weeds (such as dock) in grass can provide very significant long-term ecological and cost benefits. Although machine vision (with interface to suitable automation) provides an effective means of achieving this, the associated challenges are formidable, due to the complexity of the images. This results from factors such as the percentage of dock in the image being low, the presence of other plants such as clover and changes in the level of illumination. Here, these challenges are addressed by the application of Convolutional Neural Networks (CNNs) to images containing grass and dock; and grass, dock and white clover. The performance of conventionally-trained CNNs and those trained using ‘Transfer Learning’ was compared. This was done for increasingly small datasets, to assess the viability of each approach for projects where large amounts of training data are not available. Results show that CNNs provide considerable improvements over previous methods for classification of weeds in grass. While previous work has reported best accuracies of around 83%, here a conventionally-trained CNN attained 95.6% accuracy for the two-class dataset, with 94.9% for the three-class dataset (i.e. dock, clover and grass). Interestingly, use of Transfer learning, with as few as 50 samples per class, still provides accuracies of around 84%. This is very promising for agricultural businesses that, due to the high cost of collecting and processing large amounts of data, have not yet been able to employ Neural Network models. Therefore, the employment of CNNs, particularly when incorporating Transfer Learning, is a very powerful method for classification of weeds in grassland, and one that is worthy of further research.

**Keywords:** Weed Detection in Grass, Convolutional Neural Networks, Transfer Learning, Machine Vision.

## 1 Introduction

The presence of weeds in grass reduces productivity and weed consumption by grazing animals can constitute a considerable threat to their health. To combat this the agricultural industry often employs weed control in the form of ‘blanket spraying’, where herbicides such as glyphosate are sprayed across the land irrespective of where the

weeds actually are. As well as being expensive and detrimental to the environment (causing contamination of watercourses and negative effects on surrounding plants and wildlife), this practice also kills other broad-leaf plants, such as white clover, which are desirable since they can reduce fertilizer requirement, as well as comprising a valuable source of nutrition. One of the most common and invasive weeds is the broad-leaf weed known as dock leaf (*Rumex obtusifolius*). One way of controlling this is to employ ‘selective spraying’, where herbicide is manually sprayed on areas with high weed densities. However, since this is highly laborious, it is an expensive procedure. There are, therefore, strong motivators for the development of systems that are able to automatically implement such selective spraying with good precision. The obvious way to direct such an operation is through use of a vision system, and when the weed is of a different color to surrounding soil [1] (or where the weed rises above the surrounding plants), this can be readily achieved. However, “in-pasture” weed detection is considerably harder as the weeds and grass are both predominantly green; and the images are generally busy, with the grass often obscuring the weed [1]. For situations where the amount of weed in the image is relatively small (e.g. less than 5%), its identification is a non-trivial task, even for human observers. The task can however be achieved by machine vision systems that employ conventional image analysis algorithms. For example, use of Local Binary Patterns (LBP) to extract features from local textures from dock leaves in pasture, with a Support Vector Machine (SVM) as the classifier, has been reported to provide an accuracy of just over 80% [1]. In a previous paper [2], we described how we replicated this finding on a large dataset; and how we employed deep learning to increase the accuracy to above 90% while ensuring the system is capable of detecting weeds in different pastures under various representative outdoor lighting conditions. This paper further extends the work by again training a CNN using a large set of images and confirming accuracies of dock detection of around 95%; and then experimenting with ‘Transfer Learning’. This is a recently emerging method that allows the creation of very complex CNN models with relatively little training data and low computational requirements. This was achieved by employing networks previously trained on the ImageNet dataset [3]. The creation of reliable models from restricted datasets is certainly something worth investigating, since a traditional problem with machine learning approaches has been the large amount of data that is generally required to train a model of sufficient sophistication for the task. This paper describes work we have done in this area; and starts with an outline of relevant background research.

