

Weed Segmentation from Grayscale Tobacco Seedling Images

Petre Lameski^(✉), Eftim Zdravevski, and Andrea Kulakov

Faculty of Computer Science and Engineering, University of Saints Cyril
and Methodius in Skopje, Skopje, Macedonia
{`petre.lameski,eftim.zdravevski,andrea.kulakov`}@`finki.ukim.mk`

Abstract. Manual weed extraction from young seedlings is a hard manual labour process. It has to be continuously performed to increase the yield per land unit of any agricultural product. Precise segmentation of plant images is an important step towards creating a camera sensor for weed detection. In this paper we present a machine learning approach for segmenting weed parts from images. A dataset has been generated using bumblebee camera under various light conditions and subsequently training and test patches were extracted. We have generated various texture-based descriptors and used different classification algorithms aiming at correctly recognizing weed patches. The results show that in a case when the images are grayscale, the light conditions are varying, and the distance of the camera to the weeds is not constant machine learning algorithms perform poorly.

Keywords: Weed control · Image processing · Machine learning · Precision agriculture

1 Introduction

According to [15] the agricultural production increasing trend is insufficient to reach the population demand in the future. One of the ways to increase the production is to increase the per land unit yield since the land units available for production are limited. Automation and introduction of Robots in Agriculture is not a new concept [3], and could help in the food production efficiency while at the same time, reducing the environmental impact of the food production in agriculture.

The interest about using autonomous robots in weed control has increased in recent years [16]. Weed control is a process of eliminating undesired plants from the plantations by using chemical compounds or removing them mechanically. It has been proven that precise appliance of chemical compounds decreases the amount of pesticides and makes the products healthier [11] by using special end effectors. Both mechanical and chemical weed control require precise localization of the plants that need to be eliminated. There are several existing approaches towards using machine learning and image processing for weed control.

In [12] the authors propose a robust technique for plant segmentation from land using normalized green index. NDVI index for images is used in [8] to segment land from plants and build their descriptors to classify different types of plants (both known and weed plants). There are existing datasets that contribute towards the automated weed detection from images. In [9] authors present Near infra red (NIR) and visible RED images from carrot plantation taken under controlled light. Authors use the NDVI index to segment green areas and use machine learning approach to classify weed and usable plants.

There are many machine learning approaches towards plant images classification including several competitions [7] where machine learning approaches have proven to give promising results towards plant recognition from images [5, 18]. To precisely localize the weeds in images in order to be able to aid the control of a weeding robot, it is important to segment the image to weed and non-weed regions. In this paper we are using images from low resolution camera under various light conditions. The approach uses several texture patch features to classify patches as weed or not weed.

The problem assessed in this dataset is the problem of segmenting closely seeded plants in seedling plantations as they are traditionally seeded in the agricultural area of Prilep in Macedonia.

The main contribution of our work is the dataset and the initial results from the segmentation using several texture descriptors and feature selection processes to reduce the dimensionality. The goal of the presented work was to evaluate the possibility of using cheap cameras as sensors under variable light conditions and to train a classifier able to detect weed regions from those images.

2 Dataset and Algorithms

2.1 Dataset

The dataset is consisted of 66 randomly selected images taken under various light conditions and various heights from a cluttered tobacco seedling plantation. The dataset is then split to 45 train and 21 test images. The images are then labelled manually for their content: tobacco seedlings or land as one class and weed as another class.

From the images we extracted a total 160000 patches with size of 65×65 pixels that were used for training and from the test images 50000 patches used for testing. To obtain a balanced dataset we tuned the number of randomly selected patches to 80000 from each class for training, and 25000 from each class for testing.

The images used were converted to grayscale since we were interested in the texture characteristics for the segmentation process. The patches were labelled based on the central pixel. If the pixel belongs to a weed its label was marked as weed, else it was marked as non weed patch. Figure 1 shows an example of the images used for the dataset where the traditionally closely placed seedlings in the plantation can be observed.

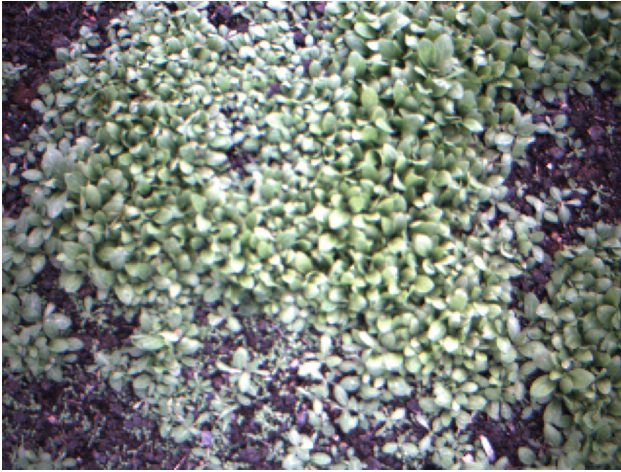


Fig. 1. Example image from the dataset containing tobacco seedlings

2.2 Texture Descriptors

To generate the classification data we used several patch descriptors.

