Plant Classification System for Crop / Weed Discrimination without Segmentation

Sebastian Haug [1] Andreas Michaels [1] Peter Biber [1] Jörn Ostermann [2]
{firstname.lastname}@de.bosch.com ostermann@tnt.uni-hannover.de
[1] Robert Bosch GmbH, Corporate Research [2] Leibniz Universität Hannover

Abstract

This paper proposes a machine vision approach for plant classification without segmentation and its application in agriculture. Our system can discriminate crop and weed plants growing in commercial fields where crop and weed grow close together and handles overlap between plants. Automated crop / weed discrimination enables weed control strategies with specific treatment of weeds to save cost and mitigate environmental impact.

Instead of segmenting the image into individual leaves or plants, we use a Random Forest classifier to estimate crop / weed certainty at sparse pixel positions based on features extracted from a large overlapping neighborhood. These individual sparse results are spatially smoothed using a Markov Random Field and continuous crop / weed regions are inferred in full image resolution through interpolation.

We evaluate our approach using a dataset of images captured in an organic carrot farm with an autonomous field robot under field conditions. Applying the plant classification system to images from our dataset and performing cross-validation in a leave one out scheme yields an average classification accuracy of $93.8\,\%$.

1. Introduction

The incentive for automating weed control in agriculture with machine vision and autonomous field robots is manifold: In conventional farming, the amount of chemical herbicides necessary can be reduced to minimize cost and reduce pollution. In organic farming, weed control is currently done manually which is both hard and tedious work and very costly. Using automated systems that are able to work 24 / 7, weed control is expected to be achieved more efficiently, environmentally friendly and cost-effectively.

Precise crop / weed discrimination is a major requirement to realize such systems for precision weed control. Weed plants growing close-to-crop or intra-row need to be regulated to avoid substantial yield loss [19]. However, these types of weed require sophisticated detection and classification methods as crop and weed are in close proximity, possibly overlap and loss of crop must be minimized.

We focus on the perception part of such robotic weed control systems and propose a new approach for plant classification that does not require segmentation into individual plants or leaves. Instead we show that crop and weed can be discriminated based on features extracted from patches representing the neighborhood of sparse keypoints arranged in a grid in image space. The image patches overlap because the patch size is significantly larger than the spacing of the grid. The system applies machine learning to train a classifier that discriminates between crop and weed. Based on the classification results for each of the sparse keypoints, continuous crop / weed regions in the image are inferred through interpolation. Figure 1 shows an input image next to the output image in which the predicted crop / weed classes are color coded. The evaluation indicates that our method achieves good performance on field images even in situations with overlap.

Figure 1: Input image (left) and output (right) of the plant classification system where the predicted plant class is color coded (crop in green, weeds in blue and red). See Figure 5 #1 for ground truth and more details.

The plant classification system is integrated into a new version of the autonomous field robot Bonirob [16] built for the use case of weed control in commercial organic carrot farms (project RemoteFarming.1). The robot will also carry a mechanical weed regulation unit [12] which acts based on the output of the plant classification system and treats weed.

The main contributions of this paper are:

- A new method for plant classification without segmentation: Feature extraction and classification are performed on overlapping image patches. They represent the neighborhood of sparse keypoints which are arranged in a grid.
- This approach enables our system to handle overlap of plants and irregular shaped leaves. No prior segmentation into plants or leaves is required.
- The sparse results are spatially smoothed and interpolated. The system outputs per pixel crop / weed estimates in full image resolution. The precision loss of cell based methods that classify large non-overlapping cells (see related work) is avoided.
- Applying the plant classification system to an image dataset captured on a commercial carrot farm results in an average classification accuracy of $93.8\,\%$.

2. Related Work

Computer vision techniques have been applied to solve plant classification tasks at different levels:

When considering only leaves, methods based on shape, color and texture have proven effective to discriminate between different types of leaves. Beghin *et al*. [3] classify leaves from 10 species based on shape and texture with an average accuracy of 85%. Kumar *et al*. [11] developed a smartphone application for classification of leaves from trees in the Northeastern United States. These approaches share with other work [6, 10] that the input is an image of a flat leaf captured on a mostly homogeneous background.

In agriculture machine vision can be applied with the goal of intelligent weed control, but in general the requirement of single leaf images is not applicable to commercial field situations. Remote-sensing has been successfully applied to estimate weed densities and distributions on field level [22] with the goal to regulate herbicide usage. But for precision agricultural activities that are considered here, plant classification must be done on a much finer scale [5].