## **2 Background Research**

### **2.1 Automated Weeding**

Although modern automation hardware can facilitate the building of an automated herbicide spraying system, the main challenge is how to selectively control the spraying. This requires a vision system capable of accurately and reliably locating weed position. While this is possible using fairly simple machine vision techniques for weeds in bare earth, creating a reliable model for detecting weeds in grassland has proved a much harder task. A few research works have investigated weed detection/recognition

in grasslands by using conventional machine learning methods [4, 5, 6]. More recently, Binch and Fox [1] presented a comparison of the current state of the art in weed detection techniques, using a single ‘Rumex vs Grass’ dataset to train and evaluate them on. They use a classifier which looks to maximize the margin by which a hyperplane separates two classes in feature space. One gap or weakness in these works is the omission of any Neural Network based approaches, in particular ‘deep’ Convolutional Neural Networks (CNNs), which have shown significant advantages over ‘traditional’ machine learning methods in recent years.

## 2.2 Machine Learning

Machine Learning is a subfield of Artificial Intelligence that allows us to solve problems that cannot be solved with explicit programming or conventional mathematical techniques. Taking ideas from the fields of Artificial Intelligence, Computer Science and Computational Statistics, it uses algorithms that can learn from data through a process of self-improvement. Relevant algorithms include Local Binary Patterns (which describe local textures as a feature extractor) and a Support Vector Machine as the classifier. The resulting models produced can then be used to make predictions on similar, domain specific data. Machine learning has been providing effective vision system solutions for 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. One such method is Neural Networks, which is a class of machine learning algorithm/model based loosely on neuronal processes in the biological brain, where a number of neurons are connected together by ‘weights’ (where the model is stored). While the first studies involving artificial ‘neurons’ were in the 1940s, perhaps the first paper to receive widespread interest was the introduction of the Perceptron by Frank Rosenblatt in 1958 [7]. However, at this time it was thought that this approach was not suitable for solving non-linear problems, which led to a long period of limited research in this area. It was not until the 1980s that there was wide realization and exploitation of the capabilities of multi-layer Neural Networks to model any relationship between inputs and outputs, whether linear or non-linear, (provided there is sufficient data from which to generalize). In 1980, Fukushima [8] proposed an architecture inspired by the human visual receptive fields that he named the ‘Neocognitron’. It described alternating layers that convolved and sub-sampled an input image. However, its usefulness was limited without a robust approach to learning the weights. A significant advance was made in 1986 by Rumelhart et al., who presented a new method of weight optimization called ‘Backpropagation’ [9] that allowed several layers of neurons to be stacked together to form a multi-layered network. Backpropagation uses the chain rule to propagate the error signal at the output back through each layer of the network, for calculation

of weights that tend to minimize the error. The work of Fukushima and Rumelhart effectively inspired and enabled the development of the powerful machine learning technique known as the Convolutional Neural Network (CNN).

**Convolutional Neural Networks.** The CNN is a method that was developed for image recognition tasks by taking inspiration from the biological brain, in particular the connective topology of neurons in the visual cortex. In 1999 LeCun described ‘LeNet-5’ [10], a 7-layer CNN that recognized hand written digits; and the power of the new approach was proven when it was used for an automated check reading system that has been used by banks ever since. A useful/intuitive way to think of convolution is as an operation that takes as input two functions: an input signal, and a filter; and outputs a third function that is a modified or ‘filtered’ version of the input. This idea has been used for some time in image processing, where specialized ‘filters’, also called convolutional kernels, have been designed to perform a number of tasks such as image smoothing, sharpening and edge detection; and the generation of specialized, system specific ‘feature detectors’. The function output from this operation is often called a ‘feature map’ and can be used as the input to traditional machine learning/classification algorithms for feature detection - to enhance the performance of the model. The CNN was a new type of Neural Network, that was trained using backpropagation and which had a new architecture. It consisted of several ‘convolutional layers’, each containing several separate convolutional filters that each output a feature map from the layer’s input image. These are then stacked together and sent to the next layer. Also, the network had several ‘fully connected layers’, similar to those traditionally found in Neural Networks. This created a powerful image classification model that combined the feature detection properties of convolutional filters with the versatility of Neural Networks. The final development, in 2009, was to employ Graphical Processing Units (GPUs) as general-purpose processors for training large machine learning models [11]. This enabled great reductions in training times, where a network that would previously take months to train could now be trained in days. This meant that the use of exceptionally large networks was possible; and led to a period of rapid development of CNN architectures. The result was a dramatic increase in the performance of vision systems, and the subsequent research into the application of CNNs across a wide range of sectors, that we see today. There has been only one significant drawback to the CNN approach – traditionally very large amounts of training data have been required to train networks that are sufficiently complex to accurately and reliably perform useful tasks in the real world.