First we generate an image of differences between each pixel from the patch and the central pixel and we generate a histogram of these differences. Then a Histogram of Oriented Gradients [4] (HOG) descriptor was calculated from the grayscale image and also on the difference image, and a HOG descriptor was calculated on the regular grayscale patch. The hog descriptor requires patches with size of 65×65 , so we ignored the rightmost column and bottommost row from each patch when generating the HOG descriptor. For the HOG descriptor we used the OpenCV [2] implementation with the following parameters: windows size 64, block size 16, block stride 8, cell size 8 and number of bins 9. The descriptor is generated for the window so for each block a different histogram is generated. Then the blocks are moved with overlapping. The idea behind the descriptor is that objects can be recognized based on the local intensity gradients distributions without the knowledge of their exact location.

We also generate a Local binary pattern [13] (LBP) histogram. Local binary pattern takes N pixels in radius R around the central pixel and calculates whether the pixel has a smaller or larger value than the central pixel. If the value is larger the bit is set to 1 and otherwise it is set to 0. All the bits are then encoded in a single number. For the LBP descriptor we used the scikit-image implementation [17] with radius of 1 and 8 points. This gives an 8 bit number of different variations of ones and zeros that can be encoded in 256 different values.

The descriptors are chosen because of their good overall performance in texture classification. The total length of the descriptor used is 3784, 256 from the histogram of differences, 256 from the LBP histogram and 1636 each for the HOG descriptor on the grayscale and the HOG descriptor on the difference image.

2.3 Pre-processing and Classification

The length of the descriptor is reduced by using ensembles of trees. First it uses Random Forest (RF) for evaluating initial dataset score, but more importantly feature importance. It is used with the default parameters, as we have noticed that tuning them does not improve the performance dramatically (unlike SVMs for example). As a wrapper approach for evaluating feature subsets during the feature engineering and elimination loop, the framework uses Extremely Randomized Trees (ERT) [6], as they are significantly faster (over 50%) than Random Forests (RF), especially when the number of features is larger. They also provide estimate for feature importance, however it is not as stable as the feature importance estimated by the Random Forests.

Because we are using different training and test datasets, we also want to avoid including features that easily over-fit to the training dataset. Therefore we use an approach based on the idea presented in [1]. For each feature we calculate two scores - one for its importance related to the target class and one for its ability to predict whether the patch comes from the train or the test set. Ideally the reduced dataset will have features that are good predictors of the target class, but bad predictors of the dataset. Using a grid search we discard some features and get optimal dataset. The implementation for Extremely Randomized Trees used is the Python implementation included in scikit-learn [14]. The algorithm creates an ensemble of classifiers from random subsets from the dataset and from the features.

3 Results and Conclusion

The results on the test set show that the best precision of 52.57% and AUC ROC of 0.535 was obtained using the ERT with selected total of 845 features. It is important to note that similar performance was obtained with both ERT and RF when using only the top 6 features (52.41% for RF and 52.39% for ERT).

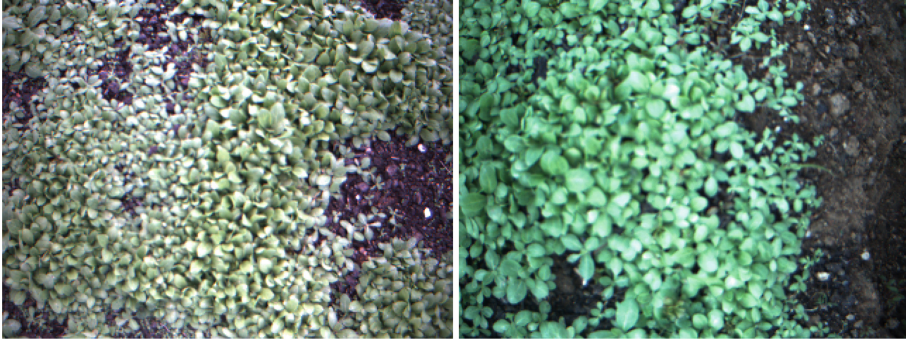
There are several observations that are important to note regarding the generated features. After the analysis of the importance of each feature, none of the features generated from the HOG or the LBP descriptor were in the top features. The most important feature characteristic found based on the dataset was the first value of the difference histogram that actually represents the number of pixels that have the same value as the other pixels in the patch and all of the top 10 features were based on the difference histogram from the patch.

Table 1 shows the top 2 classification schemes for both ERT and RF with the number of descriptor values selected and the resulting classification performance.

The results show that it is difficult to distinguish the weed patches from tobacco patches based on grayscale image patches using texture descriptors. The texture descriptors failed to emphasise the difference between the classes. One of the reasons for this difficulty could be the various light conditions in which the images were taken as it can be seen on two of the images in Fig. 2.