On ground level, camera based sensing can been applied to identify single plants or groups of a few plants and to classify them. Hemming *et al*. [9] present a robot and computer vision system which correctly classifies 51 to 95 % of plants based on color and shape features of segmented plants in top down images. They conclude that segmentation into individual plants is difficult and needs more research. Astrand & Baerveldt [2] use similar features on segmented plants for classification. They evaluate their system in greenhouse experiments with large plants (about 5 cm diameter) but do not present quantitative results on classification performance. Leaf segmentation has also been researched in field settings to enable crop / weed discrimination [8, 14]. Neto *et al*. [13] present a leaf segmentation technique that performs well on convex leaves but not for other leaf shapes (like carrots considered here). They conclude that more research is needed in these situations.

Cell based methods [1, 21] require no plant / leaf segmentation and have also been applied to crop / weed discrimination tasks. New images are processed by splitting them into grid cells (without overlap) and deciding for each cell individually whether it should be treated or not. Aitkenhead *et al*. [1] present a system that uses very coarse cells (16 per image) and analyzes them using a self-organizing neural network. They achieve a classification accuracy of approximately 75 % in experiments with plants specifically sown in a greenhouse. Tellaeche *et al*. [21] capture images with a frontal downward looking camera and partition them into non overlapping grid cells aligned to crop rows. They apply a Bayesian theoretic approach to decide whether or not to treat a grid cell. The cell based methods lack the precision of plant or per-pixel based methods and are not applicable if high precision treatment is desired.

Our approach closes this gap and avoids segmentation into plants or leaves which was determined a major problem in the literature. Although working without segmentation, the system still returns per pixel crop / weed classification results in full input image resolution.

3. Plant Classification Pipeline

The plant classification system is designed as a 6 stage pipeline performing the actual task of crop / weed discrimination and several offline steps during training. Figure 2 gives an overview of both the on- and offline processing steps and the data involved. In the following we address and explain each step in depth.

3.1. Image Acquisition

Input data consists of images captured by a multi-spectral down-looking monocular camera. The camera is a JAI AD-130 GE with 1.3 Mpx which outputs images at a resolution of 1296 px by 966 px. It is mounted on the autonomous field robot Bonirob and acquires images in the visible red (R) and near-infrared (NIR) spectrum. To avoid interference by changing environment conditions and to enable the robot to work around the clock, the space under the robot is shaded and artificial lighting is installed. The camera system is positioned approximately 45 cm above ground and the resulting ground resolution is $10\,\mathrm{px/mm}$. All images used for training and evaluation of our system were collected while the robot system was driving with 5 cm/s.

3.2. Background Removal

The goal of the background removal step is to segment the vegetation from soil. The different reflectance of

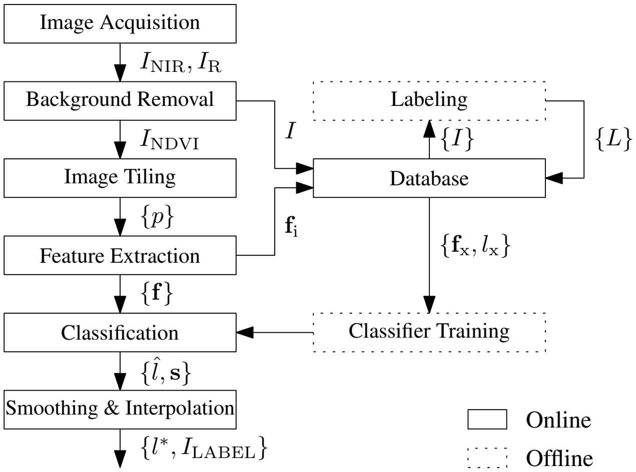
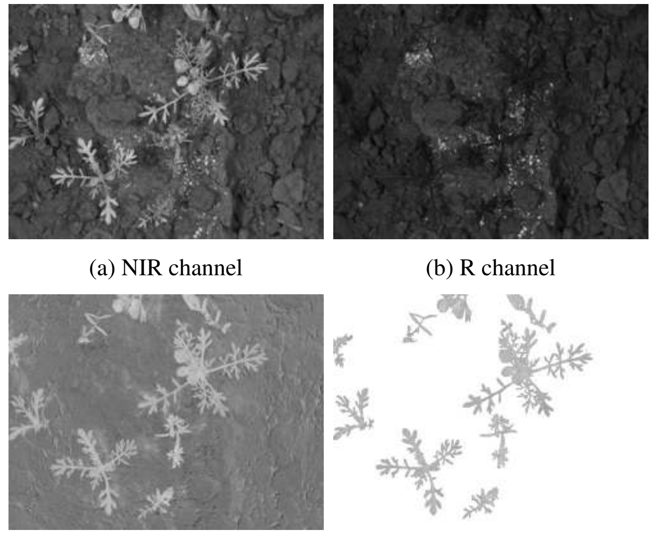


