Using Neural Networks for Maintenance Tasks in Agriculture: Precise Weed Detection
Patricio Nebot*, Joaquín Torres-Sospedra and Gabriel Recatalá

Department of Computer Science and Engineering, Universitat Jaume I, Campus del Riu Sec, Avda. Sos Baynat S/N, E12071, Castellón, Spain.
*Corresponding author. E-mail: pnebot@uji.es

Abstract
The problems derived from herbicide utilization compel people to use new methods of weed control. Their main purposes are reducing the herbicides and maintaining the crop production. These are the principal reasons why we introduce a new alternative procedure for weed detection in orange groves, which consists of two different parts or stages. In the first one, the main elements of the grove are determined. In the second one, the system focuses on the soil areas to detect weeds. In this paper, this system is described and some experiments are shown to validate it. The experimental results show that the two stages are properly carried out and it can be considered a new alternative for weed detection.

Key words: Ensembles of Artificial Neural Networks, Terrain Classification, Weed Detection and Agricultural Robotics

1. Introduction
In agriculture, weed control is a time-consuming and expensive activity. Moreover, the long-term use of herbicide is a potential source of pollution, which could damage people, animals and the environment. In fact, agricultural herbicides have been uniformly sprayed in fields and overused in a conventional way for several years. Actually, selective spraying is manually done as shown in Fig.1. Manually spraying had made severe environmental pollution. Therefore, the research efforts are being encouraged in designing weed-detecting technologies for precision spraying with selective herbicides with the final purpose of saving herbicides and reduce environmental pollution without sacrificing crop yield.

FIGURE 1: Manual spraying of an orange grove

One alternative to apply efficiently herbicides consists of using a robotic system to detect and treat weed zones autonomously. However detecting automatically weed areas on an orange grove is not trivial and involves some varying factors (weather and lightning conditions).
In this paper, we propose a two-staged system to detect weeds in orange groves based on neural networks. Although using classifiers to detect weeds is not new (see Burks et al. (2005), Karimi et al. (2005), Wang et al. (2007), López-Granados (2011), and Tellaeche et al. (2011)), our approach differs from previous work:

1. We use a simple CCD camera avoiding the use of special sensors. In this way, our approach can be considered cheaper than other that use hyperspectral vision systems, remote imagery or other special devices.
2. We detect weeds in high areas by using a camera with frontal view. Other systems use zenithal view to capture only soil, so weeds are detected in small areas.
3. In our application, the ground is not clean and it can contain “garbage” elements such as leaves, orange fruits or stones. In other researches, weed detection have been applied to more controlled environments which were much more clear than standard orange groves.
4. Usually in the literature, the weed detectors based on neural networks use a single network to perform the classification. In this paper, ensembles of neural networks are used instead of the single network. It has been proven that ensembles improve the generalization ability with respect to the single network when the networks that compose the ensemble are not correlated.

The rest of this paper is organized as follows: In Section 2, some concepts related to ensembles of neural networks are introduced. The procedure used to detect weeds is described in Section 3. The experiments carried out are shown in Section 4.

2. Neural Networks and Ensembles

Firstly, the Multilayer Feedforward Network (MF) is the network architecture used in the experiments carried out in this paper. This is a feed forward neural network with an architecture that closely resembles the layered machine proposed by Nielson in 1965 and the α-Perceptron proposed by Rosenblatt in 1961 as described in Pao (1989). Concretely, we use a Multilayer Perceptron (MLP), which consists of three layers of computational units. The neurons of the first layer apply the identity function whereas the neurons of the second and third layer apply the sigmoid function.

The Backpropagation algorithm, Rumelhart et al. (1986), has been applied to perform the training stage of the networks. In any iteration of the algorithm, the networks are adapted by using patterns (known samples or instances) from the training set. At the end of the iteration, the Mean Square Error, MSE, has been calculated by classifying all the patterns from the validation set (a set with patterns or instances not used in the training set). When the learning process has finished, the weights of the iteration with lowest error on the validation set are assigned to the final network. This regularization using validation sets is done to avoid overfitting and improve the generalization of the networks.

The ensemble model used in this paper is Simple Ensemble. This is the simplest alternative to generate ensemble and it consists of training a set of independent networks with different weight initialization. In this way, each network provides an output vector (classification) for a given unknown pattern. The average of all the outputs vectors conforms the final output vector with the global classification.

Detecting weeds using ensembles of neural networks is introduced in this paper. In this case, a capture image is processed to obtain a map or the classification image as shown in Fig. 2. Since a capture can not be considered a pattern, it has been divided into some tiles (this fact will be detailed in Section 3) which can be related to the concept of pattern. The information extracted from each tile is processed by the ensemble of neural networks. Then, the label or class associated to the pattern is obtained directly from the global output.
3. Weed detection system based on Ensembles of Neural Networks

In this section, the proposed two-stage alternative to detect weeds with neural networks is detailed in the following subsections.

3.1. First Stage: Detection of main elements of the grove.

The main aim of this stage is to detect the main elements of the grove. With main elements we mean: Orange trees, Soil and Sky.

