Efficient Detection of Colorado Potato Beetles in Ultra-High-Resolution UAV-Imaging of Potato Fields using Convolutional Neural Networks

Jana Wieme
Department of Plants and Crops, Ghent University
Coupure links 653, 9000 Ghent, Belgium
Technology and Food Unit, ILVO
Burg. van Gansberghelaan 115, 9820 Merelbeke, Belgium
jana.wieme@ugent.be
jana.wieme@ilvo.vlaanderen.be

Robin Mortier
Department of Information Technology, Ghent University
Technologiepark-Zwijnaarde 126, 9052 Ghent, Belgium
robmorti.mortier@ugent.be

Sam Leroux
Department of Information Technology, Ghent University
Technologiepark-Zwijnaarde 126, 9052 Ghent, Belgium
sam.leroux@ugent.be

Simon Cool
Technology and Food Unit, ILVO
Burg. van Gansberghelaan 115, 9820 Merelbeke, Belgium
simon.cool@ilvo.vlaanderen.be

Jan Pieters
Department of Plants and Crops, Ghent University
Coupure links 653, 9000 Ghent, Belgium
jan.pieters@ugent.be

Wouter Maes
Department of Plants and Crops, Ghent University
Coupure links 653, 9000 Ghent, Belgium
wouter.maes@ugent.be

1. Introduction

In recent decades, the Colorado potato beetle (*Leptinotarsa decemlineata*) has emerged as a major agricultural pest, causing significant damage to potato crops worldwide. It is widely regarded as one of the most invasive insect herbivores, with annual costs for ongoing management rising to tens of millions of dollars and even billions of dollars unmanaged [10]. Traditionally, fields are managed homogeneously. Precision agricultural fields, in contrast, are divided into management zones, with each zone receiving adjusted inputs at the right time [4]. Therefore, to establish the appropriate treatment within such zone, timely detection of Colorado potato beetle infestation is essential. Especially taking into account that the beetle has shown a spectacular ability to adapt and quickly develop insecticide resistance [10]. The overuse of insecticides, and insufficient use of alternative control methods (e.g., crop rotation, biological controls), even appear to accelerate this trend of insecti-

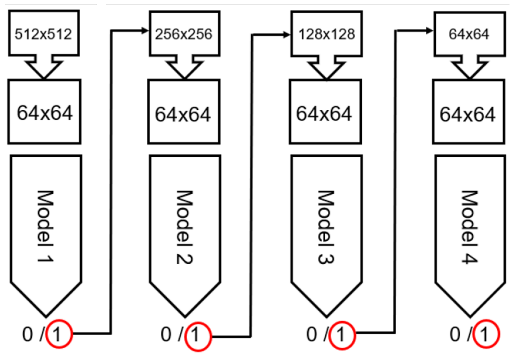

Figure 1. High-level representation of the architectural pipeline.

cide resistance [1].

New technological developments have tremendously improved the possibilities to collect data using Unmanned Aerial Vehicles (UAVs) [7]. In particular, ultra-high-resolution UAV-imaging has shown promise in agricultural applications due to their potential to collect detailed information of large areas. UAV-imagery thus provides the opportunity to introduce a game changer in detecting Colorado potato beetles, without crop disturbances in the field. Currently, deep learning has become a state-of-the-art technique to process huge amounts of data in various precision agriculture tasks [6]. While this concept has been used for generic pest detection, limited research exists on the use of deep learning for Colorado potato beetle detection specifically. Hunt et al. [5] used high-resolution UAV data of potato fields to relatively detect crop damage due to Colorado potato beetles by using NDVI thresholds, object-based image analysis and plant height. However, they concluded that the latter two methods require extensive operator intervention for success and further research is needed for more accurate detection. Roldan-Serrato et al. [9] proposed a recognition system for Colorado potato beetles based on a neural network (random subspace classifier) using a small dataset of 25 images from the Internet. In addition, Barbedo et al. [2] showed that there is a growing interest in image-based plague detection capabilities, and identified some challenges for the further development and practical applicability. On the one hand, more comprehensive image databases are needed given the high variability in agricultural data (i.e., in environmental conditions, in stages of the insect life cycle). On the other hand, the aspect of reacting adequately fast to possible pest invasions should not be disregarded. It is often difficult to perform pest monitoring in relatively short time frames, especially when high-resolution data is required to detect small specimens.