Figure 2: Overview of classification pipeline: The plant classification system itself is an online process (solid blocks), labeling of data and classifier training is done offline (dotted blocks).



(a) NIR channel (b) R channel

(c) NDVI image (d) Masked NDVI image

Figure 3: Background removal step: (a) and (b) are the input image channels, (c) the intermediate image and (d) the masked NDVI output image.

biomass and soil in the visible red and near-infrared light is used to mask soil pixels [18]. We apply the Normalized Difference Vegetation Index (NDVI) to each pixel pair in the input images I_{NIR} and I_{R} to define an I_{NDVI} image:

$$I_{\mathrm{NDVI}} = \frac{I_{\mathrm{NIR}} - I_{\mathrm{R}}}{I_{\mathrm{NIR}} + I_{\mathrm{R}}} \tag{1}$$

A threshold is selected in NDVI space using Otsu's method [15] and all pixels with NDVI values smaller than the threshold are masked. The output of the background removal step is a masked image in NDVI image space, in which only pixels belonging to plants are present. Figure 3 displays the four images (R, NIR, NDVI and masked NDVI image) for one position in the field.

3.3. Image Tiling

Keypoints are generated by applying a sparse grid (15 px by 15 px) to the image. For each keypoint located at a biomass pixel (i.e. pixel has not been masked in previous step) an image tile representing the neighborhood (80 px by 80 px) of this keypoint is extracted. In the following, computations will be done on these tiles called image patches $\{p\}$ and all results (e.g. their feature vectors) will be associated to the keypoints in the input image.

3.4. Feature Extraction

Features are extracted from all patches of the masked NDVI image generated in the previous step. A set of shape and contour features (based on typical features used in earlier work [6, 9], f_6 and f_7 are new additions of our own) and statistical features are calculated for each image patch, see Table 1 for a summary. The first seven features are extracted from a binarized version of the image patch (biomass vs. soil), the statistical features are calculated from the intensity values of biomass pixels in the image patch.

Table 1: Description of used features. Features f_1 to f_7 are contour / shape features, f_8 to f_{15} statistical features.

$\mathrm{f_i}$	**Description**
f_1	perimeter (length of contour)
f_2	area (number of pixels covered by biomass)
f_3	compactness (area / perimeter^2)
f_4	solidity (area / area of convex hull)
f_5	convexity (perimeter / perimeter of convex hull)
f_6	length of skeleton
f_7	length of skeleton / perimeter
f_8	minimum of biomass pixel intensities
f_9	maximum of biomass pixel intensities
f_{10}	range of biomass pixel intensities
f_{11}	mean of biomass pixel intensities
f_{12}	median of biomass pixel intensities
f_{13}	standard deviation of biomass pixel intensities
f_{14}	kurtosis of biomass pixel intensities
f_{15}	skewness of biomass pixel intensities

3.5. Classification

To be able to represent more than one weed class we use a multi-class classifier. We chose the Random Forest classifier [4] because it is multi-class, fast to train and able to estimate class certainty scores ($\mathbf{s}$) in addition to the most certain label ($\hat{l}$) during prediction of new samples. See Section 4 for details on classifier training.

3.6. Spatial Smoothing and Interpolation

During post-processing the classification results on the sparse grid are smoothed spatially and interpolated to full image resolution.

