Vegetation segmentation in cornfield images using Bag of Words

Yerania Campos[1]*, Erik Rodner[2], Joachim Denzler[2], Humberto Sossa[3], and Gonzalo Pajares[1]

[1] Department of Software Engineering and Artificial Intelligence, Faculty of Informatics, Complutense University, 28040 Madrid, Spain.
[2] Computer Vision Group, Friedrich Schiller University Jena, 07743 Jena, Germany.
[3] Instituto Politécnico Nacional-CIC, Av. Juan de Dios Batis, S/N, Col. Nva. Industrial Vallejo, 07738 México D.F., México

Abstract. We provide an alternative methodology for vegetation segmentation in cornfield images. The process includes two main steps, which makes the main contribution of this approach: a) a low-level segmentation and b) a class label assignment using Bag of Words (BoW) representation in conjunction with a supervised learning framework. The experimental results show our proposal is adequate to extract green plants in images of maize fields. As a classification task, an accuracy value of 95.3 percent has been achieved, it is similar to the values reported in the current literature.

Keywords: Bag-of-words, machine learning, colour vegetation indices, green detection.

1 Introduction

The rapid development of new technologies is changing the manner in which the food is produced. Advances in electronics, artificial intelligence, machine vision and other technologies have been integrated in the design and development of autonomous agricultural vehicles (AAV) capable of performing a wide range of activities in the agricultural industry. The main benefits of AAV are: save time and effort, major quality of food, environment protection and operational cost reduction [1]. Autonomous vehicles are equipped with vision-based sensors, which provide all the data needed to develop activities of localization, mapping, path planning and obstacle avoidance.

In a AAV, segmentation of vegetation is a critical step towards the development of different activities in the crop field such as counting plants for germination monitoring, detecting weeds for early season site specific weed management, or nutrient application. This task is usually performed from images acquired by the vision system and must therefore be considered in the design of agricultural vehicles. In short, a good algorithm to split an image into foreground (maize/plants) and background (soil, irrigation pipes, etc.) is highly demanded to improve the performance of the activities carried out by the AAV.

In this paper, we provide a method for vegetation segmentation in agricultural images (AI), making the main finding. The procedure includes a low level segmentation process to get regions of interest (ROIs), these are subsequently evaluated using a classifier model to determine which ROIs do not belong to vegetation. Additionally, we provide a dataset composed of maize field images and their corresponding labelled images which were made by inspection and carefully hand painted. Images were captured with a single camera mounted on board a tractor, which is part of the fleet in the RHEA project [2].

This paper is organized as follow: Section 2 provides a revision of the state of the art, Section 3 explains our work, Section 4 shows the testing we conducted to prove its efficiency, and Section 5 gives the conclusions.

* Corresponding author, Tel.: +34 1394, ext. 4375, 7546.

2 Literature review

The first attempts to develop AAV were reported in the 1960s [3], new proposals have been introduced to increase the effectiveness of the navigation systems in agricultural vehicles; they are summarized by Mousazadeh [4], Vibhute et. al [5] and Saxena [6]. This work is limited to dealing with outdoor scenes where vegetation segmentation is the first crucial step within a complex process. In this context, Table 1 provides an overview of recent proposals. Definitions of abbreviations used on this table can be consulted in Table 2, they refer to the colour vegetation indices (CVIs).

Reference	Application	Methodology	Performance/Remarks
*Haug et al. [7]	Plant classification crop/weed. Images captured on a carrot farm.	Segmentation of vegetation from soil is obtained from NDVI. To discriminate between crops and weeds machine learning is applied.	Classification accuracy 93.8%.
Hlaing and Khaing [8]	Weed and crop segmentation and classification.	Segmentation is achieved combining ExG and area thresholding algorithms.	Error rate 33.3% for misclassified plants.
Tewari et al. [9]	Herbicide applicator for weed control.	The weed percentage in an image (total number of green pixels / size image) is computed to determine the herbicide amount.	Weeding efficiency 90%
Wei et al. [10]	Fruit picking robot.	Otsu adaptive threshold algorithm and features in OTA colour space are used for fruit detection.	Fruit object extraction 95%.
Choi et al. [11]	Line extraction in paddy fields.	Preocess includes: NIR imge, gray colour, median filter, Otsu and blob noise elimination	Green segmentation performance is not provided.
*+Torres et al. [12]	Vegetation detection in herbaceous crops.	Automatic thresholding algorithm based on Otsus method.	Error between 0% and 10%. Classification rate is affected by segmentation shape and compactness parameters.
Yang [13]	Greenness identification in cornfield	Segmentation is achieved from Hue components in HSV colour space and ExG metric.	Recognition accuracy 95%. Sensible to change illumination
*Jiang et al. [14]	Crop row detection	Rows detection from binary image obtained from $Gray_1$ metric.	Crop row detection accuracy 93%. Depends on vegetation segmentation.
Meng et al. [15]	System to Inter-row weeding in maize crop field.	H component (HIS colour space) is segmented considering Hue values in the range of [120,160]. From segmentation a scanning method is applied for crop lines detection.	Average error below 2.7 cm.
*Guijarro et al. [16]	Greenness segmentation.	Combining vegetation indices (greenness) and wavelets (texture).	Useful when the quality of imaging greenness is low. Precision 92.09%.
Balasubramaniam and Ananthi [17]	Segment incomplete nutrient-deficient crop images	Fuzzy C-means colour clustering	High accuracy in extraction of deficiency region.
*Kazmi et al. [18]	Thistle detection in sugar beet fields.	Detection based on CVI, Mahalanobis distance and Linear discriminant analysis (LDA).	Accuracy up to 97%.
*Kazmi et al. [19]	Weed detection sugar beet and creeping thistle images.	BoW scheme with KNN and SVM classifiers.	Accuracy of 99% in scanned leaf images. Outdoor images were not considered.
Ye et al. [20]	Crop segmentation	Adoption of Markov random field to provide belief information from crop extraction.	92.29% accuracy, even under strong illumination changes.
Cheng et al. [21]	Rice and weed discrimination.	Harris corner detection and machine learning (decision tree).	Precision of 98.8% to distinguish weeds from rice plants.
Moorthy et al. [22]	Vegetation segmentation.	Nave Bayesian model using features from RGB and HSV colour spaces.	87% on sugar beet and maize plants.
Santos et al. [23]	3D plant modelling for plant phenotyping (stereo vision)	3D Point cloud segmented by spectral clustering.	Experiments with maize were unsuccessful.
Lonescu et al. [24]	Biomass type identification.	Texture features, local texton dissimilarity and BoW representation.	Accuracy 90%. Available for mobile devices.