The contributions of this research are twofold. First, we demonstrate the potential of ultra-high-resolution UAV imagery combined with Convolutional Neural Networks (CNNs) as a game-changer in Colorado potato beetle detection. The proposed pipeline is trained and tested on an in-field dataset of labeled images with different life stages of the Colorado potato beetle to enhance the applicability. Second, a crucial aspect of our study involves developing an efficient pipeline to expedite the processing of these ultra-high-resolution images, making our approach more accessible and practical. Therefore we consider the detection problem as a binary classification problem in combination with a multi-step approach.

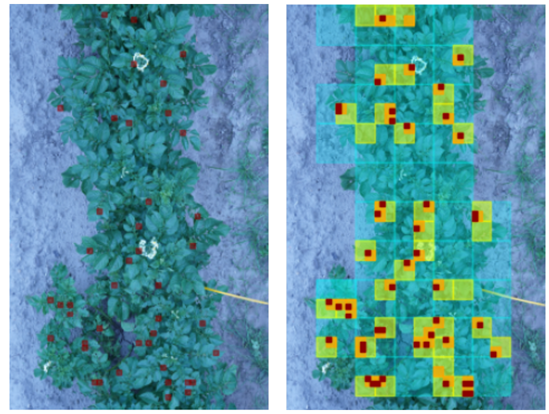

Figure 2. On-image visualization of ground truth (left) and result of the different models in the pipeline (right). Blue boxes indicate 512x512 patches proceeded to the second level, green boxes are 256x256 patches proceeded to the third level, orange boxes are 128x128 patches proceeded to the final level and red boxes indicate patches of 64x64 in which a beetle/larvae is detected by the model.

2. Materials and Methods

A dataset of 80 images, captured by a DJI M600 drone equipped with a Sony Alpha 7 III RGB camera with 135 mm lens (Zeiss) at an altitude of 10 m, was collected on the 26th of June 2019 at a trial field of PCA in Kruisem, Belgium. Originally the images had a size of 7952x5304 pixels, but these were cropped to images of 3000x5304 pixels as there was only 1 row of potato plants on each image. The ground truth is labeled by domain experts by manually drawing bounding boxes around every beetle or larvae on the cropped images. To evaluate the applicability of the method, this labeling was converted to binary classified patches of 64x64 pixels, each of which was cut from the original image. If the center of a bounding box is in a patch, the patch was labeled positive. Further labeling of new datasets can immediately be executed with this strategy, greatly reducing the labeling effort and improving usability. The labeled dataset is very unbalanced, as there are a lot more patches (97.5%) without beetles or larvae, due to the nature of these data. During training, class weights were used to address the imbalance. The division between train and test data is 70-30.

The baseline model is a binary classification model, comprising of three convolutional layers followed by two

	Classifier	Pipeline
Precision	0.213	0.702
Recall	0.939	0.526
F1 (@ 0.5)	0.347	0.601
Inference time (s)	2.910	0.236

Table 1. The results of the baseline classifier and the pipeline.

dense layers. The model uses the naïve approach where the high-resolution image is split into patches of 64x64 and processed one by one through the model. In contrast, the proposed pipeline consists of four sequential binary classification models. These models all have the same internal CNN architecture and are connected by a filtering layer, as shown in Figure 1. First, the input image is cut into parts of 512x512 pixels, which are then downscaled to a resolution of 64x64. Second, all images of 64x64 that correspond to the 512x512 data are processed through the first model, determining if they are relevant for further examination. Third, the irrelevant images are discarded based on an automatic threshold (score < 0.5), saving only the images possibly containing a beetle/larvae. Next, the remaining data of 512x512, that corresponds with the selected 64x64 data, is again cut into four parts, resulting in 256x256 patches. These are once more down-scaled to 64x64 and then used as input for the second model. This principle is further repeated through all models in the pipeline, and so the output of the last model is the final classification result. The system thus only uses relevant data further down the pipeline and always in a resolution of 64x64, drastically reducing the computational cost. The PyTorch framework [8] is used to train the models on an NVIDIA Geforce GTX1060.

3. Results

Table 3 summarizes three performance metrics: precision, recall and F1 score when using 0.5 as threshold for the binary classification at each step for both the binary classifier and pipeline. In addition, the inference time is also included, as it provides an important perspective on the efficiency of the detection pipeline. This inference time was calculated excluding the time for downscaling, since this depends on the method used. However, when using the resize function with bilinear interpolation of OpenCV [3], this amounts to only 4.99 milliseconds for a complete image. The left image in Figure 2 shows the original ground truth. The righter images visualize the result of the four levels in the pipeline, with all colored patches per level being further processed to the next one.