For smoothing a Markov Random Field approach is used to calculate a smoothed labeling l^* given the prediction from the classification (expressed by the score vector $\mathbf{s}$ for each keypoint, sum of elements in $\mathbf{s}$ is 1). The key assumption is that the final labeling should be mostly smooth as neighboring keypoints most likely belong to the same class and should have the same label. We model this with the energy function

$$E(l) = \sum_{p \in P} D_p(l_p) + \sum_{p,q \in N} V(l_p, l_q). \qquad (2)$$

The data term is based on the predicted class certainty $\mathbf{s}_p(l_p)$ (score for predicted class l at keypoint p)

$$D_p(l_p) = 1 - \mathbf{s}_p(l_p) \qquad (3)$$

and the discontinuity cost

$$V(l_p, l_q) = \min\left[|l_p - l_q|, 1\right] \qquad (4)$$

is given by the difference in labels of two keypoints p and q in the four-connected neighborhood. The discontinuity cost is truncated at 1 (assuming integer labels) to only penalize different labels, but not to prefer any class over another. This cost function is minimized using efficient belief propagation [7] and the smoothed labeling l^* is returned.

The final step is the interpolation of the sparse results to get per biomass pixel predictions in the same resolution as the input image. We use nearest neighbor interpolation to assign to each biomass pixel the smoothed label of the nearest keypoint. This label image I_{LABEL} is the output of the plant classification system. From this image connected crop / weed regions can be determined and a weeding tool can use this information to selectively treat weeds.

4. Training of the System

The acquisition of labeled training data and the training of the classifier are done offline.

During labeling an expert user is shown the masked NDVI images and asked to provide ground truth labels. Users provide labels by drawing polygons and assigning a single class label to each polygon using a web based tool [17]. The label of the polygon is then assigned to all biomass pixels that are enclosed by the polygon. The number of classes is defined in this step and more than two classes are supported to e.g. define multiple weed classes.

The classifier is trained in supervised mode with a pool of labeled images. For each image I with ground truth labels L our system also stores the extracted feature vectors $\{\mathbf{f}\}$ in a database. During training the database is queried to output feature vectors with corresponding ground truth labels. The query can be restricted to run only on a subset of all images in the database (e.g. for cross-validation). Using these feature vector and label pairs a Random Forest classifier [4] is trained. The trained classifier is stored and can be loaded into the pipeline to run the plant classification system online on new images.

5. Evaluation of Results

The plant classification system is evaluated using image data captured with the field robot in June 2013 on a commercial organic carrot farm. The dataset comprises 70 non-overlapping images of plants growing under field conditions. All images contain multiple plants, all plants are approximately of the same size and both inter and intra class overlap is present. All weeds grow close to crop (weeds further away from crop were removed with non precision weed control methods before image acquisition) and need to be treated to avoid loss of yield.

These images were labeled by an expert user using the labeling tool. For this dataset three classes were defined: One crop class for the carrot plants and two weed classes. Chamomile is a very common weed that looks similar to carrots (both true leaves are pinnate) and was labeled separately. The third class was used to label all other weeds. Table 2 summarizes the number of plants (including partial plants when overlap is present) and the resulting number of patches for the different classes in the dataset.

Table 2: Summary of key figures of dataset used for evaluation (Cham. is short for Chamomile).

	Carrot	**Weed**	**Cham.**	**Total**
Number of Plants	168	143	60	371
Number of Patches	5668	6055	3802	15525

For the evaluation all parameters of the pipeline are set to the described values. Each Random Forest is set to grow 100 trees and the number of features sampled for splitting at each node is set to the default value (square root of number of features: $\sqrt{15}$).

The plant classification pipeline is evaluated by passing images from the dataset through the pipeline. We employ a leave one out cross-validation scheme: Each image is selected as test image once. For each test image a classifier is trained using all other images from the dataset as training data. The test image is then run through the pipeline and the ground truth of the test image is used to evaluate the performance of the system.

Figure 4 shows the resulting ROC curves after leave one out cross-validation over all 70 images. The ROC curves are generated using the score vector output of the classifier prior to smoothing.

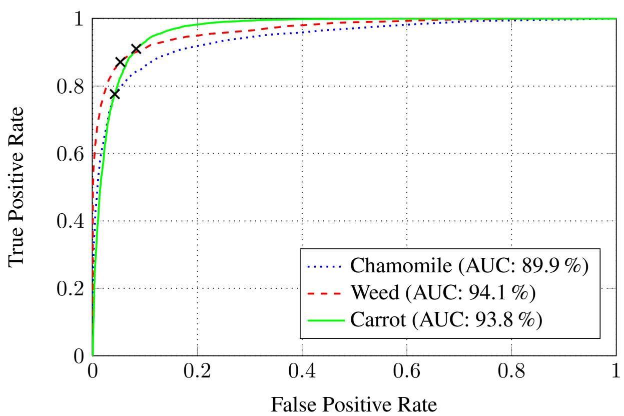


Figure 4: Plant classification ROC curves after leave one out cross-validation on the complete dataset. The crosses mark the chosen operating points. The multi-class ROC curves are generated in one-vs.-all mode.