**Transfer Learning and the ImageNet Data-Set.** The method known as ‘Transfer Learning’ involves combining a ‘pre-trained’ network with a limited amount of application specific data. The result is a CNN that has a performance comparable to that of a conventionally fully trained CNN, but a much smaller requirement for training data. In Transfer Learning, the lower convolutional layers of the pre-trained network (with ‘frozen’ weights), are combined with several, untrained, ‘fully connected’ layers. The resulting model is then trained on an application specific dataset, resulting in the training of only the ‘fully connected’ layers. The pre-trained network can be generated by

employing the ImageNet dataset, which contains over 1,000,000 images in 1000 classes (examples include animals, types of vehicles, household objects and many more). In 2012 Krizhevsky et al. developed AlexNet [12], which was a slightly deeper, more complex version of Lecun’s original CNN model, LeNet. When analyzing ImageNet, AlexNet outperformed all other approaches by wide a margin: 15.3% error versus the next best of 26.2% error. Subsequently this method was universally accepted and largely adopted by the machine vision community. In 2014 Ali Sharif Razavian et al. released a paper that explored the following idea: Since the ImageNet dataset is so diverse, the lower convolutional layers of a CNN, trained on this dataset, would contain very general feature detectors that could be used as the low-level feature detectors for a number of machine vision tasks [13]. This theory proved to be true and ImageNet trained CNNs have shown strong performance when compared to other more sophisticated methods. There are now a number of the best performing models, pre-trained on the ImageNet dataset, available to download and use freely. Implementing Transfer Learning in this way has become a very useful tool for creating high performing models without the large data and computational requirements that are often needed to train complex networks. This is important since, for large amounts of data, much effort is required for both the capturing and accurate labelling. The tests described below were undertaken to determine the extent to which Transfer Learning could provide these CNN benefits for automatic detection of weeds in grass.

