Machine Learning paradigms for Weed Mapping via Unmanned Aerial Vehicles

M. Pérez-Ortiz*, P.A. Gutiérrez†, J.M. Peña‡, J. Torres-Sánchez‡, F. López-Granados‡ and C. Hervás-Martínez†
*Dept. of Quantitative Methods, Universidad Loyola Andalucía, 14004 Córdoba, Spain.
Email: {mariaperez}@uloyola.es
†Dept. of Computer Science and Numerical Analysis, University of Córdoba, 14071 Córdoba, Spain.
Email: {pagutierrez,chervas}@uco.es
‡Institute for Sustainable Agriculture, CSIC, 14080 Córdoba, Spain.
Email: {jmpena,jtorres,flgranados}@ias.csic.es

***Abstract*—This paper presents a novel strategy for weed monitoring, using images taken with unmanned aerial vehicles (UAVs) and concepts of image analysis and machine learning. Weed control in precision agriculture designs site-specific treatments based on the coverage of weeds, where the key is to provide precise weed maps timely. Most traditional remote platforms, e.g. piloted planes or satellites, are, however, not suitable for early weed monitoring, given their low temporal and spatial resolutions, as opposed to he ultra-high spatial resolution of UAVs. The system here proposed makes use of UAV-imagery and is based on: 1) Divide the image, 2) compute and binarise the vegetation indexes, 3) detect crop rows, 4) optimise the parameters and 4) learn a classification model. Since crops are usually organised in rows, the use of a crop row detection algorithm helps to separate properly weed and crop pixels, which is a common handicap given the spectral similitude of both. Several artificial intelligence paradigms are compared in this paper to identify the most suitable strategy for this topic (i.e. unsupervised, supervised and semi-supervised approaches). Our experiments also study the effect of different parameteres: the flight altitude, the sensor and the use of previously trained models at a different height. Our results show that 1) very promising performance can be obtained, even when using very few labelled data and 2) the classification model can be learnt in a subplot of the experimental field at low altitude and then applied to the whole field at a higher height, which simplifies the whole process. These results motivate the use of this strategy to design weed monitoring strategies for early post-emergence weed control.**

I. Introduction

Crops require the use of herbicides as a tool for maintaining and ensuring the quality and quantity of crop production. Herbicides are usually broadcast over entire fields even although there are weed-free areas because weeds are usually spatially distributed in patches [1]. Nowadays, the estimated yield loss because of weeds is 9% (being aproximately €3.334M the cost of all herbicides used, 41.5% of the total pesticides sales and 40% of the cost of all the chemicals applied to agricultural land in Europe). There are, however, clear economical and environmental risks derived from over-application, which have led to the creation of the European legislation for the sustainable use of pesticides, which has set guidelines for the reduction of these chemicals [2]. On this sense, patch spraying has enabled the use of site-specific weed management (SSWM) according to the coverage of weeds [3]. One of the keys for ensuring early SSWM are precise weed maps, which should be provided timely for appropriate post-emergence weed monitoring. Until now, weed monitoring [4] has been performed either by remote detection or ground sampling [5]. In this sense, remote sensing has been seen to improve significantly the reliability of SSWM (against ground sampling), provided that the spatial and spectral resolutions of the equipment is suitable for the distinguishing differences in spectral reflectance [6]. However the appearance characteristics of crop and weeds are very similar in early growth stages. To solve this, previous works have mapped weeds at late growth stage (e.g. flowering) using piloted aircrafts or QuickBird satellite imagery. Nonetheless, because of their scarce spatial resolution, these technologies can not be successfully applied in early detection. However, a new aerial technology has recently joined the traditional ones: the Unmanned Aerial Vehicle (UAV) [7] and its advantages against airborne or satellite missions have been demonstrated in different studies [8], [9], i.e. a lower cost, more flexibility in flight scheduling and the possibility of acquiring ultra-high spatial resolution images. UAVs represent then a promising tool for multi-temporal studies in early crop and weed mapping [10], [11], which has been one of the classic limitations with traditional remote-sensed technologies.

Machine learning and image analysis has been used recently for precision agriculture using UAV-imagery in some recent works [12], [13], [14], [15]. However, this is still a mostly undeveloped area (despite its potential). In this sense, a popular choice for designing a weed management strategy using UAVs is the use of manually-defined rules [9], [16] (based on spectral differences, location and vegetation indexes). Nonetheless, we believe that remote sensing will benefit to a large extent from different techniques of image analysis and machine learning. This type of methods has been used with on-ground images with success [17], [6], [18], which motivates further research in this line. Proximal sensing, however, presents some limitations that make its use difficult in practice [18] (computational resource limitations given that it is usually performed in real-time, vibration of the equipment, changes of luminosity and others). Opposed to this, with remote sensing, the analysis should be performed prior to broadcasting, but it could be also useful to estimate *a priori* the needed quantity of herbicide and

optimise the field path that the broadcasting equipment should follow. It is now, that the most common issues with UAVs have been mastered (e.g. route planning) and that the cost of this technology is acceptable, that this technology is ready for its use. Up to this date, different studies have analysed the advantages and feasibility of this idea, and have proposed as well new techniques for weed monitoring, testing them in different experimental setups. This has shown great promise for the detection of weeds between crop rows [9], [11], but distinguishing weeds within crop rows still represents an open and complex challenge. The major difference between our proposal and the rest of papers in the literature is the use of a wide range of machine learning approaches and the combined use of these with an accurate row detection strategy, resulting in a precise and robust system to distinguish weeds either outside or within crop rows.