Table 3 gives the associated true positive rates and false positive rates for the three classes at the respective operating points.

Table 3: True positive rates and false positive rates at selected operating points on ROC curves in Figure 4.

Class	True Positive Rate	False Positive Rate
Carrot	91.0 %	8.3 %
Chamomile	77.6 %	4.2 %
Other Weed	87.1 %	5.3 %

To quantitatively evaluate the performance of the complete system with smoothing the average accuracy, precision, recall and f-score are determined before and after smoothing (see Table 4). After smoothing only labels but no scores are available and thus an ROC curve evaluation is not applicable. Smoothing improves the classification result in all performance measures.

The results of the classification pipeline can also be assessed visually. Figure 5 shows the input NIR image (column a), the color coded expert labeled ground (column b) and the color coded prediction by our pipeline (column c) side by side for four test images.

Table 4: Performance of plant classification system on dataset before and after smoothing. The measures are determined using leave one out cross-validation and are macro averaged [20] over all classes.

	Precision	Recall	F-score	Accuracy
No smoothing	86.8 %	86.4 %	86.6 %	91.5 %
After smoothing	90.4 %	89.9 %	90.2 %	93.8 %
Improvement	+3.6 %	+3.5 %	+3.6 %	+2.3 %

Smoothing is an important step to improve performance and to receive a more consistent labeling. Figure 6 shows the internal sparse classification results at the keypoints prior to interpolation (for test images #1 and #2 from Figure 5). The predictions are overlaid on top of the image before smoothing (left image) and after smoothing (right image). The classifier certainty (score) at each keypoint is visualized by plotting the dot in different sizes (larger for higher certainty) in the image before smoothing. After smoothing only labels (either carrot, chamomile or weed) are available and all dots are plotted in the same size. The class label is color coded as before.

6. Discussion

Our contribution is a new plant classification system with the application of crop / weed discrimination in commercial crop fields. The evaluation in Section 5 shows that the system achieves an average accuracy of 93.8 % when it is applied to field images where weeds grow close to crop, are of the same size and plants overlap.

Both of the two main approaches in related work do not handle this situation well. First, the plant / leaf segmentation approach fails when plants overlap. Reliable segmentation is an unsolved challenge and classification accuracy significantly decreases when overlap is present. Second, the approach of classifying large non-overlapping cells can partially cope with overlap, but there is a large loss in output accuracy. The output estimates are only available per cell and not in pixel space. This is not desired for precise mechanical treatment.

Our approach solves these issues as it neither requires plant / leaf segmented input data nor does it output only coarse per cell predictions. Another advantage is that in our method overlap is not a special case. There are no special parameters to tune to cope with overlap, the system generically handles field situations with and without overlap.

The feature extraction process at sparse keypoints combined with the smoothing and interpolation steps are the key contributions. The evaluation proves that crop / weed classi-