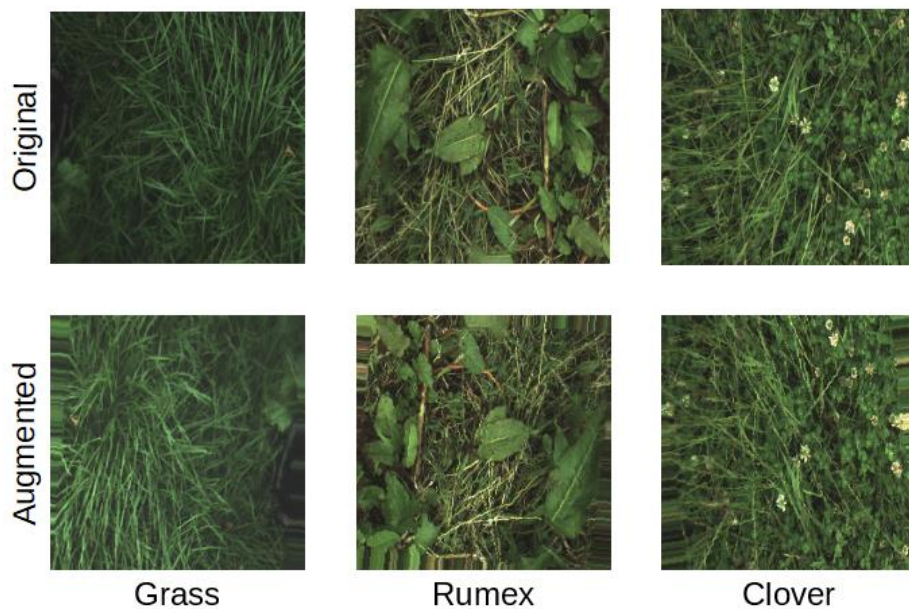
### **3 Methodology**

The work undertaken comprised three parts. The first was to design a CNN based classifier for the task of classifying weeds in grassland, trained from scratch on the subject dataset. This model was based on the ResNet architecture, since it employs a powerful yet simple learning model [14]. The second part compared the performance of several pre-trained CNNs, as pre-trained feature detectors for the task of classifying weeds in grassland, and built a ‘Transfer Learning’ based classifier, using the best model found. The third part compared the performance of Transfer Learning models with conventional CNNs for small datasets. To do this, the two network designs, identified in the first two stages, were trained on increasingly small datasets, to see how their performance varies with the data available.

#### **3.1 Dataset Processing and Usage**

The first dataset used was a two-class dataset, consisting of Rumex and grass, with each class consisting of 2025 images, resulting in a total dataset size of 4050 samples. Images were labelled as Rumex if over 5% of their pixels contained Rumex, with the remainder of the image being filled with grass. This dataset was used for much of the initial testing because, as the two datasets are so similar, it should yield quicker results than the ‘Rumex vs grass vs clover’ dataset and give almost as good an assessment. The second dataset used was a three-class dataset consisting of Rumex, grass and clo-

ver. Due to fewer clover samples being available, and balanced datasets generally yielding less biased models, each class only consists of 1600 samples, resulting in a total dataset size of 4800 samples. Each dataset was divided, with an 85/15% split, into a training set (used to train the models on), and a validation set (solely used to test the model's performance on unseen data). This allows the assessment of the model's true performance and the detection of overfitting. All tests conducted, including those with varying size training datasets, use the same validation datasets, ensuring a fair comparison. This results in a training set of 3442 samples and validation set of 608 samples for the two-class problem (Rumex vs grass), and a training set of 4080 samples and validation set of 720 samples for the three-class problem (Rumex vs grass vs clover). Data augmentation was used to artificially enlarge the dataset being used by applying random transformations such as zooming or rotating the image. This can help to reduce the problems from overfitting that are often present when modelling small yet complex datasets. The Keras TensorFlow library provided a useful tool for real-time image augmentation. An augmentation regime was chosen that consists of the following: First the pixel values are scaled between zero and one; a process that yields slightly better training results. Next the images are randomly rotated in the range of  $-10^{\circ}$  to  $10^{\circ}$ . The images are then randomly shifted in both width and height in the range of 0-2.5%. A random zoom is then applied, zooming in on the images in the range of 0-10%. The final augmentation technique used is to randomly flip images horizontally and vertically. The resulting image is then resized to 224x224 pixels before being presented to the network for training. Fig. 1 shows several examples of augmented images (without pixel scaling applied).



**Fig. 1.** Original and augmented images of: Grass, Rumex and Clover.

### 3.2 Testing

There is not space here to give detailed descriptions of all the tests undertaken; and therefore this section will focus on the testing of the Transfer Learning. Here the performance of Transfer Learning based CNNs are compared to that of ‘conventional’ CNNs (trained from scratch on the subject dataset), while the amount of data available (or the size of the training dataset), is varied. Of particular interest is how each method’s performance degrades when trained on increasingly restricted datasets; since an ability to create complex models with small amounts of data is very useful. The best conventional ResNet network and the best Transfer Learning based network were trained on several training datasets of varying size, while being tested on the same validation datasets used throughout. The training details, the training and validation accuracy, and training and validation loss, were recorded for assessment. The conventional ResNet network was trained from scratch, and the other network was trained using Transfer Learning. Each network was trained on data batches of 64 samples at a time, for 200 epochs, and employed the same learning rate decay strategy as was used in previous sections, with the initial learning rate being divided by ten after the first 150 epochs. The initial learning rate was taken from previous tests. The ‘Stochastic Gradient Descent’ optimization algorithm, with a momentum of 0.9, and ‘L1’ loss function was also be used, as in previous tests. Six dataset sizes were tested, based on the number of samples in each class, and the class sizes used were 10, 25, 50, 250, 500 and 1721 / 1360 samples (full training dataset). This resulted in full training dataset sizes of 20, 50, 100, 500, 1000 and 3442 samples for the two-class problem (Rumex vs grass); and full training dataset sizes of 30, 75, 150, 750, 1500 and 4080 samples for the three-class problem (Rumex vs grass vs clover). This gave a good idea of how performance degrades with the loss of training data, for each method tested.