This paper proposes a novel system for weed monitoring in sunflower via UAV-imagery. The analysis of the images is performed using machine learning and image analysis. The system here proposed is defined so that it minimises the information provided by the user (i.e. in this case a set of labelled patches for each class and the set of parameters for the algorithm) since this is the major bottleneck in a framework like this. The proposed methodology is composed of the following steps: divide the field image (into a set of smaller subimages), compute vegetation indexes and binarise them, detect crop rows and, finally, train a prediction model. The main novelty of our approach is the detection of crop rows using the Hough transform (HT) [19] and images taken by UAVs (although HT has been employed in on-ground studies [17], it has not been so common with remote-sensed platforms [20]).

Concerning machine learning, most standard classifiers are based on learning a discriminant function using a labelled data (i.e. supervised learning). Nonetheless, obtaining labelled data can be time-consuming and expensive, in contrast with unlabelled data. In this line, the new system proposed here is based on two main intentions: perform the analysis of the image with scarce a priori knowledge and processing of the image, and optimise all the parameters automatically. The second novelty of our work is the comparison between three areas of machine learning (unsupervised, supervised and semi-supervised learning) to study the importance and influence of labelled and unlabelled data. The objectives of this paper are the following: 1) analyse how to combine UAV images with the information that a row detection algorithm provides, to improve the performance and help to distinguish weeds within crop rows; 2) analyse the potential of several machine learning strategies in the development of new monitoring system (basing our analysis on easy to compute information, then providing a mostly unsupervised analysis that can be used in other scenarios); 3) compare several factors such as the sensors or the flight altitudes; and finally 4) analyse the generalisation ability of the final models to other scenarios, to alleviate the whole process. This paper complements our previous work in [15], including a more thorough analysis of the results and an additional experiment.

This paper is organised as follows: Section II describes the data acquisition and summarises the proposed system; Section III describes the experimental setup and analyses the results. Finally, Section IV outlines the conclusions and future work.

II. Materials and Methods

Fig. 1 presents a summary of the steps of the weed monitoring system proposed here. As can be seen, the user only provides the UAV-images, label some pixels and set the initial values for the parameters of the system. As said, the steps of our system are the following: Divide the whole image into more tractable subimages for the use of the different algorithms employed later, calculate the well-known vegetation indexes (depending on the sensor), binarise this vegetation indexes (needed step for row detection), use the HT for crop row detection, and finally, optimise the parameters and classify the pixels as soil, crop or weed.

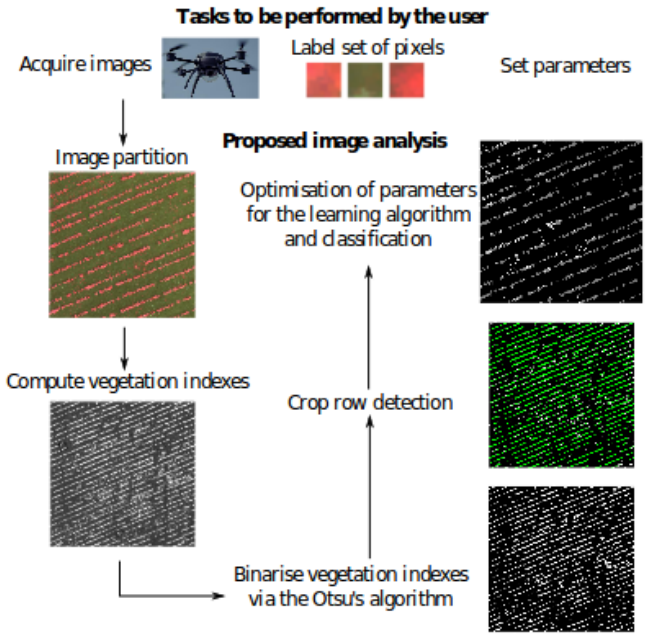

Fig. 1. Representation of different stages for the proposed algorithm for an image obtained using the TetraCam sensor.

A. Data acquisition

a) Studied area: The studied area is a sunflower field situated in Seville at the South of Spain (private farm La Monclova in La Luisiana, coordinates 37.527N, 5.302W, datum WGS84). The flights were authorised by a written agreement between the owners and our research group. The sunflower seeds were planted at the end of March 2014 at 6kg ha^{-1} in rows 0.7m apart. The set of aerial images were collected on May 15$^{\text{th}}$, just when post-emergence herbicide or other control techniques are recommended. An experimental plot of 100×100m was used for the flights. An on-ground sampling procedure was performed the day of the UAV flights, to

compare on-ground weed infestation (observed weeds versus the output of our system).