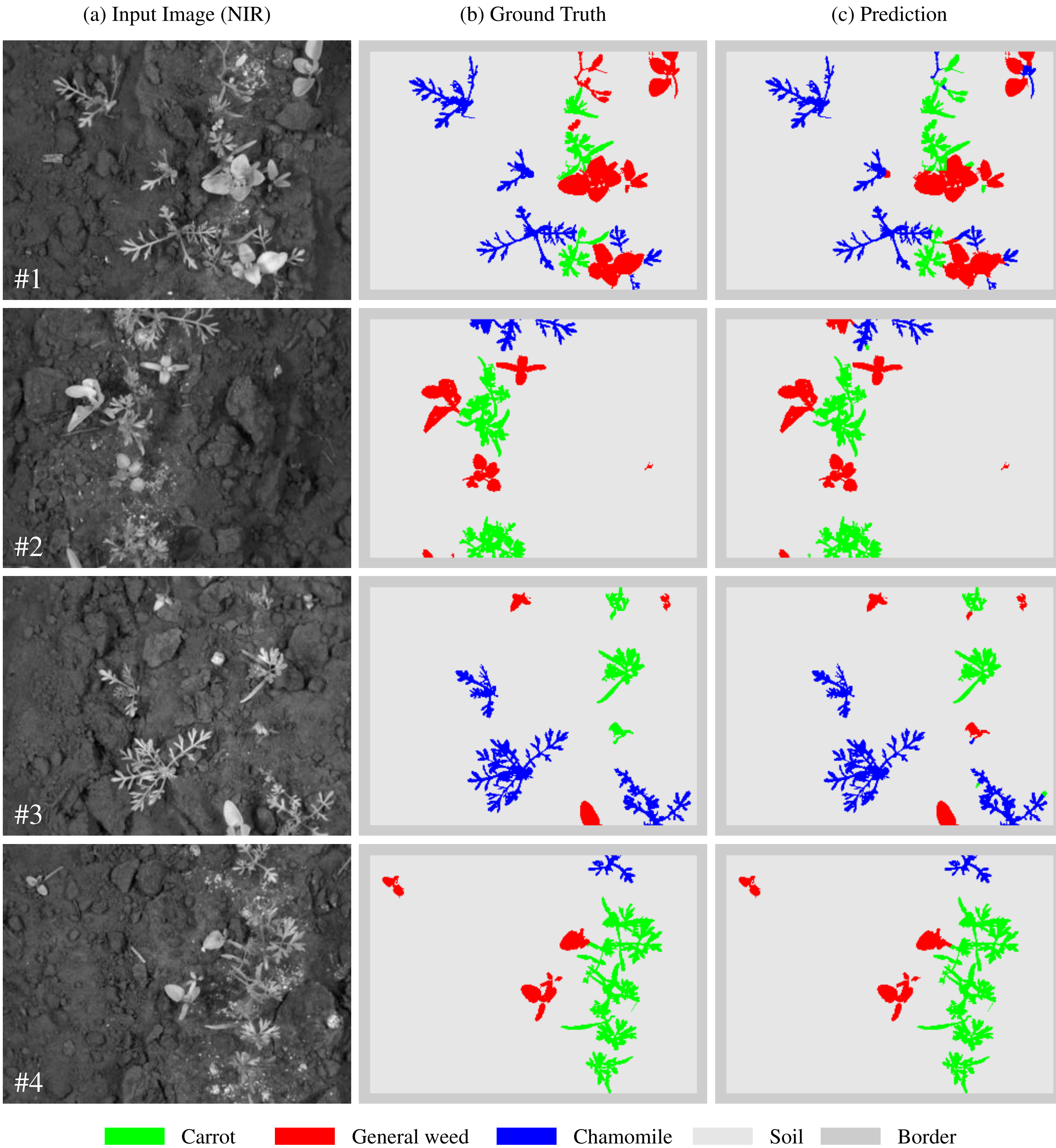


Figure 5: Results of the plant classification system generated using leave one out cross-validation. Column (a) shows the input image, column (b) the expert labeled ground truth and column c) the prediction by the plant classification system. Best viewed in color.

fication on real image data from a commercial farm is possible with our method and that high accuracies are achieved.

The output of our system is a labeled image that can be used for selective weed treatment. Additionally, the pixel labels can be used to calculate metrics like weed coverage or the crop / weed area ratio, that help farmers when applying

(a) Sparse classification results *before* smoothing (b) Sparse classification results *after* smoothing