## 4 Results

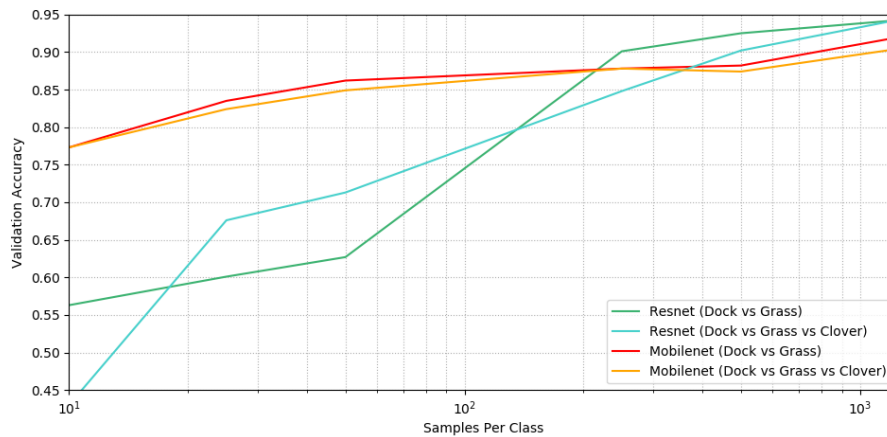
### 4.1 Final Models

The best conventionally trained model identified was a 26-layer ResNet network with 16 filters in the first convolutional layer and dropout applied to convolutional layers with a dropout rate of 0.15. The best Transfer Learning based model used the convolutional layers of an ImageNet pre-trained MobileNet network [15], which is a streamlined version of Xception [16] that was developed by Google for use in their mobile devices. This model is being included because Xception was an important architecture that introduced a new model for convolutional layers. Also, since any weed detection model would have to run on mobile hardware, it is useful to see how a mobile architecture could perform. The MobileNet network employed two ‘dense’ layers of 1024 and N neurons, where N is the number of classes in the dataset, and dropout was applied to the dense layers with a dropout rate of 0.7. These models were trained for the Rumex vs grass dataset and for the Rumex vs grass vs clover dataset. For the first dataset the best validation accuracy achieved by the ResNet network was 95.6% and the best achieved by the MobileNet network was 93.2%. For the second dataset, the ResNet

network achieved a validation accuracy of 94.9% and the MobileNet network achieved 91.1%. The conventionally trained ResNet network still outperformed the MobileNet network, providing the best classifier for both datasets presented.

## 4.2 Transfer Learning vs Conventional Networks

The performance of conventionally trained neural networks was compared to that of Transfer Learning based approaches, as the data available for the models to learn from was reduced. The two models identified above were trained on several datasets of diminishing size while being tested on a single validation dataset. Fig. 2 shows the validation accuracy achieved by each model plotted against the number of data samples per class in the training dataset used.



**Fig. 2.** CNN (Fully Trained ResNet26 vs Transfer Learned MobileNet) validation accuracy against Samples per class for identification of Rumex (Dock) in Grass and with Grass and Clover.

While the conventionally trained network achieved higher accuracy scores if there were over 250-500 data samples per class in the dataset, this performance quickly dropped as the data available was reduced, and for the smaller dataset the ResNet network was able to achieve little more than random guessing. The pre-trained MobileNet network on the other hand was far more robust towards reduction in dataset size; and was able to achieve competitive results with as little as 25 samples per class in the dataset, thereby showing the true strengths of this method.

