An Effective Classification System for Separating Sugar Beets and Weeds for Precision Farming Applications

P. Lottes   M. Hoeferlin   S. Sander   M. Müter   P. Schulze Lammers   C. Stachniss

Abstract— Robots for precision farming have the potential to reduce the reliance on herbicides and pesticides through selectively spraying individual plants or through manual weed removal. To achieve this, the value crops and the weeds must be identified by the robot's perception system to trigger the actuators for spraying or removal. In this paper, we address the problem of detecting the sugar beet plants as well as weeds using a camera installed on a mobile robot operating on a field. We propose a system that performs vegetation detection, feature extraction, random forest classification, and smoothing through a Markov random field to obtain an accurate estimate of the crops and weeds. We implemented and thoroughly evaluated our system on a real farm robot on different sugar beet fields and illustrate that our approach allows for accurately identifying the weed on the field.

I. Introduction

One target of sustainable farming is to increase yield while reducing reliance on herbicides and pesticides. Precision farming techniques seek to address this challenge by monitoring key indicators of crop health and targeting treatment only to plants that need it. Doing this manually, is a time consuming and expensive activity. There has, however, been a great progress on autonomous farming robots that target to automate the work on the field.

In order to build autonomous robots for farming applications, several challenges need to be addressed. These challenges include robust perception, fast and effective actuators, rough terrain navigation, long-term autonomy, and several others. In this paper, we investigate the first of those challenges, namely a part of the perception problem. Our aim is to develop an effective, vision-based perception system that can identify the value crop and distinguish it from weeds growing on the field. By automatically separating both classes of plants, we enable the robot to manually remove the weed or to perform spraying actions on a per-plan basis. An illustration of our field robot and an example classification results is depicted in Figure 1.

The contribution of this paper is a vision-based classification system for mobile robots to separate value crops from weeds. The crop under consideration here are sugar beets, an important crop in Germany and other countries in Northern Europe. Our proposed system executes several steps. It first separates the vegetation from the remaining parts of the
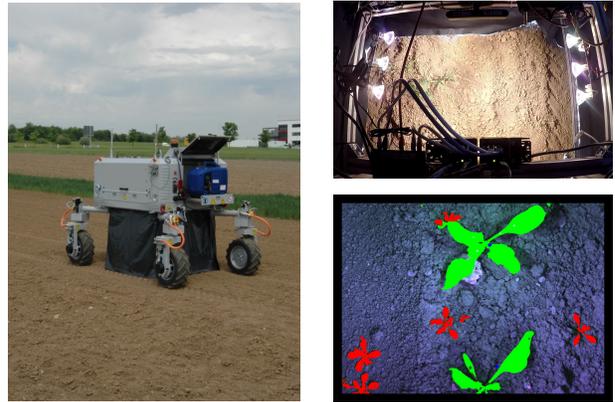
Fig. 1: Left: Bonirob V3 robot operating on a sugar beet field. Right: Downward looking camera capturing images of object space illuminated by halogen spots and example image with sugar beet/weed classification illustrated through green (crop) and red (weed) labels.

image, i.e., mostly soil. Then, we compute a series of features in the image regions that correspond to vegetation and exploit a random forest for performing the classification based on our features computed over several image regions and channels. In a last step, we take the neighborhood information between classified regions into account though Markov random fields and in this way improve the individually computed classifications of the random forest. Our approach can also exploit spatial priors, for example, if value crops were roughly planted at a known distance. We implemented the proposed system as ROS modules and evaluated them on different real field robots. For our experiment, we used different versions of BOSCH's Bonirob system and the sensor for classification is a single 4-channel JAI camera.

We evaluated our approach on sugar beet plants at different growth stages and weed plants that grew on test fields near Stuttgart, Germany. As the evaluations suggest, our system provides accurate classification results. In our precision farming scenario, it is important to keep the number of false negatives, i.e., the number of sugar beet plants that are classified as weeds, small. This type of misclassification should be avoided as this would lead to the elimination of the value crop by the robot. In contrast, not detecting a weed is less critical. The evaluation of our approach suggest that the majority of weed plants get correctly classified while the number of false positives stays small.

II. Related Work

Extracting semantic information about the environment is relevant topic in robotics [14], [17], [11]. In the context

of agricultural applications, several vision-based crop and weed detection approaches for specific plants have been proposed. Whereas traditional methods on plant phenotyping typically bring the plant into a specialized, static sensor array, several innovative solutions have been developed for on-field operation. This includes our previous work by Müter et al. [11], which focused on the manual removal of weeds through the design and control of a mechanism for intra-row weeding. Related to that Nieuwenhuizen [12] presents an approach on the automated detection and control of volunteer potato plants.