*Otsu as threshold strategy.
+ Unmanned aerial vehicle (UAV).
The normalized difference vegetation index (NDVI) value is obtained from a multi-spectral camera.

Table 1. The current state-of-the-art in vegetation detection for agricultural applications.

Abbreviation	Expression
Normalization	$R_n = R^*/(R^* + G^* + B^*)$, $G_n = G^*/(R^* + G^* + B^*)$, $B_n = B^*/(R^* + G^* + B^*)$, $R^* = R/max(R)$, $G^* = G/max(G)$, $B^* = B/max(B)$
Gray	$0.2898 * R_n + 0.5870 * G_n + 0.1140 * B_n$
Gray_1 [14]	$1.262 * G_n - 0.884 * R_n - 0.311 * B_n$
ExG [25]	$2 * G_n - R_n - B_n$
ExR [26]	$1.4 * R_n - G_n$
CIVE [27]	$0.441R_n - 0.811G_n + 0.385B_n - 18.78$
ExGR [28]	$ExG - ExR$
NDI [29]	$(G_n - B_n)/(G_n + B_n)$
GB [25]	$G_n - B_n$
RBI [30]	$(R_n - B_n)/(R_n + B_n)$
ERI [30]	$(R_n - G_n) * (R_n - B_n)$
EGI [30]	$(G_n - R_n) * (G_n - B_n)$
EBI [30]	$(B_n - G_n) * (B_n - R_n)$
VEG	$G_n * R_n^a * B_n^{(a-1)}$

Table 2. Colour channels and colour vegetation indices.

3 Proposed methodology

Bag of Words. It was initially introduced for text analysis [31], the success of this representation is based on the high discriminative power of some words and the redundancy of language in general. Subsequently, this technique was adapted in applications of computer vision [3, 18, 24, 32], where a *visual word* is a sparse vector of occurrence counts of a *visual vocabulary* of local image features. The visual vocabulary is usually obtained by quantifying the image features into visual words.

The process to determine whether a ROI is vegetation by using the BoW representation consists of two stages: training and testing, Fig. 1. On the first, a classifier model for three classes is built with features extracted from the ROIs. The model is used to predict the label of a new ROI into the second stage. The three classes involved are; vegetation (v), soil (s) and one more identified as others (o). The last class includes elements that did not identify with the two predominant classes.

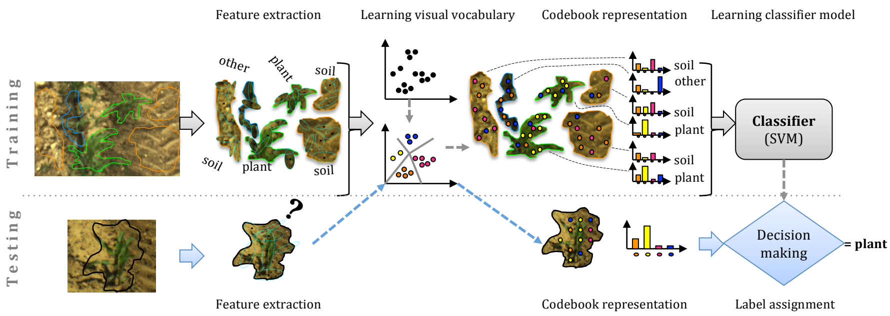


Fig. 1. Bag of Words scheme for agricultural images.

Feature selection plays an important role in the performance of the classifier function. This topic has been widely discussed in the literature, the researchers conclude that feature selection depends on the nature of problem [33]. Our work focuses on finding an appropriate set of features for characterization of vegetation. Because of this, descriptors proposed recently by Kazim et al. [19], mix of different CVIs, are used for vegetation characterization, Table 3. Additionally, the local SIFT [34] and SURF[35] descriptors are also included for analysis.

Descriptor	Colour vegetation indices
CVI_2	ExG, GB
CVI_4	ExG, CIVE, GB, ERI
CVI_9	ExR, ExGR, NDI, GB, RBI, ERI, EGI, R_n, G_n
CVI_{14}	ExG, CIVE, ExR, ExGR, NDI, GB, RBI, ERI, EGI, EBI, R_n, G_n, B_n, Gray

Table 3. Composition of the CVI descriptors [19]. See Table 2 for the expressions of the indices.