b) UAV flights and sensors: A quadrocopter platform (model md4-1000, microdrones GmbH, Siegen, Germany) was used for this study taking images at different flight heights. The coordinates of each corner were collected with a GPS. Then, the flight route was programmed into the UAV.The images were taken at three different altitudes: 30, 60 and 100m, with two different sensors mounted separately: a six-band multispectral camera, model TetraCam mini-MCA-6 (TetraCam Inc., Chatsworth, CA, USA), and a still point-and-shoot camera, model Olympus PEN E-PM1 (Olympus Corporation, Tokyo, Japan). The TetraCam mini-MCA-6 is a multispectral sensor composed of six individual digital channels arranged in a 2×3 array. Each channel has a focal length of 9.6mm and a 1.3 megapixel ($1,280 \times 61,024$ pixels) CMOS sensor. The camera has user configurable band pass filters of 10-nm full-width at half-maximum and centre wavelengths at blue (B, 450nm), green (G, 530 nm), red (R, 670 and 700 nm), R edge (740 nm) and near-infrared (NIR, 780 nm). The Olympus camera acquires 12-megapixel images in true colour (R, G and B bands) with 8-bit radiometric resolution and is equipped with a $14 - 42$mm zoom lens. A sequence of 60% end or longitudinal lap and 30% side or lateral lap images were collected corresponding to each flight mission sensor and altitudes. Further information concerning the configuration of the UAV and the specifications of this platform and the sensors used can be found in [10].

c) Image preprocessing: Several overlapping images were collected to cover the whole experimental field. This is because UAVs fly at low altitudes (because of the Spanish regulation the maximum altitude is 120m for UAVs with a weight lower than 25kg) which makes it necessary to take a sequence of multiple images. A crucial step is then the combination of these images using orthorectification and mosaicking. To do so, the Agisoft Photoscan Professional Edition (Agisoft LLC, St. Petersburg, Russia) software was employed. Further details concerning mosaicking can be found in [21].

B. Image partition and labelling

One of the objectives of this system is to provide timely weed mapping responses. To do so, the processing step has to be light in terms of: 1) user intervention (therefore the need to use very few labelled information and optimise automatically all the parameters of the system) and 2) computational time (which is performed dividing the process into independent chunks of computation).

Thus, after the orthorectification and mosaicking steps, the mosaicked image is partitioned into multiple subimages of approximately 1000×1000 pixels. Each subimage is separately processed, i.e. an independent classifier is trained for every subimage. Note, however, that the training labelled samples are always the same, and the only information that change is the unlabelled one. This partition is also important to consider potential differences in the spectral information of the experimental field or lighting changes.

To alleviate the process of labelling, the expert is asked to label only three patches of pixels, each corresponding to one of the classes considered: soil, crop and weeds. In this case, the patches are chosen of size 10×10 pixels, to analyse whether it is possible to learn a model with such little labelled information. Fig. 1 includes a visual representation of these patches for the TetraCam sensor.

C. Vegetation indexes (VI) and crop row detection

The term VI refers to a mathematical expression that combines the surface reflectance at two or more wavelengths of an image in order to highlight a particular property of vegetation. In this work, two widely-used indices for vegetation estimation are compared to analyse their performance: 1) the normalised difference vegetation index (NDVI) [22], [11] is used with the images obtained from the multispectral camera (TetraCam) to consider the near infrared (NIR) wavelength; and 2) the excess green index (ExG) [23], [11] is used with visible images (Olympus sensor).

One of the main hypothesis of this paper is that weed discrimination can be improved using the relative location of weeds with respect to crop lines. The Hough transform (HT) [19] is a powerful technique for the detection of complex patterns in binary images (e.g. lines or circles). This method relies on a parametrisation (defined beforehand) that characterises the patterns that the algorithm is aiming to find (e.g. the parameters for the case of a line are the corresponding slope and intercept). This technique converts a difficult global detection problem in the image space into an easier local peak detection problem in the parameter space.

The HT method is applied to the VI considered. As stated before, the HT method is used for binary images. Because of this, a thresholding procedure is necessary to binarise the VI. In this paper, the Otsu's method [24] is selected, given its demonstrated robustness and simplicity [11].

The HT method detects only the first and last points of a crop row, but these rows should have a predefined width (which is associated to the crop and height). Consequently, the pixels adjacent to the lines should also be considered part of the crop row. In [15] we develop a method to automatically optimise this parameter and the rest of parameters of the proposed system. In this case, the distance from each point to the nearest crop row is computed. A threshold is then applied to these distances to decide the points of the buffer for the crop row. This threshold is optimised using an evaluation metric which captures the amount of pixels with a high value of VI that are covered by the detected crop rows [15].

D. Features

The detected crop rows have to be incorporated in the classification process. To do so, we propose a new feature to incorporate to the learning process of the classifier (apart from the spectral information and VI, which would be the most straightforward idea). This feature is computed by multiplying

the VI and the crop row binary mask (the result being zero for points outside the buffer of the crop row and the VI value for points within the buffer). Ideally, this will allow to distinguish between pixels which lie very far from the crop rows (bare soil or weeds in between crop rows) and pixels which lie closer to these crop rows (crop or weeds within crop rows). However, at the same this feacture also incorporates information about the VI (e.g. in order to distinguish between the different pixels within the buffer, which is the most common handicap in weed mapping).

E. Machine learning paradigms

As said, three paradigms are considered in this paper to approach the previously-mentioned problem. Note that unsupervised learning alone does not match the definition of our problem. Moreover, labelling all the pixels in the experimental field (or, at least, a majority of them) is a hard and tedious work, and almost intractable for practical applications. Because of this, we consider a different strategy, and guide the learning process using only a 10×10-pixel window for each group of pixels (i.e., crop, weed and soil). The idea is to check if this is feasible and help us to capture the nature of these three classes.patch of pixels for each group. Semi-supervised learning can also be considered, where both labelled and unlabelled data is included in the model.