## 5 Discussion

One of the main questions posed here has been what performance gains are possible by using Convolutional Neural Networks for weed classification in grasslands - in particular identifying Rumex in grass. A review of the current state of the art (non-CCN)



techniques for this weed detection stated that: ‘... In their study, the single best algorithm had a classification accuracy of 82.88% and the best ensemble of models had a classification accuracy of 83.4%.’ In our work, the best classification accuracy achieved was 95.6%, from the 26-layer ResNet network, showing considerable improvements over previous techniques. However even more impressive is the fact that the best ‘transfer learning’ based model achieved a classification accuracy of 83.5% using a dataset of just 50 samples. This shows higher results than previous techniques of [1], while only requiring a training dataset that could be collected and processed by a single person in a day - clearly showing the advantages possible with these methods.

Initial testing conducted on training a ResNet network from scratch on the subject datasets showed that a smaller architecture, in both depth and width, led to a higher validation accuracy from the resulting model. Two possible reasons for this are that the network could either be overfitting the dataset and may yield better results from techniques such as dropout, or that the network is so big that it learns exceptionally slowly so that, given time, it may lead to a superior model. One interesting result to come from these tests was that applying dropout to the convolutional layers of the network, with a very small dropout rate, produced gains in validation accuracy- a technique that is generally believed to only yield positive results when applied to dense layers. Out of the ImageNet pre-trained models, MobileNet showed significantly superior results to any other model assessed. The main problem found with using this model was its tendency to over-fit the dataset it is being trained on. While this problem was significantly reduced with the use of dropout in the dense layers, it was still present for the Rumex vs grass vs clover dataset.

In the tests that looked at the performance of both conventionally trained networks and Transfer Learning’ based networks when trained on increasingly small datasets, the clear advantage of Transfer Learning was illustrated. Where the conventionally trained network’s performance sharply dropped as the data available was reduced, the Transfer Learning based model continued to give competitive results from very small datasets. This is a very useful feature for small businesses and projects where acquiring and processing the large amounts of data required for conventionally trained networks is not always possible.

## **6 Conclusion**

This work has addressed the problem of building a model for classifying weeds in grasslands - a problem that has stalled the development of autonomous ‘selective’ herbicide spraying systems for some time. The recent leaps in image classification accuracy made using Convolutional Neural Networks, combined with the potential substantial ecological and cost savings of automated selective spraying, have provided the motivation for this work.

The main goals were to:

- Assess the performance of Convolutional Neural Networks for weed classification, looking at the gains possible over the previous state-of-the-art, taken from [1].
- Examine the performance of both conventionally trained networks, and the use of networks pre-trained on the ImageNet dataset - a process known as 'Transfer Learning'.

From the testing conducted two models were identified for the task of weed classification: a 26-layer ResNet network that was trained from scratch on the subject datasets, and a Transfer Learning based model that used the convolutional layers of an ImageNet pre-trained MobileNet network in conjunction with two 'dense' layers trained on the subject datasets. While both models showed improved classification accuracy when compared to previous techniques, it was the conventionally trained ResNet network that provided the best results, with a classification accuracy of 95.6% on the two-class dataset and 94.9% on the three-class dataset. This shows a considerable step in performance when compared to 'traditional' models and opens many options for agricultural automation. The final observation was how little data was required to create Transfer Learning based models using ImageNet pre-trained networks. This method showed competitive results when trained on as few as 25 samples per-class in the dataset and substantially reduces the data required to train complex models. This method therefore comprises a viable option for businesses that, due to the high costs of collecting and processing large amounts of data, previously could not employ the use of Neural Network based models.

To summarize, the tests conducted in this work have shown the improvements possible for weed classification models when using Convolutional Neural Networks and Transfer Learning. The results suggest that these methods have potential to provide reliable and practical multi-species weed classification in real time in-the-field; thereby enabling development of automated selective weed eradication systems, and realization of all the important associated long-term benefits.