3.1 Classification model

Consider a set of N interest regions $R = \{R_1, \dots, R_N\}$, each element is a set of pixels $R_i = \{r_1, \dots, r_m\}$, $\mid R_i \mid= m$. The number of pixels in each region is different. Also, the set of labels associated to each region $L = \{l_1, \dots, l_N\}, l_i \in \{v, s, o\}$ is given. Examples of ROIs and their associated labels can be seen in Fig. 1. R is split into two complementary sets: R_A and R_B $(R_A \cap R_B, \emptyset)$, $\mid R_A \mid= a$ and $\mid R_B \mid= b$. The same with the label set: L_A and L_B $(L_A \cap L_B, \emptyset)$, $\mid L_A \mid= a$ and $\mid L_B \mid= b$. R_A and L_A are used to train the classifier function while R_B and L_B are used for parameter estimation.

Training process. Input: R_A, L_A and the vocabulary size K. Output: classification function Ψ.

1. Feature extraction: Consider a region $R_i \in R_A$. For each pixel in R_i, a feature descriptor is computed: $F_i = F_i^1, \dots, F_i^m$, $F_i^j \in \Re^z$, z is the dimension of the descriptor. The same applies for all elements in R_A having as result a set of descriptors: $F_A = \{F_1, \dots, F_a\}$.
2. Visual vocabulary: Descriptors in F_A are used to train a clustering method to obtain K-centres, we apply k-means [36]. Each centre represents a visual word. The set of K-visual words is the visual vocabulary: $W = \{w_1, \dots, w_K\}$, $w_k \in \Re^z$. Also, from k-means, at each descriptor in F is associated the label of the nearest centre. For example, F_i is represented for $D_i = \{D_i^1, \dots, D_i^m\}, D_i^j \in \{1, \dots, K\}$ and the set of labelled features $D_A = \{D_1, \dots, D_a\}$.
3. Codebook: For each element in D_A, the frequency of each visual word is computed. The vector of counts is divided by the number of pixels in the ROI at which it belongs to in order to get a normalized vector. The frequency vectors are the codebooks: $CB_A = \{H_1, \dots, H_a\}$.
4. Classification function: CB_A, L_A and a method of cross validation [37], used to find the best parameter values, are processed during the learning process. The decision function chosen for classification is the one provided by support vector machines (SVM) with parameter c wich tells the SVM optimization how much misclassifying is allowed at each training [38].

Testing process. Input: R_B, L_B, W and Ψ. Output: Performance model.

5. Feature extraction: Apply Step 1 to get descriptors in R_B: $F_B = \{F_1, \dots, F_b\}$.
6. Visual words: At each descriptor in F_B is associated the label of the nearest cluster in W: $D_B = \{D_1, \dots, D_b\}$.
7. Codebook: Apply step 3 to get the codebooks in D_B: CB_B.
8. Class assignment: Ψ is used to predict the labels in CB_B: $L_B^* = \{l_1^*, \dots, l_b^*\}, l_i^* \in \{v, s, o\}$.
9. Performance model: The true labels (L_B) and the labels obtained in the previous step (L_B^*) are processed with the first expression in Table 4 to compute the accuracy value.

ID	Description	Expression
OSR	Overall success rate/Accuracy	$(n_{TP}+n_{TN})/(n_{TP}+n_{TN}+n_{FP}+n_{FN})$
TPR	True positive rate/Recall/Sensitivity	$n_{TP}/(n_{TP}+n_{FN})$
TNR	True negative rate/Specificity	$n_{TN}/(n_{TN}+n_{FP})$
PPV	Positive Predictive Value/Precision	$n_{TP}/(n_{TP}+n_{FP})$
NPV	Negative Predictive Value	$n_{TN}/(n_{TN}+n_{FN})$
F	F-measure	$(2*n_{TP})/(2*n_{TP}+n_{FN}+n_{FP})$

Table 4. Statistical measures for performance evaluation [39]; n_{TP}, n_{TN}, n_{FP} and n_{FN} represent the number of true positives, true negatives, false positives and false negative respectively.

3.2 Image vegetation segmentation

An image I_{rgb} is segmented by classifying each pixel as foreground or background with the help of the visual vocabulary W and the classifier function Ψ.

Image to Interest regions. Without knowledge of the image structure, the first step is to find nearly uniform regions - ROIs. The principle is that pixels in small regions tend to contain elements of the same class. Ideally, each ROI would contain a single class of elements; vegetation, soil or other. However, improvement of the labelling process, using the BoW and the learning strategy together, is not guaranteed. In short, each image pixel is assigned to a unique region: $IR = \{IR_1, \ldots, IR_p\}$, p is the number of ROIs in the image. To group pixels into multiple ROIs four algorithms were tested: K-means [36], Self-organization maps (SOM) [40], Fuzzy C-means (FCM) [41] and Over-segmentation (OS) [42].