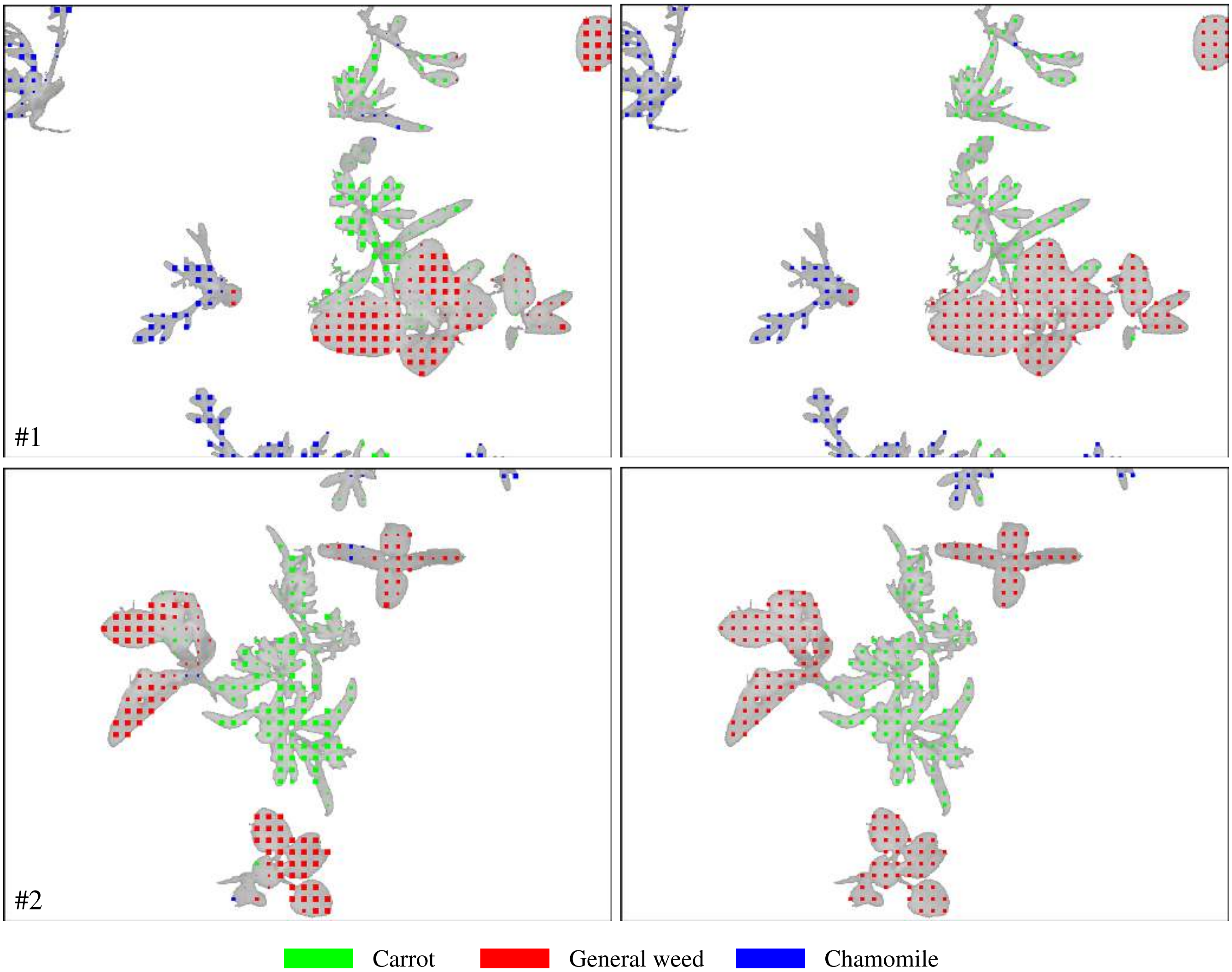


Figure 6: Result of spatial smoothing for the center part of test images #1 and #2. The dot is plotted at the keypoint location, the predicted class is color coded and in (a) the dot size represents the certainty of the prediction. See Figure 5 for ground truth data. Best viewed in color.

precision agriculture farm management techniques.

One limitation of our method is that multiple plants of the same class that overlap (intra class overlap) are not split into different plant regions in the output label image. They get represented by one connected component. However, this is no drawback for the goal of precision weed control. Here the overlap between different classes (inter class overlap) is most important.

In the future, the spatial arrangement of crop and weed in the field could be used as a priori information. Row crops are cultivated in one or multiple parallel straight rows. The lower probability of crops outside rows can be fused into the output to further improve the results.

7. Conclusions

A plant classification system for crop / weed discrimination that does not require segmentation into individual plants is presented. Features are instead extracted on large overlapping image patches representing the neighborhood of sparse keypoints arranged in a grid in image space. The per patch classification results are spatially smoothed using a Markov Random Field. Per pixel crop / weed predictions in full image resolution are derived from the smoothed keypoint results using nearest neighbor interpolation. The proposed method is designed to deal with real world field situations where crops and weed grow close together and plants overlap. The output label image of the plant classification system can be used to control a precision weeding tool to treat the weeds selectively.

To analyze the system a dataset of images was captured in an organic carrot farm under commercial field conditions. The performance is analyzed by testing our system with all images in leave one out cross-validation mode and comparing the output with expert labeled ground truth. Visual analysis indicates good results, empirically our system achieves an average accuracy of $91.5\,\%$ before smoothing that increases when applying smoothing by $+2.3\,\%$ to $93.8\,\%$.