Note that the pixel information is used differently depending on the paradigm considered. For unsupervised learning, it is used only for initialising the data centroids, but the training data correspond to each subimage. For supervised classification, labelled data are used for the training phase, and the rest (i.e. the subimage) is only considered for prediction. For semi-supervised approaches, it is used for the training phase, in conjunction with all of the unlabelled data (i.e. the rest of the image). For more information concerning the system proposed for weed mapping refer to [15].

d) Unsupervised learning: The most widely used clustering method is the k-means algorithm [25]. Two versions of this algorithm are included in the experimentation to analyse whether averaging a small part of the image could be sufficient to initialise the centroids. These versions are the following: k-means (which computes the initial data centroids by averaging each of the labelled 10×10 subimages) and repeated k-means (Rk-means), which chooses a random pixel of each 10×10 subimage (a procedure that is repeated 30 times) and the final centroids are the ones which produce the best results in train.

e) Supervised learning: Supervised learning concerns the design of a prediction model using labelled samples. In our case, we test the performance of these methods when very few labelled data is considered (300 pixels, 100 for each class). We test two well-known and widely used classification methods: k-nearest neighbours (k-nn) and Support Vector Machines (SVM) [26]. Both linear and kernel versions of the SVM algorithm are tested, to compare the results (between the two methods and the semi-supervised linear version).

f) Semi-supervised learning: Semi-supervised approaches [27] are based on the idea of deriving a decision boundary sufficiently smooth with respect to the data structure, composed by labelled and unlabelled information to improve the robustness and precision of the model. The motivation to use semi-supervised learning in this problem is to analyse whether the inclusion of unlabelled data helps to the construction of the decision model when the amount of labelled data is low (compared to the number of unlabelled examples). In this case, a light linear semi-supervised SVM algorithm is used [27].

g) Output of the classifier: The output of the proposed system can be used for different treatments: a binary apply/not apply herbicides to weed infested field section or the application of different herbicides (e.g. to control broadleaved, grass or resistant weeds). These treatment maps will be given afterwards to the treatment equipment software to properly apply the herbicides. Moreover, the detection of crops could also be useful for plant counting or to position the equipment according to crop rows. For more information concerning the system proposed for weed mapping refer to [15].

III. Experiments

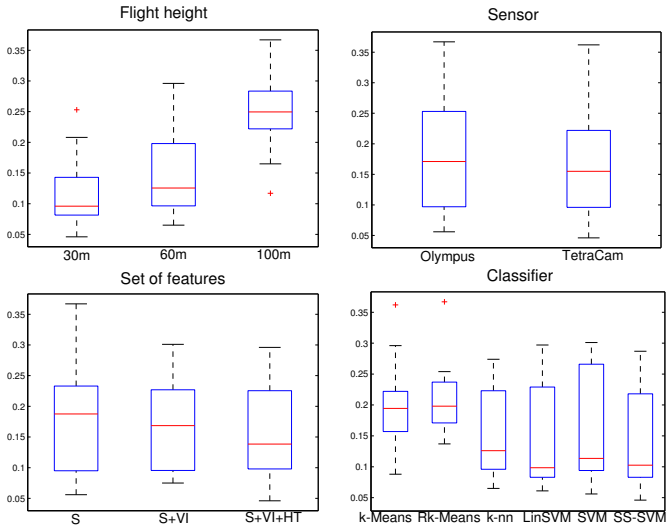

Fig. 2. Boxplot of the MAE results obtained for the different factors taken into account in this study.

Two different experiments are considered in this section in order to verify and study the potential of the methods tested. The former experiment intends to test the difference in mean performance among the different chosen paradigms and sets of input features, as well as to analyse the impact of the flight height and sensor used. The latter experiment tests whether it is possible to use previously trained models to generalise well on images with similar lighting and spectral settings and whether the performance of the method could be improved by choosing more fine-tuned models (e.g. train the model at a subplot at 30m but use it at 100m).

A. Methods tested

As stated before, different learning paradigms are tested in these experiments to perform a thorough analysis of the

techniques which show the best and most robust results for the purpose considered in this paper. To do so, we use two centroid-guided unsupervised algorithms (k-means and Repeated k-means or Rk-means), a semisupervised SVM method (SS-SVM), and three supervised learning techniques (k-nn, linear SVM version or LinSVM and kernel SVM). The selection of these methodologies is not arbitrary: k-means is one of the pioneer and most widely-used methods for unsupervised learning, Rk-means is included in order to analyse the impact of the choice of the centroids, the k-nn method is chosen because the prediction is also based on a distance relation (as for the k-means) methods and finally, the SVM methodology is one of the most successful machine learning methods (both the linear and nonlinear versions were included to analyse the difficulty in separating the data).

For optimising the parameters of the chosen methods, a 5-fold cross-validation procedure is used. For the k-nn method the number of nearest neighbours is cross-validated among the following values: $\{1, 3, 5, 7\}$. For the LinSVM and SVM techniques the cost parameter is chosen from one of these values: $\{10^{-2}, 10^{-1}, 10^{0}, 10^{1}, 10^{2}\}$ (and so is the kernel parameter associated to SVM). For the regularisation parameters of SS-SVM the following values are used: $\{10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 10^{0}\}$. The evaluation metric for cross-validation in this case is the well-known accuracy.