Borregaard et al. [2] perform crop versus weed classification using narrowband reflectance at 694 nm and 970 nm. Feyaerts and van Gool [7] conducted a multi-spectral machine vision study with the aim to design an online weed detection system for selective spraying. They collected multi-spectral images using six channels with different wave length (441, 446, 459, 883, 924 and 988 nm) in the field. They report crop versus weed classification rates of 80% for sugar beet plants and 91% for weeds. Related to our method, Haug et al. [8] present a method to classify carrot plants and weeds in RGB- and NIR-images without needing a pre-segmentation of the scenes into agglutinative objects. They achieve an average accuracy of 94% for carrot plants on an evaluation set of 70 images where both, intra- and inter-row overlap is present.

Other researchers have investigated the use of texture computed from grayscale and color images to identify plant species. Shearer and Holmes [16] used color co-occurrence matrices in the hue, saturation and intensity color space. Related to that, Burks et al. [4] evaluated color texture classification of different weed species using a neural network classifier. Both works report that using statistical parameters extracted from co-occurrence matrices provide high discriminative power to identify or separate plants but accentuate that more research is needed for testing under uncontrolled field conditions.

Several work has been conducted in the context of leaf image classification and segmentation [19], [10], [5]. In the work by Wang et al. [19], leaf images are segmented using morphological operators and shape features are extracted and used in a moving center hypersphere classifier to infer plant species. Kumar et al. [10] start from segmented images of leaves using a binary classifier on global image signatures as a validity test and curvature features compared with a given database to extract the best match. To cover a variety of leaf shapes, also deformable leaf models and morphology descriptors have been exploited by Cerutti et al. [5].

Tellaeche et al. [18] present a vision-based approach for selective weed spraying. They capture images inclined downwards with respect to the horizontal plane of field scenes and subdivide them in to grid cells. For each cell a decision is made based on structural and area features using Bayesian decision theory. A further cell-based approach by Aitkenhead et. al [1] fragments images in a top-down fashion containing seedlings of crop and weeds into 16 cells and classify each of them using a self organized neural network.

Fig. 2: Top: RGB and NIR image. Bottom: NDVI and masked NDVI image according to the vegetation detection.

They attain a classification performance close to 80%, but as in case of [18] at the comparably low resolution of the cells. In contrast, we provide labels for the full image resolution to allow a high precision treatment in object space.

Hemming and Rath [9] propose a vision based system which distinguishes carrots, cabbage and weeds using a fuzzy logic classifier. For leaf classification, they perform a pre-segmentation into individual plants to extract shape and color features per segment. They evaluate their approach in open field experiments and achieve classification accuracies of 72% up to 88%, but report that pre-segmentation of plants entails problems and embodies a limiting factor.

The contribution of this work is a visual 4-channel detection system for mobile robots operating on the field that allows for separating weeds from a value crop, here done for sugar beets. The proposed system performs vegetation detection, feature extraction, random forest classification, and smoothing through a Markov Random Field to obtain an accurate estimate of the crops and weed.

III. Vision-Based Plant Classification

The primary objective of our proposed plant classification system is to enable mobile field robots to distinguish crops and weeds in agricultural field environments. Here, we consider sugar beets, a popular value crop in Northern Europe, especially in Germany. The classification is performed on a mobile robot, see Figure 1, which perceives the field using a 4-channel camera that offers in addition to RGB information also one near-infra-red (NIR) intensity measurement per pixel, i.e., RBG+NIR. See Figure 2 for example images. The NIR information is especially useful for separating the vegetation from the soil and other background due to the high reflectivity of chlorophyll and thus (healthy) plants in the NIR spectrum.

The main goal of the proposed system is identifying crop and weeds on a per-pixel basis in the RGB+NIR camera images. Our overall pipeline works in four steps: First, we identify the vegetation using the NIR information, which leads to a vegetation mask $\mathcal{I}_{\mathcal{V}}$, see Figure 2 (bottom right) for an example. This step is highly effective as it allows

us to compute the features on the subsequent processing steps only for the regions that correspond to vegetation. Second, we compute a set of features for the image regions that correspond to vegetation. Third, we perform random forest classification of the vegetation based on the features extracted in the previous step. This yields a binary probability distribution representing the fact that the pixel corresponds to our crop or to weed. Fourth, we exploit the information from the random forest in a Markov Random Field to obtain smoothed results by utilizing the neighborhood information. In the subsequent Subsections III-A-III-E, we provide a detailed description of the four steps.

A. Vegetation Detection

The goal of vegetation detection is to eliminate the irrelevant background from the image $\mathcal{I}$ so that the subsequent classification task operates on regions that correspond to vegetation. Due to the hight reflectivity of chlorophyll in the NIR spectrum [15], it is comparably easy to separate vegetation from soil or other objects. We compute a vegetation mask

$$\mathcal{I}_{\mathcal{V}}(i,j) = \begin{cases} 1, & \text{if } \mathcal{I}(i,j) \in vegetation \\ 0, & \text{otherwise} \end{cases}, \quad (1)$$

with the pixel location (i,j).