4. Discussion

The results show that the pipeline has a better overall performance, and a drastic increase in speed as it clearly filters irrelevant parts. The rather low recall for the pipeline indicates that further optimization is needed. By analyzing the variation in the threshold per level, the balance between precision and recall can be further optimized. The F1 score of the pipeline is significantly higher than that of the classifier, though. In addition, when further analyzing the results, the number of beetles on the image is found to have an impact on the inference time of the model. Therefore, this can be intuitively explained by the fact that with more beetles/larvae, more relevant patches also need to be processed. Averaged over the entire test set, however, the inference time in Table 3 shows that the pipeline is still much faster. In conclusion, our research proposes a novel pipeline that promises more efficient detection of Colorado potato beetles in potato fields using ultra-high-resolution UAV-imagery and CNNs. Still, the detection can be further improved by optimizing all levels; this will be the focus of future work.

References

[1] Andrei Alyokhin, David Mota-Sanchez, Mitchell Baker, William E. Snyder, Sandra Menasha, Mark Whalon, Galen Dively, and Wassem F. Moarsi. The red queen in a potato field: integrated pest management versus chemical dependency in colorado potato beetle control. *PEST MANAGEMENT SCIENCE*, 71(3):343–356, MAR 2015. 2

[2] Jayme Garcia Arnal Barbedo. Detecting and classifying pests in crops using proximal images and machine learning: A review. *AI*, 1(2):312–328, 2020. 2

[3] G. Bradski. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 2000. 3

[4] C. Hedley. The role of precision agriculture for improved nutrient management on farms. *Journal of the Science of Food and Agriculture*, 95(1):12–19, 2015. 1

[5] E. Hunt and Silvia Rondon. Detection of potato beetle damage using remote sensing from small unmanned aircraft systems. *Journal of Applied Remote Sensing*, 11:026013, 05 2017. 2

[6] A. Kamilaris and F. X. Prenafeta-Boldú. Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, 147:70 – 90, 2018. 2

[7] Wouter H. Maes and Kathy Steppe. Perspectives for remote sensing with unmanned aerial vehicles in precision agriculture. *Trends in Plant Science*, 24(2), 2019. 2

[8] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. Pytorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems 32*, pages 8024–8035. Curran Associates, Inc., 2019. 3

[9] Lucero Roldán-Serrato, Tetyana Baydyk, Ernst Kussul, Alberto Escalante-Estrada, and María Teresa González Rodriguez. Recognition of pests on crops with a random sub-

space classifier. In *2015 4th International Work Conference on Bioinspired Intelligence (IWOBI)*, pages 21–26, 2015. 2

[10] Sean D. Schoville, Yolanda H. Chen, Martin N. Andersson, Joshua B. Benoit, Anita Bhandari, Julia H. Bowsher, Kristian Brevik, Kaat Cappelle, Mei-Ju M. Chen, Anna K. Childers, Christopher Childers, Olivier Christiaens, Justin Clements, Elise M. Didion, Elena N. Elpidina, Patamarerk Engsontia, Markus Friedrich, Inmaculada Garcia-Robles, Richard A. Gibbs, Chandan Goswami, Alessandro Grapputo, Kristina Gruden, Marcin Grynberg, Bernard Henrissat, Emily C. Jennings, Jeffery W. Jones, Megha Kalsi, Sher A. Khan, Abhishek Kumar, Fei Li, Vincent Lombard, Xingzhou Ma, Alexander Martynov, Nicholas J. Miller, Robert F. Mitchell, Monica Munoz-Torres, Anna Muszewska, Brenda Oppert, Subba Reddy Palli, Kristen A. Panfilio, Yannick Pauchet, Lindsey C. Perkin, Marko Petek, Monica F. Poelchau, Eric Record, Joseph P. Rinehart, Hugh M. Robertson, Andrew J. Rosendale, Victor M. Ruiz-Arroyo, Guy Smagghe, Zsofia Szendrei, Gregg W. C. Thomas, Alex S. Torson, Iris M. Vargas Jentzsch, Matthew T. Weirauch, Ashley D. T. Yates, George D. Yocum, June-Sun Yoon, and Stephen Richards. A model species for agricultural pest genomics: the genome of the colorado potato beetle, leptinotarsa decemlineata (coleoptera: Chrysomelidae). *SCIENTIFIC REPORTS*, 8, JAN 31 2018. 1