Interest regions to vegetation detection. The IR set is processed following steps 5 through 8 above to get the label in each region: $L^*_{IR} = \{l^*_1, \ldots, l^*_p\}, l^*_i \in \{v, s, o\}$. At each pixel in I_{rgb} is assigned the label of the IR at which it belongs: $I_{lab}(x, y) = l^*_i$ if $I_{rgb}(x, y) \in IR_i$. The final vegetation segmentation is achieved with the expression 1.

$$I_{bin}(x,y) = \begin{cases} 1 & \text{if } I_{lab}(x,y) = v, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

4 Experimental results

4.1 Image dataset

A collection of $\Omega = 168$ images, which were acquired under different illumination conditions and different plant growth state, were selected and manually segmented. Really, the images in Ω are subimages of size 920×950 obtained from the original images with resolutions of 2336×1752, based on the camera system geometry [43]. Unique pixels only contain a main component (vegetation, soil or perhaps other unidentified component), so no mixed information can be considered as relevant in this regard.

From Ω, $\Omega_1 = 26$ images were used solely for building the classifier function and the remaining $\Omega_2 = 142$ for measuring the success in the segmentation process. Table 5 displays some representative colour images (first row) and their corresponding hand-labelled images (second row). It should be noted that the labelled images have three different classes; green to identify vegetation (v), light-brown for soil (s), and dark-brown for elements on the border between green plants and soil or any different item on the image (o). Manual segmentation on the vegetation borders is even difficult to carry out under the supervision of an expert. Moreover, we noted that the vegetation detection accuracy using a model with three classes (v, s, o) is greater than the accuracy achieved with a binary classifier (v, s). Under this scenario, the inclusion of the third class in the classifier design is justified while the segmentation performance is carried out from a binary image where foreground comes from pixels with label v as described in subsection 3.2.

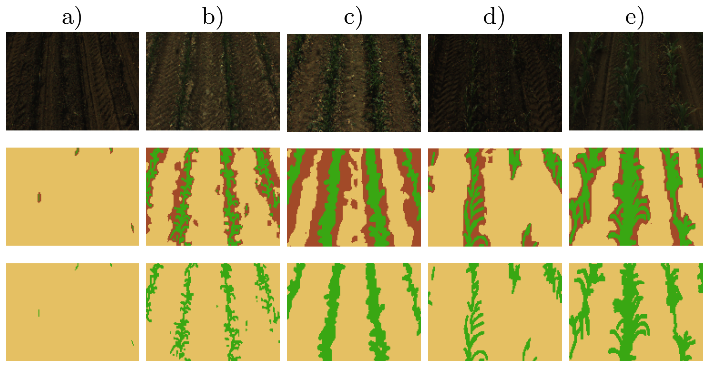


Table 5. First row: RGB images. Second row: hand-labelled images (v, s and o). Third row: Binary image, foreground (v) and background (joint s and o).

Classifier function estimation. Ω_1 is divided into two sets: 20 images for training and 6 images for accuracy evaluation. 1005 regions in the first set (346-vegetation, 171-soil and 488-others) and 739 regions in the second set (399-vegetation, 166-soil and 174-others). Two models were considered; linear and nonlinear. In both cases, the penalization parameter (c) was selected from the range $[0.1, 22]$ with intervals of 0.5. For the nonlinear model, a radial base function (RBF) with parameter γ was used as kernel. The searching consists on testing with pairwise (c, γ) and the one with the best cross-validation accuracy is picked - γ takes the same range values than c. This process was repeated several times changing the visual vocabulary size K; varying from 50 to 2000 with intervals of 50. Classifiers with highest performance are given in Table 6.

Descriptors		SVM-Linear			SVM-RBF		
Abbreviation	Size	W	c	$OSR(\%)$	W	(c, γ)	$OSR(\%)$
COM	1	1790	16.6	85.17	590	$21.1, 6.6$	91.50
CVI_2	2	2000	13.6	80.25	1490	$19.6, 21.1$	93.83
CVI_4	4	1400	9.6	78.49	1970	$21.6, 20.6$	95.31
CVI_9	9	1900	17.6	82.31	1670	$14.1, 21.6$	94.65
CVI_{14}	14	1950	17.1	81.20	1490	$17.1, 20.6$	94.84
SIFT	128	1550	21.6	68.12	1650	$19.1, 21.6$	90.99
SURF	64	1650	21.1	66.68	1950	$18.1, 18.1$	90.38

Table 6. Accuracy (%) of the classifier model for three classes (v, s, o).

From Table 6; the SVM-RBF has the best performance. The highest rates were achieved with the descriptors proposed by Kazim et al. [19]. They reported an accuracy of 97% with CVI_{14} to detect creeping thistle. We have similar rates for maize field images, the highest performance is 95.31% with CVI_4 and parameter values of $(c, \gamma) = (21.6, 20.6)$ and $K = 1970$.

Extraction of ROIs. The quality segmentation of 50 images, randomly selected from Ω_2, was used as criteria to select the partitioning method. The segmentation was carried out with a linear classifier model with COM as feature descriptor. The qualitative performance is summarized in Table 7; the average values of OSR and TPR are similar in all cases. The parameter estimation of each method was made as follow: For KM and FCM, the number of clusters was selected from $\{5, 10, 20, 30, 40, 50, 60\}$, KM has shown good performance with 30 clusters, while FCM works better with 10. For SOM, matlab default parameters were used; the row vector of dimension sizes ($[8, 8]$), the number of training steps for initial covering of the input space (100), the initial neighbourhood size (3), a hexagonal layer topology function and the link distance function were used to find the distances between the layer's neurons. In the case of OS, we set $(k, \sigma, min) = (0.1, 300, 100)$ to get small regions [42]. Visual results for a single image are displayed in Table 8; the partitions obtained with different methods are in the first row, while the true labelled image followed by the segmentation results and their respectively performance values $(TPR\%, OSR\%)$ appear in the second row.