To separate the vegetation, we exploit specific reflectance of healthy vegetation using the normalized difference vegetation index (NDVI) according to Rouse et al. [15] using the NIR channel $\mathcal{I}_{NIR}$ and the red channel $\mathcal{I}_R$ on a per-pixel basis:

$$\mathcal{I}_{NDVI}(i,j) = \frac{\mathcal{I}_{NIR}(i,j) + \mathcal{I}_R(i,j)}{\mathcal{I}_{NIR}(i,j) - \mathcal{I}_R(i,j)} \quad (2)$$

Figure 2 (bottom left) shows an example of a NDVI image $\mathcal{I}_{NDVI}$ for sugar beet plants and weeds. On the field, the reflectivity of chlorophyll typically leads to a bimodal intensity distribution in $\mathcal{I}_{NDVI}$ for healthy vegetation and allows us to perform a threshold-based classification on the $\mathcal{I}_{NDVI}$ information for every pixel.

In most cases, a threshold-based classification based on the $\mathcal{I}_{NDVI}$ will lead to small residual errors. They are typically caused by lens errors, especially chromatic aberration, resulting in slightly different mappings of the red and the near-infra-red light from the work space to pixels on the chip. Most of the residual errors ca be eliminated through basic image processing techniques such as (i) requiring a minimum brightness in $\mathcal{I}_{NIR}$, (ii) using morphological opening and closing to fill gaps and to remove noise at contours, and (iii) removing regions corresponding a few pixels only. Figure 2 (bottom right) depicts the application of the vegetation mask $\mathcal{I}_{\mathcal{V}}$ on the $\mathcal{I}_{NDVI}$ image.

B. Feature Extraction

As plants are expected to have a minimal size, we do not perform a classification for each pixel but use a grid tessellation of the image after applying the vegetation-mask. We compute features around keypoints $\mathcal{K}$ and in our current implementation, the keypoints are spaced 10 pixel by 10 pixel apart, which corresponds to a size of 3 mm by 3 mm in object
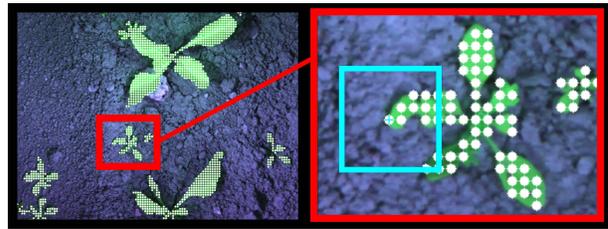

Fig. 3: Keypoints (white) for classification at a 3 mm distance on the object and an example neighborhood (blue) representing the region that is considered for the feature computation.

space. This leads to substantial computational savings when computing the features and we found no loss in performance in our experiments. The features computed at each keypoint take the local neighborhood into account. In our current implementation, the neighborhood $\mathcal{P}(\mathcal{K})$ of a keypoint $\mathcal{K}$ has a size of 80 pixel by 80 pixel. Figure 3 illustrates the arrangement of the keypoints on an image and including the neighborhoods. The features we compute can be categorized into two main groups:

- **Statistical features**: Our set $\mathcal{F}_{St}(x, \mathcal{P}(\mathcal{K}))$ of statistical features includes *min, max, range, mean, standard deviation, median, skewness, kurtosis, and entropy*. These statistical features are computed in the local neighborhood $\mathcal{P}(\mathcal{K})$ of a $\mathcal{K}$ on different inputs x. The input x is not only computed in the raw image itself but the channel G, B separately, the NDVI image as well as on the magnitude of the first and second gradient of the NDVI image, i.e.,
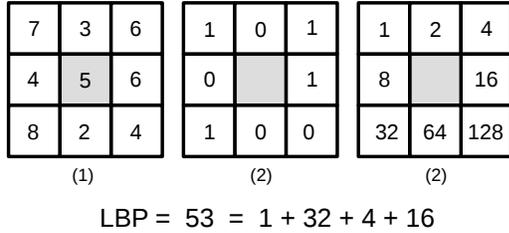
$$\nabla \mathcal{I}_{NDVI} = \left| \frac{\partial \mathcal{I}}{\partial x} + \frac{\partial \mathcal{I}}{\partial y} \right| \quad (3)$$

$$\Delta \mathcal{I}_{NDVI} = \left| \frac{\partial^2 \mathcal{I}}{\partial x} + \frac{\partial^2 \mathcal{I}}{\partial y} \right| \quad (4)$$

The magnitudes of the directed gradients provide information about structures or homogeneous regions as well as a source for specific statistics of a local neighborhood.

In addition to that, we compute the statistical features also on the texture. To achieve this, we employ local binary patterns (LBP) according to Ojala and Pietikääinen [13] as features to encode texture information. The LBP operator performs thresholding operations within a 8-connected neighborhood based on the value of its center pixel and converts this pattern as binary number. Figure 4 gives an example for the computation of a LBP number and the associating contrast measure for a pixel. We compute the joint probability distribution $p(LBP, C)$ for each $\mathcal{P}(\mathcal{K})$ based on $\mathcal{I}_{NDVI}$ and $\Delta \mathcal{I}_{NDVI}$ considering each vegetation pixel and respectively extract the statistical features from it.