Table 1. Classification performance on the dataset

Classifier	Number of features	Training time [s]	Accuracy %	AUC ROC
ERT	845	117.96	52.57	0.5348
RF	459	83.59	52.47	0.5382
RF	6	0.97	52.42	0.5122
ERT	6	0.93	52.39	0.5106

**Fig. 2.** Images taken under various light conditions and different heights

In order to improve the accuracy several things can be considered. The images could be acquired from a higher resolution camera in order to be able to distinguish the leaves better. On the current data even a human distinguishes with difficulty the weed areas on some images. The images can be taken using multi-spectral camera in order to obtain the best light frequencies that can be used to classify the different species of plants on the plantation. The NDVI index generated from the difference between NIR and visible light shows promising results so it could be used for generating new datasets using adequate cameras.

Another approach would be using deep learning to automatically generate the descriptors and perform the classification [10]. Considering other types of classifiers could improve or degrade the results by certain percentage but it will not improve the overall classification significantly, since during the preprocessing phase we consider only the features that give the highest information gain and are the best choice for the characterization of the classes. We are confident that images in different colour spectre should be considered for the task of weed detection.

It is important to note that any system that will be used for weed detection and removal should be tuned to have as small as possible false positive rate since a false positive would identify a useful plant seedlings as weed. This must be done so that the detection system will be viable for usage on real agricultural plantations and on real commercial systems.

Acknowledgment. The work presented in this paper was partially funded by the University of Sts. Cyril and Methodius in Skopje, Faculty of Computer Science and Engineering

References

1. Boull, M.: Tagging fireworkers activities from body sensors under distribution drift. In: 2015 Federated Conference on Computer Science and Information Systems (FedCSIS), pp. 389–396, September 2015
2. Bradski, G.: The opencv library. *Dr. Dobb's Journal of Software Tools* (2000)
3. Cosmin, P.: Adoption of artificial intelligence in agriculture. *Bull. Univ. Agric. Sci. Vet. Med. Cluj-Napoca* **68**(1), 1–10 (2011). Agriculture
4. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR 2005, vol. 1, pp. 886–893. IEEE (2005)
5. Dimitrovski, I., Madjarov, G., Kocev, D., Lameski, P.: Maestra at lifeclef 2014 plant task: Plant identification using visual data. In: CLEF (Working Notes), pp. 705–714 (2014)
6. Geurts, P., Ernst, D., Wehenkel, L.: Extremely randomized trees. *Mach. Learn.* **63**(1), 3–42 (2006)
7. Goëau, H., Joly, A., Bonnet, P., Selmi, S., Molino, J.F., Barthélémy, D., Boujemaa, N.: Lifeclef plant identification task 2014. In: CLEF2014 Working Notes for CLEF 2014 Conference, Sheffield, UK, 15–18 September 2014, pp. 598–615. CEUR-WS (2014)
8. Haug, S., Michaels, A., Biber, P., Ostermann, J.: Plant classification system for crop/weed discrimination without segmentation. In: 2014 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 1142–1149, March 2014
9. Haug, S., Ostermann, J.: A Crop/weed field image dataset for the evaluation of computer vision based precision agriculture tasks. In: Agapito, L., Bronstein, M.M., Rother, C. (eds.) ECCV 2014. LNCS, vol. 8928, pp. 105–116. Springer, Heidelberg (2015). doi:[10.1007/978-3-319-16220-1_8](https://doi.org/10.1007/978-3-319-16220-1_8)
10. Hung, C., Xu, Z., Sukkarieh, S.: Feature learning based approach for weed classification using high resolution aerial images from a digital camera mounted on a uav. *Remote Sens.* **6**(12), 12037–12054 (2014)
11. Jeon, H.Y., Tian, L.F.: Direct application end effector for a precise weed control robot. *Biosyst. Eng.* **104**(4), 458–464 (2009)
12. Jeon, H.Y., Tian, L.F., Zhu, H.: Robust crop and weed segmentation under uncontrolled outdoor illumination. *Sensors* **11**(6), 6270–6283 (2011)
13. Mäenpää, T.: The Local binary pattern approach to texture analysis: Extensions and applications. Oulun yliopisto (2003)
14. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E.: Scikit-learn: machine learning in python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011)
15. Ray, D.K., Mueller, N.D., West, P.C., Foley, J.A.: Yield trends are insufficient to double global crop production by 2050. *PloS one* **8**(6), e66428 (2013)
16. Slaughter, D., Giles, D., Downey, D.: Autonomous robotic weed control systems: A review. *Comput. Electron. Agric.* **61**(1), 63–78 (2008)

17. van der Walt, S., Schönberger, J.L., Nunez-Iglesias, J., Boulogne, F., Warner, J.D., Yager, N., Gouillart, E., Yu, T.: The scikit-image contributors: scikit-image: image processing in Python. *Peer J.* 2, e453 (2014). <http://dx.doi.org/10.7717/peerj.453>
18. Yanikoglu, B., Tolga, Y., Tirkaz, C., FuenCaglartes, E.: Sabanci-okan system at lifeclef 2014 plant identification competition. In: Working Notes of CLEF 2014 Conference (2014)