In the future, the arrangement of the crops growing in rows can be leveraged to further improve the classification results. Additionally, the complete robotic system comprising this classification system and a treatment unit will be deployed and compared with manual weed control.

Acknowledgments

The project RemoteFarming.1 is partially funded by the German Federal Ministry of Food, Agriculture and Consumer Protection (BMELV).

References

[1] M. Aitkenhead, I. Dalgetty, C. Mullins, A. McDonald, and N. Strachan. Weed and crop discrimination using image analysis and artificial intelligence methods. *Computers and Electronics in Agriculture*, 39(3):157–171, 2003. 2

[2] B. Åstrand and A.-J. Baerveldt. An agricultural mobile robot with vision-based perception for mechanical weed control. *Autonomous robots*, 13(1):21–35, 2002. 2

[3] T. Beghin, J. S. Cope, P. Remagnino, and S. Barman. Shape and texture based plant leaf classification. In *Advanced Concepts for Intelligent Vision Systems*, pages 345–353. Springer, 2010. 2

[4] L. Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001. 4

[5] S. Christensen, H. T. Søgaard, P. Kudsk, M. Nørremark, I. Lund, E. Nadimi, and R. Jørgensen. Site-specific weed control technologies. *Weed Research*, 49(3):233–241, 2009. 2

[6] J.-X. Du, X.-F. Wang, and G.-J. Zhang. Leaf shape based plant species recognition. *Applied mathematics and computation*, 185(2):883–893, 2007. 2, 3

[7] P. F. Felzenszwalb and D. P. Huttenlocher. Efficient belief propagation for early vision. *International journal of computer vision*, 70(1):41–54, 2006. 4

[8] J. M. Guerrero, G. Pajares, M. Montalvo, J. Romeo, and M. Guijarro. Support vector machines for crop/weeds identification in maize fields. *Expert Systems with Applications*, 39(12):11149–11155, 2012. 2

[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. 2, 3

[10] P. J. Komi, M. R. Jackson, and R. M. Parkin. Plant classification combining colour and spectral cameras for weed control purposes. In *Industrial Electronics, 2007. ISIE 2007. IEEE International Symposium on*, pages 2039–2042. IEEE, 2007. 2

[11] N. Kumar, P. N. Belhumeur, A. Biswas, D. W. Jacobs, W. J. Kress, I. C. Lopez, and J. V. Soares. Leafsnap: A computer vision system for automatic plant species identification. In *Computer Vision–ECCV 2012*, pages 502–516. Springer, 2012. 2

[12] A. Michaels, A. Albert, M. Baumann, U. Weiss, P. Biber, A. Kielhorn, and D. Trautz. Approach towards robotic mechanical weed regulation in organic farming. *Autonomous Mobile Systems 2012*, pages 173–181, 2012. 1

[13] J. C. Neto, G. E. Meyer, and D. D. Jones. Individual leaf extractions from young canopy images using gustafson–kessel clustering and a genetic algorithm. *Computers and Electronics in Agriculture*, 51(1):66–85, 2006. 2

[14] C. M. Onyango and J. Marchant. Segmentation of row crop plants from weeds using colour and morphology. *Computers and electronics in agriculture*, 39(3):141–155, 2003. 2

[15] N. Otsu. A threshold selection method from gray-level histograms. *Automatica*, 11(285-296):23–27, 1975. 3

[16] A. Ruckelshausen, P. Biber, M. Dorna, H. Gremmes, R. Klose, A. Linz, F. Rahe, R. Resch, M. Thiel, D. Trautz, et al. Bonirob – an autonomous field robot platform for individual plant phenotyping. *Precision agriculture*, 9:841, 2009. 1

[17] B. C. Russell, A. Torralba, K. P. Murphy, and W. T. Freeman. Labelme: a database and web-based tool for image annotation. *International journal of computer vision*, 77(1-3):157–173, 2008. 4

[18] I. Scotford and P. Miller. Applications of spectral reflectance techniques in northern european cereal production: a review. *Biosystems engineering*, 90(3):235–250, 2005. 3

[19] D. Slaughter, D. Giles, and D. Downey. Autonomous robotic weed control systems: A review. *Computers and electronics in agriculture*, 61(1):63–78, 2008. 1

[20] M. Sokolova and G. Lapalme. A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45(4):427–437, 2009. 5

[21] 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. 2

[22] K. Thorp and L. Tian. A review on remote sensing of weeds in agriculture. *Precision Agriculture*, 5(5):477–508, 2004. 2