- **Shape features**: $\mathcal{F}_{Sh}(\mathcal{P}(\mathcal{K}))$ describe different aspects of the shape of the vegetation in the local neighborhood $\mathcal{P}(\mathcal{K})$ of a keypoint $\mathcal{K}$. Thus, the shape features only need to be computed on the vegetation mask $\mathcal{I}_{\mathcal{V}}$,

LBP = 53 = 1 + 32 + 4 + 16

C = 3.5 = (7 + 6 + 6 + 8) / 4 − (3 + 4 + 2 + 4) / 4

Fig. 4: Example for the computation of a LBP number and the corresponding contrast measure C for a pixel given its 8-connected neighborhood. (1) input, (2) threshold operation by value of center pixel, and (3) binomial weights according to [13].

which is a binary image. We consider several features such as contours, relations to geometric primitives and geometrical ratios of windowed pieces of vegetation. We exploit all features as defined by Haug et al. [8]. Additionally, we also use

- radius of *minimum enclosing circle* of the contour
- *aspect ratio* of the major and minor axes of the enclosing ellipse of the contour
- *area change under smoothing*, which describes how the area of vegetation changes due to a smoothing with different kernel sizes and is given by the ratio of the areas.
- *form factor* F, which provides a measure of the shape of an object

$$ F = \frac{4\,\pi\,\text{area}}{\text{perimeter}^2} \tag{5} $$

Table I gives a summary of all features used in our classification system and described on which inputs they are computed.

C. Relative Plant Arrangement Prior

Due to the fact that crop is often sowed in an automated fashion, prior information about the arrangement of the plants on the field can further improve the classification performance. We incorporate a plant arrangement as an additional feature. We use the probability for a certain keypoint $\mathcal{K}_i$ corresponding to a crop based on it relative location to previously classified crops.

To represent the relative arrangement of crops, we use a coordinate system that is defined according to the crop row, which defines the x-direction. Given this coordinate system, we compute the absolute difference $\boldsymbol{D}_i$ of the keypoint i in x and y direction between the keypoint and all positively classified J crops:

$$ \boldsymbol{D}_i = (|\boldsymbol{d}_{i1}|, \ldots, |\boldsymbol{d}_{iJ}|)^\top \text{ with } |\boldsymbol{d}_{ij}| = (|\Delta x_{ij}|, |\Delta y_{ij}|)^\top \tag{6} $$

Based on these distance, we compute

$$ \mathcal{F}_{23} = p(\omega_c|\boldsymbol{D}) = \frac{p(\boldsymbol{D}|\omega_c)\,p(\omega_c)}{\sum_\omega p(\boldsymbol{D}|\omega)\,p(\omega)} \tag{7} $$

Nr.	Feature Set
↓	**_Statistical Features_** $\mathcal{F}_{St}(x)$
$\mathcal{F}_1$	min
$\mathcal{F}_2$	max
$\mathcal{F}_3$	range
$\mathcal{F}_4$	mean
$\mathcal{F}_5$	standard deviation
$\mathcal{F}_6$	median
$\mathcal{F}_7$	skewness
$\mathcal{F}_8$	kurtosis
$\mathcal{F}_9$	entropy
	All statistical features $\mathcal{F}_{St}(x)$ are computed for x in $\mathcal{I}_{GREEN}, \mathcal{I}_{BLUE}, \mathcal{I}_{NDVI}, \Delta\mathcal{I}_{NDVI},$ $LBP(\mathcal{I}_{NDVI}), LBP(NDVI)$
↓	**_Shape features_** $\mathcal{F}_{Sh}$ computed on binary image $\mathcal{I}_\mathcal{V}$
$\mathcal{F}_{10}$	Perimeter of contour
$\mathcal{F}_{11}$	Area of contour
$\mathcal{F}_{12}$	Convexity
$\mathcal{F}_{13}$	Compactness
$\mathcal{F}_{14}$	Solidity
$\mathcal{F}_{15}$	Rectangularity
$\mathcal{F}_{16}$	Aspect ratio two axes of the enclosing ellipse
$\mathcal{F}_{17}$	Area change under smoothing
$\mathcal{F}_{18}$	Form factor
$\mathcal{F}_{19}$	Radius of min. enclosing circle (MEC)
$\mathcal{F}_{20}$	MEC / major axis of the enclosing ellipse
$\mathcal{F}_{21}$	Length of skeleton (medial axis)
$\mathcal{F}_{22}$	Skeleton to perimeter ratio
↓	**_Other features_**
$\mathcal{F}_{23}$	plant arrangement prior, see Eq. (7)
$\mathcal{F}_{24}$	$\mathcal{F}_9(\nabla\mathcal{I}_{NDVI})/\mathcal{F}_9(\Delta\mathcal{I}_{NDVI})$

TABLE I: Features uses by our classification system

where ω refers to a class label, with ω_c corresponding to crops and ω_w to weeds. The class conditional probability distribution $p(\boldsymbol{D}|\omega_c)$ and the relative frequency of the classes are learned from training data or can be provided if the information is known from sowing. For $p(\boldsymbol{D}|\omega_w)$, we consider a uniform distribution, due to the fact that in reality weed can grow anywhere. Note that this feature is only available if at least n neighboring images have already been classified. In our current implementation, we select $n{=}5$. Our tests have shown that this choice favors a good trade off between uncertainty of the relative pose estimation, caused by odometry errors, and the probability to find crop in this image sequence which in necessary to compute $\boldsymbol{D}_i$.