For HT, the number of lines to detect is set to 200 (as this number is only orientative). The rest of the parameters (i.e. the minimum length of the lines and the gap between lines) are chosen by cross-validation from the following set: $\{10, 30, 50\}$ (using the evaluation metric defined in [15]).

B. Performance evaluation

For evaluating the results, 32 ground truth frames are considered. For each of these frames, we compute the approximate percentage of soil, crop and weed pixels and these are compared to the ones obtained by our system.

To evaluate the performance of the different methodologies, the Mean Average Error (MAE metric) is used. This measure computes the general mean deviation from the expected percentages, computed differently for each class $\mathcal{C}_j$:

$$MAE_{\mathcal{C}_j} = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{100}|pp_{ji} - tp_{ji}|, \qquad (1)$$

where $\mathcal{C}_j \in \{c, w, s\}$ (referring to the crop, weed and soil classes, respectively), N is the number of frames of ground truth (32 in this case), pp_{ji} corresponds to the predicted percentage for class $\mathcal{C}_j$ and true ground frame with index i and tp_{ji} to the analogue true percentage. Then, we average MAE for the three classes and the final error measure is:

$$MAE = \frac{(MAE_c + MAE_w + MAE_s)}{3}, \qquad (2)$$

this coefficient ranging from 0 to 1 (being 0 the ideal value).

C. First experiment: Study of the impact of different factors

This experiment explores the difference in performance with respect to four factors:

- Input features: input features include the original spectral information, a vegetation index (VI, ExG for Olympus and NDVI for TetraCam) and the proposed feature, to check whether the proposal of a combined use of spectral information, VI, and the HT information reduces the prediction errors.
- Classifiers: six methods are compared. The objective is to check whether semi-supervised methods could lead to a more precise and robust prediction.
- Flight height: images taken at 30, 60 and 100 meters are used.
- Camera: two sensors (visible Olympus and multispectral TetraCam) are compared.

Table I presents the performance results obtained for the different set of input features, classifiers, flight altitudes and sensors. Different conclusions can be drawn from this table: Firstly, analysing the results of k-means and Rk-means, it can be seen that, in general, the unsupervised approach does not yield satisfactory results (competitive results are obtained in specific cases but the performance of these methods is not robust). Moreover, repeating the cluster computation (i.e. using Rk-means instead of k-means) does not result in more consistent results. Secondly, concerning the supervised and semi-supervised approaches, all of them result in a similar performance, although SS-SVM presents the best results in mean, meaning that using unsupervised information complements satisfactorily the model and helps to stabilise it (to see this analyse the mean per method). Concerning SVMs, it can be seen that the linear version yields better results in mean than the kernel version, which means that the problem is generally linearly separable and the kernel mapping is not necessary. Furthermore, the kernel version is also seen to perform better when using the Olympus sensor, which could indicate that the feature proposed and the VI is more appropriate to be used with multispectral sensors (such as the TetraCam), because it leads to linearly separable decision regions. Thirdly, with relation to the flight height, it is shown that the highest resolution (i.e. 30m) presents the best result, followed by 60m and 100m. The performance gap between 30m and 60m could be acceptable depending on the application, but up to our knowledge, the results at 100m are not satisfactory in general. Finally, regarding the sensors, the Olympus camera produces better results when taking into account only the spectral information (or spectral plus VI information), but not when using our proposed feature. The TetraCam camera, however, produces outstanding results under this framework, despite the lower spatial resolution of this camera when compared to the visible one. In this case, therefore, the improvement is due to the use of the NIR band, which helps to discriminate the classes of the problem in a better manner. Note that the best results are obtained using this camera, the complete set of features, the SS-SVM method and a flight height of 30m (with

TABLE I

MEAN MAE RESULTS OBTAINED FOR THE FOUR DIFFERENT FACTORS CONSIDERED THE BEST METHOD FOR EACH COMBINATION OF FEATURES, FLIGHT HEIGHT AND SENSOR IS HIGHLIGHTED IN BOLD FACE, AND THE SECOND ONE IN ITALICS.

MAE	Spectral Information						Spectral Information + VI						Spectral Information + VI + Hough Trans.						
	Olympus			TetraCam			Olympus			TetraCam			Olympus			TetraCam			
Methods	30	60	100	30	60	100	30	60	100	30	60	100	30	60	100	30	60	100	Mean
k-means	0.181	0.198	0.293	0.137	0.222	0.362	0.166	0.198	0.171	0.155	0.196	0.257	0.208	0.296	**0.117**	0.088	0.157	*0.193*	0.200
Rk-means	0.194	0.198	0.367	0.137	0.222	0.237	0.172	0.198	0.171	0.155	0.190	0.252	0.253	0.254	*0.230*	0.149	0.214	**0.165**	0.209
k-nn	0.065	*0.093*	**0.274**	0.096	0.132	**0.223**	0.082	**0.095**	**0.234**	0.096	0.130	*0.223*	0.121	0.128	0.247	0.096	0.124	0.218	0.149
LinSVM	*0.061*	**0.085**	0.297	**0.083**	**0.103**	0.229	0.079	0.099	0.291	**0.083**	*0.096*	0.231	*0.098*	**0.124**	0.266	*0.064*	*0.088*	0.221	*0.144*
SVM	**0.056**	0.098	0.292	*0.094*	0.204	0.252	**0.075**	*0.097*	*0.266*	*0.093*	0.270	0.301	*0.098*	*0.125*	0.265	0.102	**0.065**	0.288	0.169
SS-SVM	0.063	**0.085**	*0.287*	0.098	*0.107*	*0.227*	*0.081*	0.121	0.280	**0.083**	**0.094**	**0.218**	**0.089**	0.126	*0.253*	**0.046**	**0.065**	0.203	**0.140**
Mean	0.103	0.126	0.301	0.108	0.165	0.255	0.109	0.135	0.236	0.111	0.163	0.247	0.144	0.176	0.230	**0.091**	**0.119**	**0.215**	