Case	Measure	KM	SOM	FCM	OS
Average	OSR(%)	86.3	86.45	81.47	85.9
	TPR(%)	65.47	68.79	35.92	62.59
Best	OSR(%)	91.1	88.1	86.7	87.8
	TPR(%)	89.3	96.7	84.7	77.2
Worst	OSR(%)	74.8	80.8	61.1	68
	TPR(%)	50	60.9	37.61	42.8

Table 7. Performance evaluation of different partitioning methods.

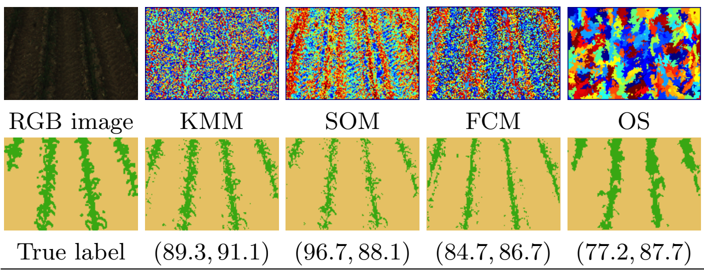


Table 8. First row: RGB image split into multiples regions with different algorithms. Second row: True labelled image followed by the segmentation results with their performance values $(TPR\%, OSR\%)$.

From Table 7, SOM was selected as partitioning method due to its TPR value is the highest.

4.2 Comparative analysis

We experimentally compare our algorithm with different methodologies proposed in precision agriculture and computer vision. A brief description of these methods is given below.

Precision Agriculture:

(i) Vegetation indices: A threshold value, selected from the CVIs, is used to get green pixels from an AI by thresholding. The resulting binary image is filtered to remove noise. Otsu's threshold technique is typically applied [18, 8, 13, 14, 44]. Our comparative analysis includes ExG, ExGR, Gray_1, CIVE, VEG, and COM indices and a 5×5 median filter for noise remotion [18].

(ii) Yang el at. [13]: The RGB image is transformed to HSV colour space. From Hue, the smallest (h_1) and largest (h_2) values are extracted. Channels R, G and B are processed separately according to expression 2. The ExG metric is computed with the new R^*, G^*, and B^*. The resulting colour image is segmented with the process described in (i).

$$A^*(x,y) = \begin{cases} 0 & \text{if } H(x,y) < h_1 \text{ or } H(x,y) > h_2, \text{ where H is a colour channel} \\ A(x,y) & \text{otherwise} \end{cases} \quad (2)$$

(iii) Hlaing and Khaing [8]: For each pixel in the RGB image, the absolute values of green minus red and green minus blue are calculated. If both of these distance values are greater than the threshold (T), the pixel is classified as plant. If none or only one is greater than T, the pixel is classified as background. T value is set to 20 as suggested by authors.

(iv) Tewari et al. [9]: For each pixel, when G colour intensity is greater than R and B colour intensity values simultaneously, the pixel is assumed to be green pixel. Otherwise, the pixel is assumed to be background.

Computer vision:

(v) Brust et al. [45]. A semantic segmentation process is carried out by using convolutional patch nerworks (CN). Authors reported good results in multi-classification task for urban scenes. As part of their contributions, they provide an open source CN library (CN24) which includes a pre-trained model able to identify multiple classes in urban scenes (building, window, sidewalk, car, road, vegetation, sky and unababeled). In our dataset, different CN architectures were tested for vegetation segmentation considering the three interest classes. The results obtained with different CN architectures and also with the pre-trained model were compared. The best results were achieved with the pre-trained model, these are reported in the comparative analysis.

(vi) Fröhlich et al. [46]. The semantic segmentation approach is based on the massive use of random decision forests (RDF) and the computation of several basic as well as high-level contextual features during learning (ICF).

The performance evaluation of methods above described was computed with images in Ω_2 and metrics in Table 4. The numerical results are provided in Table 9, as can be seen, the accuracy values with CVIs metrics (except COM) are over 83% (columns 2-7), the best performance is achieved with Tewari; 87.34% and 75.59% of OSR and TPR respectively. For a single image, the vegetation segmentation obtained with methods in Table 9 are shown in Table 10.

The results reported by Yang et al. and Hlaing and Khaing were computed with a dataset where plants are well defined (usually, one plant per image). On the first paper, an accuracy of 95% is reported, in the second case, authors do not provide vegetation segmentation results. In our dataset (many plants per image), the performance of these two proposals is poor, below 80%.

It is well known from the literature that convolutional networks have been shown high performance in various segmentation tasks. In our case, we tested different CN architectures with CN24 framework in our dataset and we could not find a configuration able to increase the performance

ID	ExG	ExGR	CIVE	VEG	COM	$Gray_1$	Yang	Hlaing	Tewari	ICF	CN24	**BoW**
OSR	86.95	85.83	83.76	85.72	75.95	87.71	79.40	70.64	87.34	75.29	82.64	86.11
TPR	71.67	83.51	74.02	67.27	53.22	74.16	60.25	13.65	75.59	53.90	71.43	73.24
TNR	89.58	85.20	85.14	88.35	90.58	89.40	84.08	76.60	89.84	82.04	84.89	90.39
PPV	66.38	40.17	44.25	60.16	73.36	63.32	44.28	4.23	60.31	54.16	39.71	58.60
NPV	93.00	97.47	95.74	92.99	78.61	94.00	91.08	90.44	92.49	83.95	93.72	89.51
F	67.10	53.05	50.27	61.29	55.59	67.67	44.94	3.24	64.96	52.35	43.32	61.60

Table 9. Performance evaluation for vegetation segmentation including our proposal. Metrics into rows, and methods into columns. Metrics can be consulted in Table 4.

value, even so, the accuracy is into the average of the accuracy values in the Table 9. ICF shows similar performance than CN24, it is important to mention that although it has low performance in vegetation detection, results can be relevant in the context of crop line detection given that vegetation on the crop line is preserved and well limited.