D. Random Forest Classification

For the classification, we apply a random forest [3] because it provides comparably robust classification results. As an ensemble method, random forests avoid overfitting to some degree and implicitly estimate confidences for the classes labels. Random forests can be used for classification and regression. The key idea is to construct as large number of decision trees ("a forest") at training time by randomizing the use of features and elements from the training data. During the operational phase, they output a class label or a distribution over the class label based on the output of the individual decision trees. Random forests run efficiently on large datasets and trained classifiers are fast to evaluate
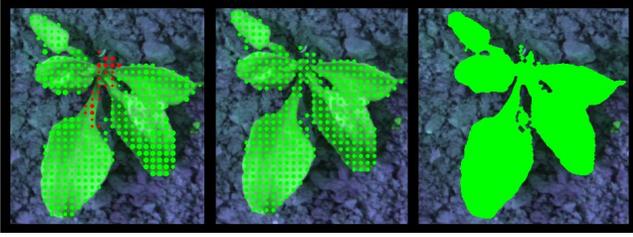
Fig. 5: Left to right: Random forest classification, MRF smoothing at keypoints, interpolation at every pixel. The MRF smoothing eliminates the few wrongly classified keypoints at the center.

and are also easy to parallelize if performance matters. An interesting property of the random forest allows for dealing with missing data, for example, if feature is not available. This is relevant for our application in case the relative plant arrangement prior is not known.

As a result of the random forest classification, we obtain for every keypoint a probability if it corresponds to a sugar beet plant or to weed. Note that since random forests are capable of solving multi-class problems, they have the potential to also separate different weed classes, which can be used to select specific weed removal or spraying treatments. Such investigation, however, is future work and have not been addressed here.

E. Smoothing Thorough Markov Random Fields

The classification system described so far, computes each label assignment independently of the other nearby labels. In order to improve the classification results and to exploit the topological relationships between keypoints, we apply a Markov random field (MRF). We compute a global classification based on the individually computed class labels $\omega_\mathcal{K}$ of the keypoints by considering their spatial distribution and class confidences $p(\omega_\mathcal{K}|\mathcal{F})$. We achieve this by minimizing the energy function

$$E(\omega_\mathcal{K}) = \sum_\mathcal{K} \left(D(p(\omega_\mathcal{K}|\mathcal{F})) + \sum_{\mathcal{K}' \in \mathcal{N}_4(\mathcal{K})} V(\omega_{\mathcal{K}'}, \omega_\mathcal{K}) \right) \tag{8}$$

through belief propagation. Here, $E(\omega_\mathcal{K})$ describes the quality of a labeling under the key assumption that neighboring labels vary slightly, but also can change fiercely at class borders. Therefore, two energy terms are needed. The first one D considers the confidence of a class label and through this defines the energy which is needed to change the label. The term V describes the energy for smoothing the four-connected neighborhood, i.e., how many neighboring labels agree. We minimize Eq. (8) using efficient belief propagation according to [6]. This optimizes the classification, see Figure 5 for an example.

IV. EXPERIMENTS

The evaluation is designed to illustrate the performance of our plant classification system and to support the three main claims made in this paper. These three claims are: (i) our approach is suitable for classifying sugar beets and
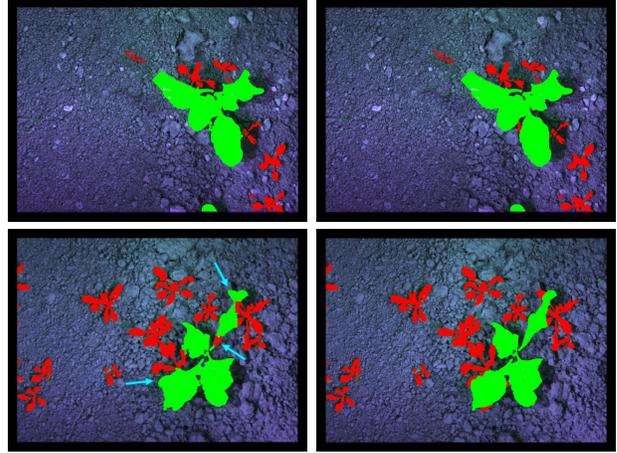

Fig. 6: Example results of our system. Left: classification results. Right: ground truth. The bottom left image depicts an example for a classification errors (blue arrows).