Firstly, the vision system captures an image from an orange grove with the VGA camera (640×480 pixels). An example of capture is shown in Fig.3 along with the truth image, the truth image contains the classification of elements done by a human expert. Concretely the areas are catalogued as Sky (white), Land or Soil (red), Orange trunks (yellow) and Orange Crown (green). In this example image, there is not any orange trunk due to the characteristics of the grove.

The resolution of the truth image is lower than the original capture (159×119 pixels) according to the classification performed for terrain classification. The original capture has been divided into overlapping tiles (or areas) of 8×8 pixels. Each tile is considered as a pattern for the neural networks and the ensembles used to perform the classification. The use of overlapping tiles has been introduced in order to increase the resolution and precision of the classification image. The size of the classification image would have been lower (80×60 pixels) without the use of overlapping tiles.

The procedure used to perform the terrain classification used in this paper is fully described in Sung et al. (2010) and it is briefly described in this section. The features used as inputs in the ensemble are calculated as follows. Firstly, the two-level Daubechies wavelet transform is applied to each HSI channel from the image provided by the vision system. With this procedure, seven sub-band images are obtained for each channel. Then two features, Mean
and Energy are calculated with using only those pixels of the sub-band images which correspond to the tile of the captured image. Although this procedure provides 42 inputs features (1- The image contains 3 channels. 2- For each channel 7 sub-band images are generated with wavelets. 3- 2 values are calculated for each subchannel ) only 24 of them were used due to the unuseful data introduced by the discarded ones. In the first experiments, these 24 features will be also used to perform the classification along with the two spatial coordinates (x and y) related to the centre of the tile.

Then, each network of the ensemble processes all the tiles of the capture and provides an output vector. The individual output vectors are averaged in order to obtain the global output vector. This final output vector is interpreted to obtain the class (element of the grove) as shown in Fig.4.

3.2. Second Stage

Once the first classification is done, weeds are detected only on those areas which corresponds to Soil (those zones shown in red in Figure 3).

Firstly, a binary mask is calculated. This mask determines if a pixel of the classification image corresponds to soil or not. Moreover, an erode/dilate filter has to be applied to this mask to obtain better results. An example of this mask (before and after filtering) can be found in Fig.5. This final mask allows determining which pixels from the original image will be used by the second classifier in order to detect weeds.

At this stage, the tile size and the input vector for the classifiers are changed. Concretely, the input features contain the RGB values for all the pixels of the tile. Moreover, the capture is divided into 2x2 pixels tiles.
The soil elements are processed by another ensemble of neural networks (the classification task is totally different in this second stage). The input vector of the second ensemble is only composed by twelve elements (4 pixels denoted with the 3 normalized RGB values) because the current task, determining if weeds are present or not, is less complex than the first terrain classification. This detection is "simpler" because the elements can be better differentiated (there are less colour similarities between the soil and the weed areas). Moreover, the spatial coordinates are not useful in weed detection since weed areas can be located in any position of the soil.

An example of this second classification was shown in Fig. 2, where the original capture is shown with the final classification. In this image, the zones drawn in black refer to weed zones, whereas the other colours refer to the main elements of the grove, which were detected in the first stage of the classification procedure, as shown in Figures 2 to 4.

4. Experimental Results

Table 1 shows the training parameters used to perform terrain-based classification and weed detection, whereas Fig. 6 shows some examples of high-level weed detection done by our system.

<table>
<thead>
<tr>
<th>TABLE 1: Training parameters for the MF networks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Hidden nodes</td>
</tr>
<tr>
<td>Adaptation step</td>
</tr>
<tr>
<td>Momentum rate</td>
</tr>
<tr>
<td>Number of epochs</td>
</tr>
<tr>
<td>Number of networks</td>
</tr>
</tbody>
</table>

FIGURE 6: Examples of high-level weed detection with the proposed system
5. Conclusions and future work

This paper has introduced a new alternative for weed detection in orange groves. This system is carried out in two stages where the main elements are first determined and then weeds are detected on Soil areas. We consider that it is an important contribution to autonomous weed detection.

Firstly, the application is done in low controlled environments. Although the basic structure of the grove is known, there can be different configurations in placing the trees. Moreover, some unknown elements, garbage, can be placed in Soil. The proposed system deals properly with all these issues and it could be adapted to other groves (e.g. olive groves). Generally, other related works introduce weed detection in more controlled environments.

Secondly, the system uses general-use purpose devices to perform the different tasks such as the weed detection introduced in this paper. Using specific sensors has a direct impact on economical costs and power consumption. Although vision based sensors are not new, they are commonly used in zenital view in the literature.

Finally, the system is able to adapt the detection task to the lightning and weather conditions. In this way, the neural networks are more flexible than using rule-based systems. Moreover, the weights of the networks could be specifically set depending on the location, kind of grove and lightning conditions. Furthermore, feedback could also improve the classification rate.

Acknowledgments

This research was partially supported by UJI-Bancaixa with the project P1-1B2009-39 entitled “RETA”.

References