Finally, BoW representation has a OSR of 86.11% with a percentage of vegetation correctly identified of 73.24%. The rate of elements well classified is 90.39%, however the overlapping between green plants and background is 61.6%, similar values as such obtained with other proposals.

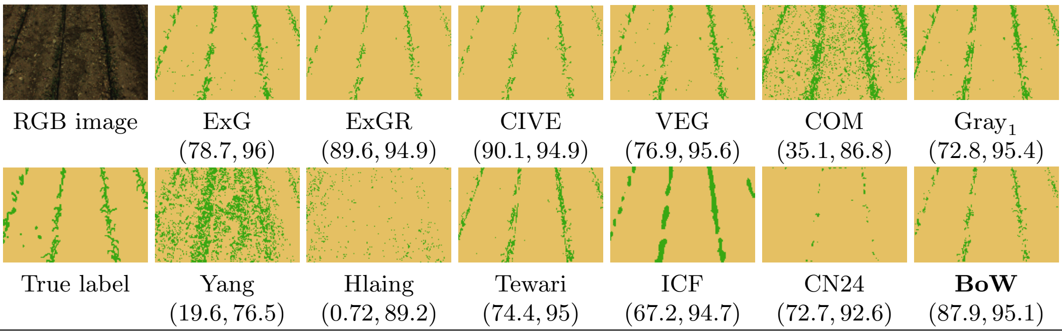


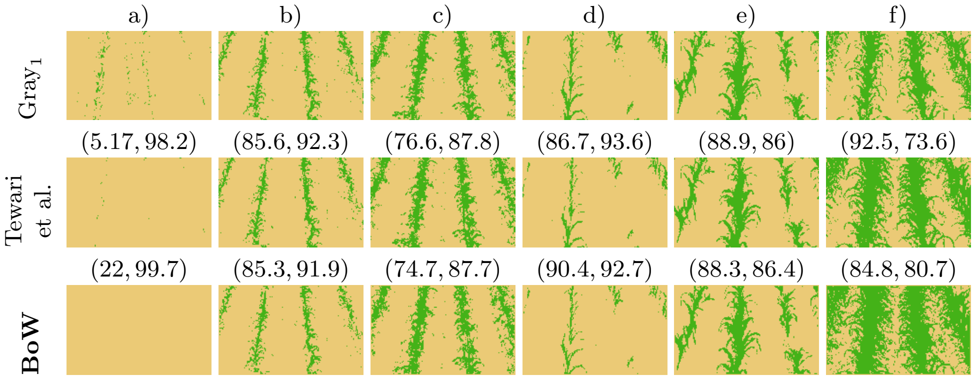
Table 10. Segmentation of an image with methodologies in Table 9. In brackets, the performance values ($TPR\%, OSR\%$).

In addition to the results above displayed, images in Table 5 were processed with; $Gray_1$, Tewari et al. and BoW, they have the best performance in Table 9, see Table 11.

On this section, a comparative analysis of different techniques for vegetation detection has been reported. Results presented were computed using the Image Processing Toolbox MATLAB 2013a for 64 bits under Windows 7 and Intel Core 2 CPU, 3 GHz, 4 GB RAM.

5 Conclusions

A wide range of computational vision tasks in agricultural applications could increase their performance if they start with an efficient vegetation segmentation process. On this paper, we presented an alternative method to identify vegetation in cornfield images, its performance (under different illumination conditions and growth stages) is similar to those reported in the current state-of-the-art. The accuracy achieved to discriminate between three classes is over 95%; however, segmentation method needs additional improvements. This is because although the classifier



	a)	b)	c)	d)	e)	f)
$Gray_1$	$(5.17, 98.2)$	$(85.6, 92.3)$	$(76.6, 87.8)$	$(86.7, 93.6)$	$(88.9, 86)$	$(92.5, 73.6)$
Tewari et al.	$(22, 99.7)$	$(85.3, 91.9)$	$(74.7, 87.7)$	$(90.4, 92.7)$	$(88.3, 86.4)$	$(84.8, 80.7)$
BoW	$(0, 99.8)$	$(87.3, 91.8)$	$(72.8, 87.2)$	$(92.1, 92.1)$	$(87.6, 86.5)$	$(78, 77.8)$

Table 11. Segmentation results under differents scenarios. In Table 5, the RGB images and their corresponding true labelled image (image per column). Performance values in brackets $(TPR\%, OSR\%)$.

achieves good performance, the segmentation algorithm depends on the method used to get the ROIs of the image. As future work, we suggest the use of probabilistic models [47] in order to improve the image segmentation results. Another possible future line of research is the deep analysis of results obtained with IFC method, segmented images are promising for crop line detection. To conclude, a set of 168 images and their corresponding handmade-labelled images are publicly available (https://www.fdi.ucm.es/profesor/pajares/ACIVS/), they are part of the contributions of this work. The dataset can be useful for performance evaluation on future researches.