Parameter	Dataset A	Dataset B	Dataset C
no. images	1024	694	974
no of crops	1315	831	1145
$p(\omega_c)$ / $p(\omega_w)$	0.74 / 0.26	0.66 / 0.34	0.68 / 0.32
growth stage	4-leaf	late 4-leaf	2-leaf/early 4-leaf

TABLE II: Information about the datasets. *A* and *B* are collected on the same field with a temporal difference of one week. *C* is collected on another field.

weeds under real world conditions, (ii) we illustrate that our approach also provides good classification results on real fields if the grow stage of the vegetation has changed, and (iii) we show that using pose information from sowing can have a significant impact to the classification results.

A. Experimental Setup

All experiments have been conducted with different generations of the Bonirob field robot, shown in figure 1. For the image acquisition, we used a 4-channel JAI AD-130 GE camera pointing downwards on the field approximately 70 cm above soil. The RGB+NIR images were captured with a resolution of $1296{\times}966$ pixels using a Fujinon TF15-DA-8 lens, which yields a ground resolution of roughly $3\,\frac{px}{mm}$ and a field of view of 24 cm in driving direction and 31 cm orthogonal to it. The images were captured with a frequency of 1 Hz while the robot was moving over the field with approximately $150\,\frac{mm}{sec}$. For aligning the images, we used the odometry information from the wheel encoders.

We used the robot on sugar beet fields near Stuttgart, Germany. In the evaluation reported below, we mainly use three datasets, here called *A*, *B* and *C*, see Table II. The datasets *A* and *B* have been collected on the same field with a temporal difference of one week. Dataset *C* has been collected on a different field. To allow for a ground truth evaluation, we manually labeled all sugar beet plants and weeds in the images. We compute all evaluation parameters by a keypoint-wise comparison between prediction and ground truth with a ground resolution of 3 mm. We refer to sugar beet plants (ω_c) as *positives* and weeds (ω_w) as *negatives*.
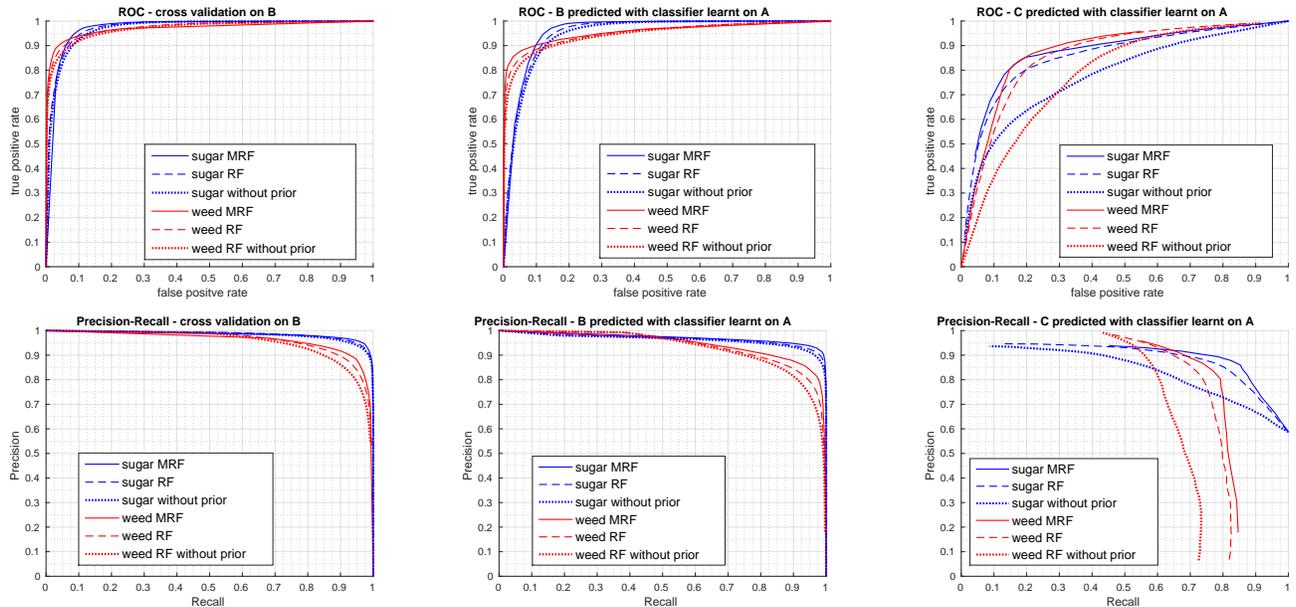
Fig. 7: ROC curves (top) and precision recall plots (bottom). Left: Performance evaluated using 15-fold cross validation on dataset *B*. Middle: Evaluation using a classifier trained on *A* and tested on *B*. Right: classifier trained on *A* and tested on *C*, i.e., a dataset from a different sugar beet field with a substantially different growth stage of the plants. The term label "RF" refers to random forest-only classification and "MRF" to the combination of random forest and MRF. "without prior" refers to the approach neglecting the relative plant arrangement prior, "sugar" to sugar beet plants and "weed" to weeds.