a relatively low performance loss when performing the flight at 60m).

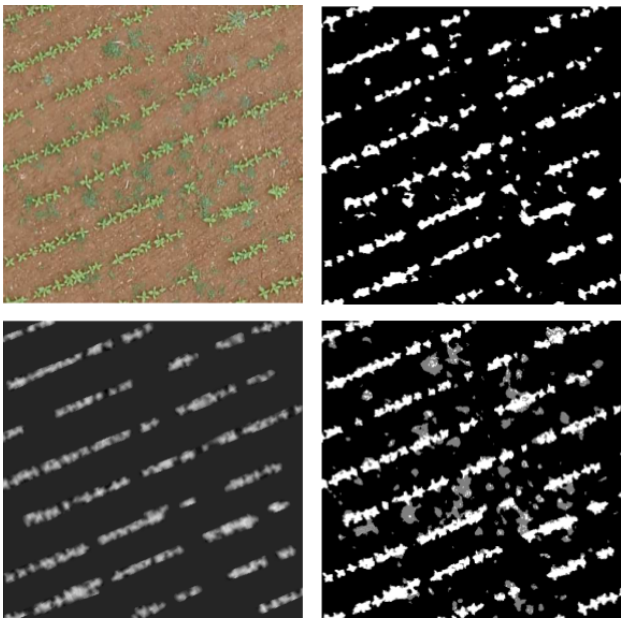

Fig. 3. Representation of the complete set of features used and the prediction produced for an image captured at 30m using the Olympus camera, where a large quantity of weed canopies can be appreciated. The first figure (top-left) represents the spectral information, the second figure (top-right) the VI (ExG), the third figure (bottom-left) the proposed feature which combines detected crop rows and VI and finally the last figure (bottom-right) represents the output of the algorithm. This results have been obtained using the SS-SVM method.

Fig. 2 shows the performance boxplots for each of the four factors considered in this paper. The results presented in the previous paragraph can also be verified using this figure.

Fig. 3 shows the representation of the complete set of features used for a portion of the experimental field: 1) spectral information (in this case, to help the reader visualise the difference between the two sensors we include the case of the visible Olympus camera); 2) VI (ExG in this case, where white pixels represent vegetation); and 3) our proposed feature which mixes the data from the VI and detected crop rows. As can be seen, this feature incorporates the VI concerning only sunflower in the crop row, which helps distinguish the crop and weeds. Furthermore, the HT detects sowing fails, which is an important characteristic. The figure at the bottom-

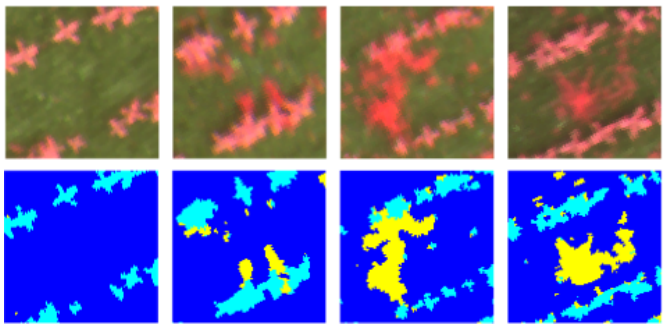

Fig. 4. Results of the SS-SVM method (S + VI + HT). The plots at the top of the figure show the image at 30m using the TetraCam sensor. The plots at the bottom show the prediction obtained for the four true ground patches where each colour represents a class (dark blue represents soil, cyan represents the crop and yellow represents weeds).

right part of the figure represents the output of the SS-SVM method trained in this region. Each colour represents a class: black (soil), white (crop) and gray (weeds). It can be seen, that although this subimage presents a large concentration of weeds, all of them are detected accurately, even small ones or those that lie within crop rows.

Finally, Fig. 4 shows the prediction provided for four of the true ground patches using the SS-SVM method at 30m. From this image it can be confirmed that weeds close to the crop row are detected, although some pixels are misclassified (mostly pixels belonging to crops).

D. Second experiment: Generalising with previously computed models

The use of previously trained models could highly reduce the required computational time and alleviate the process of weed mapping. Different hypotheses arise in this topic: Is it possible to train models at a height and generalise with success to other heights? Is it possible to apply the model learnt at an experimental field to a different field with the same crop? Note that this is a rather complex topic, given that the experimental fields could have very different characteristics (e.g. lighting or soil characteristics).