Acknowledgments

The first author acknowledges The National Council of Science and Technology of Mexico (CONACyT) for the doctoral grant number 210282 to undertake doctoral studies. H. Sossa thanks CONACyT under call: Frontiers of Science (grant number 65) for the economic support. The research leading to these results has been funded by the European Unions Seventh Framework Programme (FP7/2007-2013) under Grant Agreement No. 245986. We would like to express our sincere gratitude to Jena University research team for their fruitful comments and suggestions for significant improvement of this work, especially to Sven Sickert who help providing results with ICF algorithm.

References

1. Behmann, J., Mahlein, A.K., Rumpf, T., Römer, C., Plümer, L.: A review of advanced machine learning methods for the detection of biotic stress in precision crop protection. Precision Agriculture **16** (2014) 239–260
2. RHEA: Robot fleets for highly effective agriculture and forestry management. `http://www.rhea-project.eu/` (2016)
3. Li, M., Imou, K., Wakabayashi, K., Yokoyama, S.: Review of research on agricultural vehicle autonomous guidance. International Journal of Agricultural and Biological Engineering **2** (2009) 1–16
4. Mousazadeh, H.: A technical review on navigation systems of agricultural autonomous off-road vehicles. Journal of Terramechanics **50** (2013) 211 – 232
5. Vibhute, A., Bodhe, S.: Applications of image processing in agriculture: a survey. International Journal of Computer Applications **52** (2012)

6. Saxena, L., Armstrong, L.: A survey of image processing techniques for agriculture. Proceedings of Asian Federation for Information Technology in Agriculture (2014)
7. Haug, S., Michaels, A., Biber, P., Ostermann, J.: Plant classification system for crop /weed discrimination without segmentation. In: IEEE Winter Conference on Applications of Computer Vision. (2014) 1142–1149
8. Hlaing, S.H., Khaing, A.S.: Weed and crop segmentation and classification using area thresholding. International Journal of Research in Engineering and Technology (2014)
9. Tewari, V., Kumar, A.A., Nare, B., Prakash, S., Tyagi, A.: Microcontroller based roller contact type herbicide applicator for weed control under row crops. Computers and Electronics in Agriculture **104** (2014) 40 – 45
10. Wei, X., Jia, K., Lan, J., Li, Y., Zeng, Y., Wang, C.: Automatic method of fruit object extraction under complex agricultural background for vision system of fruit picking robot. Optik - International Journal for Light and Electron Optics **125** (2014) 5684 – 5689
11. Choi, K.H., Han, S.K., Han, S.H., Park, K.H., Kim, K.S., Kim, S.: Morphology-based guidance line extraction for an autonomous weeding robot in paddy fields. Computers and Electronics in Agriculture **113** (2015) 266 – 274
12. Torres-Snchez, J., Lpez-Granados, F., Pea, J.: An automatic object-based method for optimal thresholding in {UAV} images: Application for vegetation detection in herbaceous crops. Computers and Electronics in Agriculture **114** (2015) 43 – 52
13. Yang, W., Zhao, X., Wang, S., Chen, L., Chen, X., Lu, S.: A New Approach for Greenness Identification from Maize Images. In: Intelligent Computing Theories and Methodologies: 11th International Conference, ICIC 2015, Fuzhou, China, August 20-23, 2015, Proceedings, Part I. Springer International Publishing, Cham (2015) 339–347
14. Jiang, G., Wang, Z., Liu, H.: Automatic detection of crop rows based on multi-rois. Expert Syst. Appl. **42** (2015) 2429–2441
15. Meng, Q., Qiu, R., He, J., Zhang, M., Ma, X., Liu, G.: Development of agricultural implement system based on machine vision and fuzzy control. Computers and Electronics in Agriculture **112** (2015) 128 – 138 Precision Agriculture.
16. Guijarro, M., Riomoros, I., Pajares, G., Zitinski, P.: Discrete wavelets transform for improving greenness image segmentation in agricultural images. Computers and Electronics in Agriculture **118** (2015) 396–407
17. Balasubramaniam, P., Ananthi, V.P.: Segmentation of nutrient deficiency in incomplete crop images using intuitionistic fuzzy c-means clustering algorithm. Nonlinear Dynamics **83** (2015) 849–866
18. Kazmi, W., Garcia-Ruiz, F.J., Nielsen, J., Rasmussen, J., Andersen, H.J.: Detecting creeping thistle in sugar beet fields using vegetation indices. Computers and Electronics in Agriculture **112** (2015) 10 – 19 Precision Agriculture.
19. Kazmi, W., Garcia-Ruiz, F., Nielsen, J., Rasmussen, J., Andersen, H.J.: Exploiting affine invariant regions and leaf edge shapes for weed detection. Computers and Electronics in Agriculture **118** (2015) 290 – 299
20. Ye, M., Cao, Z., Yu, Z., Bai, X.: Crop feature extraction from images with probabilistic superpixel markov random field. Computers and Electronics in Agriculture **114** (2015) 247 – 260
21. Cheng, B., Matson, E.T.: A Feature-Based Machine Learning Agent for Automatic Rice and Weed Discrimination. In: Artificial Intelligence and Soft Computing: 14th International Conference, ICAISC 2015, Zakopane, Poland, June 14-18, 2015, Proceedings, Part I. Springer International Publishing, Cham (2015) 517–527
22. Moorthy, S., Boigelot, B., Mercatoris, B.: 30. In: Effective segmentation of green vegetation for resource-constrained real-time applications. Precision agriculture '15 (2015) 257–266
23. Santos, T.T., Koenigkan, L.V., Barbedo, J.G.A., Rodrigues, G.C.: 3D Plant Modeling: Localization, Mapping and Segmentation for Plant Phenotyping Using a Single Hand-held Camera. In: Computer Vision - ECCV 2014 Workshops: Zurich, Switzerland, September 6-7 and 12, 2014, Proceedings, Part IV. Springer International Publishing, Cham (2015) 247–263
24. Ionescu, R.T., Popescu, A.L., Popescu, M., Popescu, D.: Biomassid: A biomass type identification system for mobile devices. Computers and Electronics in Agriculture **113** (2015) 244 – 253