## References

1. Binch, A., Fox, W. "Controlled comparison of machine vision algorithms for Rumex and Urtica detection in grassland." *Computers and Electronics in Agriculture*, 140:123-138 (2017).
2. Zhang, W., Hansen, M. F., Volonakis, T. N. , Smith, M., Smith, L., Wilson, J., Ralston, G., Broadbent, L., Wright, G., SoilEssentials, Aralia Systems Ltd. "Broad-leaf weed detection in pasture". *IEEE 3rd International Conference on Image, Vision and Computing (ICIVC)*: IEEE Available from: [http://eprints.uwe.ac.uk/38754\(ICIVC\)](http://eprints.uwe.ac.uk/38754(ICIVC)), Chongqing, China, 27-29 June (2018).

3. Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., Fei-Fei, L. "ImageNet: a large-scale hierarchical image database", 2009 IEEE Conference on Computer vision and Pattern Recognition, Miami, FL, 248-255 (2009).
4. Kounalakis, T., Triantafyllidis, G. A., and Nalpantidis, L. "Weed recognition framework for robotic precision farming." In 2016 IEEE International Conference on Imaging Systems and Techniques (IST), 466-471 (2016).
5. van Evert, F. K., Polder, G., Van Der Heijden, G. W. A. M., Kempenaar, C., and Lotz, L. A. P. "Real-time vision-based detection of *Rumex obtusifolius* in grassland." *Weed Research* vol. 49, no. 2, 2009, 164-174 (2009).
6. van Evert, F. K., Samsom, J., Polder, G., Vijn, M., van Dooren, H. J., Lamaker, A., van der Heijden, G. W., Kempenaar, C., van der Zalm, T., and Lotz, L. A. P. "A robot to detect and control broad-leaved dock (*Rumex obtusifolius* L.) in grassland". *Journal of Field Robotics* vol. 28, no. 2, 264-277 (2011).
7. Rosenblatt, F. "The perceptron: A probabilistic model for information storage and organization in the brain". *Psychological Review*, 65(6), 386-408 (1958).
8. Fukushima, K. "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, 193-202 (1980).
9. Rumelhart, D., Hinton, G., and Williams, R. "Learning representations by back-propagating errors". *Nature* 323 (6088): 533—536 (1986).
10. LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. "Gradient-based learning applied to document recognition." In *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278-2324 (1998).
11. Raina, R., Madhavan, A., and Ng, A. "Large-scale deep unsupervised learning using graphics processors". <http://dx.doi.org/10.1145/1553374.1553486> (2009).
12. Krizhevsky, A., Sutskever I., and Hinton, G. E. "ImageNet Classification with Deep Convolutional Neural Networks". Part of: *Advances in Neural Information Processing Systems* 25 NIPS (2012).
13. Razavian A. S., Azizpour, H, Sullivan, J., Carlsson S. "CNN Features off-the-shelf: An Astounding Baseline for Recognition". *arXiv:1403.6382* (2014).
14. He, K., Zhang, X., Ren, S., and Sun, J. "Deep Residual Learning for Image Recognition". *arXiv:1512.03385* (2015).
15. Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., and Adam H. (2017). *MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications*. *arXiv:1704.04861 [cs.CV]* (2017).
16. Chollet, F. "Xception: Deep Learning with Depthwise Separable Convolutions". *arXiv:1610.02357 [cs.CV]* (2017).

## Acknowledgements

The image data employed in this work came from the following project: "GrassVision - Automated application of herbicides to broad-leaf weeds in grass crops". This was funded by The UK Technology Strategy Board, Agri-Tech Catalyst - Early Stage Feasibility - Round 5 (Ref. 132339); and involved CMV, Soil Essentials Ltd and Aralia Systems Ltd. The project that this paper is based upon was undertaken by Arlo Byrne in CMV at UWE, with supervision by Dr. Matthew Studley and Dr. Mark Hansen.