For this last experiment, we test whether the models trained (at a specific flight height) generalise well on images at different heights. This could be useful to reduce the computational time and alleviate the whole process. To do so, we considered

all of the different models trained before (recall that for each image there was as many models as divisions made for the experimental field). Each model is used with the whole image, thus obtaining as many results as models. The mean of these results can be seen in TABLE II, where one can appreciate e.g. that when using the image at 30m for the Olympus camera but with the model trained at 100m the performance deteriorates (to see this, analyse TABLE I. In this case, the base performance is 0.046, as opposed to the obtained results: 0.102 and 0.170 for 60m and 100m respectively). On the contrary, when using an image for the generalisation at a higher height (e.g. 100m) but using a model trained at a lower height (30m for TetraCam) the performance is seen to improve significantly (from 0.203 to 0.124). This result is important, because it means that learning a model at a lower height (even using a relatively small section of the field) and then using this model at a higher height (which would in this case cover all of the experimental field) is feasible and leads to promising results. As said before, this is crucial to shorten flight times and the amount of images to be mosaicked. For the case of the TetraCam sensor at 60m, the same conclusions are not applicable, which motivates further improvement of these models.

TABLE II
MEAN RESULTS OBTAINED WHEN GENERALISING PREVIOUSLY TRAINED MODELS WITH DATA ASSOCIATED TO OTHER IMAGES (SAME LOCATION BUT DIFFERENT HEIGHT). THE "TRAIN" ROW INDICATES THE HEIGHT AT WHICH THE MODEL WAS TRAINED. THE "TEST" ROW REFERS TO THE HEIGHT OF THE IMAGE THAT IS USED FOR GENERALISING THE MODEL.

Train	30m		60m		100m	
Test	60m	100m	30m	100m	30m	60m
MAE Olympus	**0.117**	0.148	**0.102**	**0.129**	**0.170**	**0.146**
MAE TetraCam	0.124	**0.124**	0.331	0.402	0.189	0.193

IV. CONCLUSIONS

This paper presents a novel computational approach for weed monitoring in sunflower using unmanned aerial vehicles, with the main purpose of designing site-specific weed control treatments in early post-emergence. One of the main difficulties for this objective is the spectral similitude between crop and weeds pixels, thus needing a more sophisticated method for providing an accurate prediction. This system makes use of the Hough transform (HT) for detecting crop rows in order to complement this spectral data and increase its robustness and precision. Several machine learning paradigms are explored (unsupervised, semi-supervised and supervised) with the purpose of minimise the intervention of the final user. The results show that both supervised and semi-supervised methods perform considerably well (even when only a very small subset of labelled training patterns is available). Furthermore, this work validates some hypotheses in the literature concerning factors such as the flight height and the sensor. Finally, an additional experiment shows that it is possible to train a model at a lower height in a subplot of the experimental field and apply it successfully to the rest of the field using a flight height of 100m, which could simplify the process and the computational time to apply this system in the practice.

Concerning future work, firstly, although the classical HT was initially designed to detect lines, this technique has been also extended to arbitrary shapes, this improvement could be used to allow more degrees of freedom in the row detection. Other crops (apart from sunflower) could be considered, such as wheat (which is a narrow-row crop) to explore the applicability of the proposal to other scenarios. The use of object-based analysis would be of special interest, to avoid the well-known salt-and-pepper effect.Finally, the applicability of a model (trained in a specific experimental field) to a different field (e.g. other location) should be analysed, as it could highly reduce the required computational time. However, it is an ambitious objective, since experimental fields could have very different characteristics. Because of this, the paradigm of incremental learning could be used. Under this setting, a learnt model could be slightly modified to better fit new data at the expense of a low computational load.

ACKNOWLEDGMENT

This work was financed by the Recupera 2020 Project (an agreement between CSIC and Spanish MINECO, EU-FEDER funds) and partially subsidized by the TIN2014-54583-C2-1-R project of the Spanish Ministerial Commission of Science and Technology (MICYT), FEDER funds and the P11-TIC-7508 project of the "Junta de Andalucía" (Spain). Research of Mr. Torres-Sánchez and Dr. Peña was financed by the FPI and Ramón y Cajal Programs, respectively.

REFERENCES

[1] M. Jurado-Expósito, F. López-Granados, L. García-Torres, A. García-Ferrer, M. Sánchez de la Orden, and S. Atenciano, "Multi-species weed spatial variability and site-specific management maps in cultivated sunflower," *Weed science*, vol. 51, no. 3, pp. 319–328, 2003.

[2] "European Commission, Directive 2009/128/EC of the European Parliament and of the Council establishing a framework for Community action to achieve the sustainable use of pesticides," 2009. [Online]. Available: http://faolex.fao.org/docs/pdf/eur113943.pdf

[3] G. Pajares, "Overview and current status of remote sensing applications based on unmanned aerial vehicles (uavs)," *Photogrammetric Engineering & Remote Sensing*, vol. 81, no. 4, pp. 281–330, 2015.

[4] B. Somers, S. Delalieux, W. Verstraeten, J. Verbesselt, S. Lhermitte, and P. Coppin, "Magnitude- and shape-related feature integration in hyperspectral mixture analysis to monitor weeds in citrus orchards," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 11, pp. 3630–3642, 2009.