FNR	0.01%	0.5%	1%	2.5%	5%
TNR	55%	72%	80%	86%	91%

TABLE III: Limiting the false negative rate, i.e. the percentage of sugar beets classified as weeds to the values shown in the first row of the table, yields to the given true negative rates (TNR), i.e. percentage of correctly classified weeds that will be eliminated.

FPR	2.5%	5%	10%	15%	30%
TPR	72 %	85%	96%	98%	99.5%

TABLE IV: Limiting the a false positive rate (FPR) as specified in the first row, yields the true positive rates (TRP).

We illustrate the performance of the classification results by ROC curves and Precision-Recall plots. For such plots, we varied the threshold for class labeling concerning to the estimated confidences to compute the classification performance of the random forest. During all experiments, we only use independent image scenes, e.g., no overlapping images, to avoid redundant data in the evaluation.

B. Evaluation

The first set of experiments is designed to illustrate the performance of our sugar beet classification system. We apply a K-fold cross validation to B and a stratification of each training set according to the learned class priors $p(\omega)$. We chose $K = 15$ and exploit class priors from ground truth data.

We present here the results from dataset B because it is the most difficult dataset due to a substantial plant/weed overlaps. The resulting ROC- and precision-recall plots are depicted in the first column of Figure 7. We achieve a true

positive rate (TPR/recall) of 96% with a precision of 95% for sugar beets and a false positive rates (FPR) of 10%. In terms of weeds our system shows a TNR of 90% and a FNR of 3%. Thus, the system classifies the majority of plants correctly. Two classification examples vs. ground truth are shown in Figure 6.

Precision farming robots are also envisioned to manually eliminate weeds. Thus, a wrong classification of a sugar beet plant will lead to its elimination, while a false positive means that a weed will not be treated. In order to avoid eliminating value crops, the plant classification system has to avoid false negatives with a higher priority than false positives. Therefore, we provide in Table III the expected true negative rate *for a given false negative rate*. The inverse, which may be relevant for other application in the context of phenotyping, is shown in Table IV.

We furthermore evaluated the performance of our approach when the training is performed on a field with a different growth stage than in the test dataset. To do so, we classified dataset B using the random forest as well as the MRF and used only dataset A for learning. In both datasets relevant for weed removal the sugar beets are in 4-leaf-stage, but the plant size differs significantly, especially for the weeds. The resulting ROC- and precision-recall curve are shown in the second column of Figure 7.

Here, our system provides a comparable precision as before with a TPR of 90% and precision of 95% for sugar which means that the number of crops, which are classified as weed is still small. The main difference in performance measure is given by TNR of 90% with a precision of 82% for weeds, probably caused by its notable altered visual and

Rank	Feature
1	$\mathcal{F}_{23}$ plant arrangement prior
2	$\mathcal{F}_3(LBP(\Delta\mathcal{I}_{NDVI}))$
3	$\mathcal{F}_7(LBP(\Delta\mathcal{I}_{NDVI}))$
4	$\mathcal{F}_{24} \rightarrow \mathcal{F}_9(\nabla\mathcal{I}_{NDVI})/\mathcal{F}_9(\Delta\mathcal{I}_{NDVI})$
5	$\mathcal{F}_3(LBP(\Delta\mathcal{I}_{NDVI}))$
6	$\mathcal{F}_3(\mathcal{I}_{GREEN})$
7	$\mathcal{F}_9(LBP(\mathcal{I}_{NDVI}))$
8	$\mathcal{F}_9(LBP(\Delta\mathcal{I}_{NDVI}))$
9	$\mathcal{F}_6(\mathcal{I}_{BLUE})$
10	$\mathcal{F}_{17}$ area change under smoothing

TABLE V: The 10 most expressive features over all datasets. See Table I for a description of the features.

physical appearance.

For the next experiment, we increase the difficulty for classification and feed our trained system with images of an unknown fields that contains sugar beet of an earlier growth stage in a 2-leaf and partially early 4-leaf stage (dataset *C*). In most real world applications, one would avoid such situations by providing a classifier trained for the appropriate growth state but it is worth investigating the loss in performance. The results are depicted in the third column of Figure 7. Obviously, the performance decreases but the achieved TPR of 85% and TNR of 79% indicate that our proposed system also delivers usable results. The performance is sufficient for selective spraying applications but probably not for mechanical weed removal, as too many crops would be removed by the robot.

In all experiments, we also evaluated the effect of using the relative plant arrangement prior from sowing or learned from training data. The dotted lines in all plots show the identical classification system but ignoring the relative plant arrangement information. In all cases, the prior helps to obtain better classification results. On average, we obtain a gain in classification accuracy of 3% over both classes for the first and second experiment. In case of the classification in an unknown environment the impact of the prior information becomes even larger, i.e. 9% improvement in classification accuracy for dataset *C*. In sum, the results illustrate that using spatial prior information about the crop arrangement is a valuable information for classification systems.