25. Woebbecke, D., Meyer, G., Von Bargen, K., Mortensen, D.: Color indices for weed identification under various soil, residue, and lighting conditions. Transactions of the ASAE **38** (1995) 259–269
26. Meyer, G., Mehta, T., Kocher, M., Mortensen, D., Samal, A.: Textural imaging and discriminant analysis for distinguishingweeds for spot spraying. Transactions of the ASAE **41** (1998) 1189
27. Kataoka, T., Kaneko, T., Okamoto, H., et al.: Crop growth estimation system using machine vision. In: Advanced Intelligent Mechatronics, 2003. AIM 2003. Proceedings. 2003 IEEE/ASME International Conference on. Volume 2., IEEE (2003) b1079–b1083
28. Meyer, G.E., Neto, J.C.: Verification of color vegetation indices for automated crop imaging applications. Computers and Electronics in Agriculture **63** (2008) 282–293
29. Woebbecke, D.M., Meyer, G.E., Von Bargen, K., Mortensen, D.A.: Plant species identification, size, and enumeration using machine vision techniques on near-binary images. In: Applications in Optical Science and Engineering, International Society for Optics and Photonics (1993) 208–219
30. Golzarian, M.R., Frick, R.A.: Classification of images of wheat, ryegrass and brome grass species at early growth stages using principal component analysis. Plant Methods **7** (2011) 1–11
31. Salton, G., Mcgill, M.J.: Introduction to Modern Information Retrieval. McGraw-Hill, Inc., New York, NY, USA (1986)
32. Bosch, A., Muoz, X., Mart, R.: Which is the best way to organize/classify images by content. Image and Vision Computing **25** (2007) 778 – 791
33. Bishop, C.M.: Pattern Recognition and Machine Learning (Information Science and Statistics). Springer-Verlag New York, Inc., Secaucus, NJ, USA (2006)
34. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vision **60** (2004) 91–110
35. Bay, H., Ess, A., Tuytelaars, T., Gool, L.V.: Speeded-up robust features (surf). Computer Vision and Image Understanding **110** (2008) 346 – 359 Similarity Matching in Computer Vision and Multimedia.
36. MacQueen, J.: Some methods for classification and analysis of multivariate observations. In: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability. Volume 1., Berkeley, Calif., University of California Press (1967) 281–297
37. Duda, R.O., Hart, P.E., Stork, D.G.: Pattern classification. John Wiley & Sons (2012)
38. Chang, C.C., Lin, C.J.: LIBSVM: A library for support vector machines. ACM Transactions on Intelligent Systems and Technology **2** (2011) 27:1–27:27 Software available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm`.
39. Labatut, V., Cherifi, H.: Accuracy measures for the comparison of classifiers. CoRR **abs/1207.3790** (2012)
40. Kohonen, T., ed.: Self-organizing Maps. Springer-Verlag New York, Inc., Secaucus, NJ, USA (1997)
41. Dunn, J.C.: A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters. Journal of Cybernetics **3** (1973) 32–57
42. Felzenszwalb, P.F., Huttenlocher, D.P.: Efficient graph-based image segmentation. International Journal of Computer Vision **59** (2004) 167–181
43. Romeo, J., Guerrero, J.M., Montalvo, M., Emmi, L., Guijarro, M., Gonzalez-de Santos, P., Pajares, G.: Camera sensor arrangement for crop/weed detection accuracy in agronomic images. Sensors **13** (2013) 4348
44. Guijarro, M., Pajares, G., Riomoros, I., Herrera, P., Burgos-Artizzu, X., Ribeiro, A.: Automatic segmentation of relevant textures in agricultural images. Computers and Electronics in Agriculture **75** (2011) 75 – 83
45. Brust, C., Sickert, S., Simon, M., Rodner, E., Denzler, J.: Convolutional patch networks with spatial prior for road detection and urban scene understanding. CoRR **abs/1502.06344** (2015)
46. Fröhlich, B., Rodner, E., Denzler, J.: Semantic Segmentation with Millions of Features: Integrating Multiple Cues in a Combined Random Forest Approach. In: Computer Vision – ACCV 2012: 11th Asian Conference on Computer Vision, Daejeon, Korea, November 5-9, 2012, Revised Selected Papers, Part I. Springer Berlin Heidelberg, Berlin, Heidelberg (2013) 218–231
47. Larlus, D., Verbeek, J., Jurie, F.: Category level object segmentation by combining bag-of-words models with dirichlet processes and random fields. International Journal of Computer Vision **88** (2010) 238–253