[5] P. Gutiérrez, F. López-Granados, J. M. Peña-Barragán, M. Jurado-Expósito, M. T. Gómez-Casero, and C. Hervás-Martínez, "Mapping sunflower yield as affected by Ridolfia segetum patches and elevation by applying Evolutionary Product Unit Neural Networks to remote sensed data," *Computers and Electronics in Agriculture*, vol. 60, no. 2, pp. 122–132, 2008.

[6] F. López-Granados, "Weed detection for site-specific weed management: mapping and real-time approaches," *Weed Research*, vol. 51, no. 1, pp. 1–11, 2011.

[7] T. Moranduzzo and F. Melgani, "Automatic car counting method for unmanned aerial vehicle images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 3, pp. 1635–1647, March 2014.

[8] A. Lucieer, D. Turner, D. H. King, and S. A. Robinson, "Using an unmanned aerial vehicle (UAV) to capture micro-topography of antarctic moss beds," *International Journal of Applied Earth Observation and Geoinformation*, vol. 27, pp. 53–62, 2014.

[9] J. M. Peña, J. Torres-Sánchez, A. I. de Castro, M. Kelly, and F. López-Granados, "Weed mapping in early-season maize fields using object-based analysis of unmanned aerial vehicle (UAV) images," *PloS one*, vol. 8, no. 10, p. e77151, 2013.

[10] J. Torres-Sánchez, F. López-Granados, A. I. De Castro, and J. M. Peña, "Configuration and specifications of an unmanned aerial vehicle (UAV) for early site specific weed management," *PloS one*, vol. 8, no. 3, p. e58210, 2013.

[11] J. Torres-Sánchez, J. M. Peña, A. I. de Castro, and F. López-Granados, "Multi-temporal mapping of the vegetation fraction in early-season wheat fields using images from UAV," *Computers and Electronics in Agriculture*, vol. 103, pp. 104–113, 2014.

[12] E. I. Papageorgiou, A. T. Markinos, and T. A. Gemtos, "Fuzzy cognitive map based approach for predicting yield in cotton crop production as a basis for decision support system in precision agriculture application." *Applied Soft Computing*, vol. 11, no. 4, pp. 3643–3657, 2011.

[13] Y. Zheng, Q. Song, and S. Chen, "Multiobjective fireworks optimization for variable-rate fertilization in oil crop production." *Applied Soft Computing*, vol. 13, no. 11, pp. 4253–4263, 2013.

[14] C. Hung, Z. Xu, and S. Sukkarieh, "Feature learning based approach for weed classification using high resolution aerial images from a digital camera mounted on a uav," *Remote Sensing*, vol. 6, no. 12, p. 12037, 2014.

[15] M. Pérez-Ortiz, J. P. na, P. Gutiérrez, J. Torres-Sánchez, C. Hervás-Martínez, and F. López-Granados, "A semi-supervised system for weed mapping in sunflower crops using unmanned aerial vehicles and a crop row detection method," *Applied Soft Computing*, vol. 37, pp. 533 – 544, 2015.

[16] M. Pérez-Ruiz, P. G. de Santos, A. Ribeiro, C. Fernandez-Quintanilla, A. Peruzzi, M. Vieri, S. Tomic, and J. Agüera, "Highlights and preliminary results for autonomous crop protection," *Computers and Electronics in Agriculture*, vol. 110, pp. 150 – 161, 2015.

[17] 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.

[18] X. P. Burgos-Artizzu, A. Ribeiro, M. Guijarro, and G. Pajares, "Real-time image processing for crop/weed discrimination in maize fields," *Computers and Electronics in Agriculture*, vol. 75, no. 2, pp. 337–346, 2011.

[19] R. O. Duda and P. E. Hart, "Use of the hough transformation to detect lines and curves in pictures," *Commun. ACM*, vol. 15, no. 1, pp. 11–15, Jan. 1972.

[20] L. Comba, P. Gay, J. Primicerio, and D. R. Aimonino, "Vineyard detection from unmanned aerial systems images," *Computers and Electronics in Agriculture*, vol. 114, pp. 78 – 87, 2015.

[21] D. Gómez-Candón, A. De Castro, and F. López-Granados, "Assessing the accuracy of mosaics from unmanned aerial vehicle (UAV) imagery for precision agriculture purposes in wheat," *Precision Agriculture*, vol. 15, no. 1, pp. 44–56, 2014.

[22] A. A. Gitelson, Y. J. Kaufman, R. Stark, and D. Rundquist, "Novel algorithms for remote estimation of vegetation fraction," *Remote Sensing of Environment*, vol. 80, no. 1, pp. 76–87, 2002.

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

[24] N. Otsu, "A Threshold Selection Method from Gray-level Histograms," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[25] A. K. Jain, "Data clustering: 50 years beyond k-means," *Pattern Recognition Letters*, vol. 31, no. 8, pp. 651–666, Jun. 2010.

[26] C.-W. Hsu and C.-J. Lin, "A comparison of methods for multiclass support vector machines," *IEEE Transactions on Neural Networks*, vol. 13, no. 2, pp. 415–425, Mar. 2002.

[27] O. Chapelle, B. Schlkopf, and A. Zien, *Semi-Supervised Learning*, 1st ed. The MIT Press, 2010.