Finally, we looked at the relevance of the individual feature for our classification task over all datasets. The ten most important features are listed in Table V. As can be seen, the NDVI information and its Laplacian, texture information and the pose arrangement prior are key supporters for the classification task.

V. Conclusion

Robots for precision farming must be able to distinguish the crops on the field from weeds. We addressed the problem of detecting sugar beet plants and weeds using a camera installed on a mobile robot operating on a real field. We developed a system that performs vegetation detection and feature extraction, classification, and smoothing. Our system uses statistical and shape features computed on the different channels of the image and has the ability to exploit a spatial arrangement prior from sowing. To decide which area of an image corresponds to sugar beets and weeds, we combine random forest classification and in addition exploit the neighboring information through a Markov random field. We implemented our approach as ROS modules and thoroughly evaluated it on a real farm robot on three different sugar beet fields and illustrate that our approach allows for accurately identifying the weed on the field.

Despite these encouraging results, there is further space for improvements. We currently extend of our work to decrease its runtime, which currently takes around 4 s per image due to the large number of features that are computed.

References

[1] M.J. Aitkenhead, I.A. Dalgetty, C.E. Mullins, A.J.S. McDonald, and N.J.C. Strachan. Weed and crop discrimination using image analysis and artificial intelligence methods. *Computers and electronics in Agriculture*, 39(3):157–171, 2003.

[2] T. Borregaard, H. Nielsen, L. Norgaard, and H. Have. Crop-weed discrimination by line imaging spectroscopy. *J. Agric. Eng. Res.*, 72(4):389–400, 2000.

[3] L. Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001.

[4] T.F. Burks, S.A. Shearer, R.S. Gates, and K.D. Donohue. Backpropagation neural network design and evaluation for classifying weed species using color image texture. *Transactions of the American Society of Agricultural Engineers*, 43(4):1029–1037, 2000.

[5] G. Cerutti, L. Tougne, J. Mille, A. Vacavant, and D. Coquin. A model-based approach for compound leaves understanding and identification. In *IEEE Intl. Conf. on Image Processing*, pages 1471–1475, 2013.

[6] P. F. Felzenszwalb and D. P. Huttenlocher. Efficient belief propagation for early vision. *Int. Journal on Computer Vision*, 70, October 2006.

[7] F. Feyaerts and L. van Gool. Multi-spectral vision system for weed detection. *Pattern Recognit. Lett.*, 22:667–674, 2001.

[8] S. Haug, A. Michaels, P. Biber, and J. Ostermann. Plant classification system for crop / weed discrimination without segmentation. In *Proc. of the IEEE Winter Conf. on Applications of Computer Vision (WACV)*, pages 1142–1149, 2014.

[9] J. Hemming and T. Rath. Computer-vision-based weed identification under field conditions using controlled lighting. *Journal of Agricultural Engineering Research*, 78(3):233 – 243, 2001.

[10] N. Kumar, P.N. Belhumeur, A. Biswas, D.W. Jacobs, W.J. Kress, I. Lopez, and J.V.B. Soares. Leafsnap: A computer vision system for automatic plant species identification. In *Proc. of the European Conference on Computer Vision (ECCV)*, 2012.

[11] M. Müter, P. Schulze Lammers, and L. Damerow. Development of an intra-row weeding system using electric servo drives and machine vision for plant detection. In *Proc. of the Agricultural Engineering Conference*, 2013.

[12] A.T. Nieuwenhuizen. *Automated detection and control of volunteer potato plants*. PhD thesis, Wageningen University, 2009.

[13] T. Ojala and M. Pietikinen. Unsupervised texture segmentation using feature distributions. *Pattern Recognition*, (32):477–486, 1999.

[14] A. Ranganathan and F. Dellaert. Semantic modeling of places using objects. In *Proc. of Robotics: Science and Systems (RSS)*, 2007.

[15] J. W. Rouse, Jr., R. H. Haas, J. A. Schell, and D. W. Deering. Monitoring Vegetation Systems in the Great Plains with Erts. *NASA Special Publication*, 351:309, 1974.

[16] S.A. Shearer and R.G. Holmes. Plant identification using color co-occurrence matrices. *Transactions of the American Society of Agricultural Engineers*, 33(6):2037–2044, 1990.

[17] C. Stachniss, O. Martínez-Mozos, A. Rottmann, and W. Burgard. Semantic labeling of places. In *Proc. of the Int. Symposium of Robotics Research (ISRR)*, San Francisco, CA, USA, 2005.

[18] A. Tellaeche, X.P. Burgos-Artizzu, G. Pajares, and A. Ribeiro. A vision-based method for weeds identification through the bayesian decision theory. *Pattern Recognition*, 41(2):521–530, 2008.

[19] X.-F. Wang, D. Huang, J. Du, H. Xu, and L. Heutte. Classification of plant leaf images with complicated background. *Applied Mathematics and Computation*, 205:916–926, 2008.