https://doi.org/10.17221/131/2022-PPS

Verification of a machine learning model for weed detection in maize (*Zea mays*) using infrared imaging

Adam Hruška, Pavel Hamouz*

Faculty of Agrobiology, Food and Natural Resources, Czech University of Life Sciences Prague, Prague, Czech Republic

**Corresponding author: Hruskaa@af.czu.cz*

Citation: Hruška A., Hamouz P. (2023): Verification of a machine learning model for weed detection in maize (*Zea mays*) using infrared imaging. Plant Protect. Sci., 59: 292–297.

Abstract: The potential of the framework of precision agriculture points towards the emergence of site-specific weed control. In light of the phenomena, the search for a cost-effective approach can help the discipline to accelerate the practical implementation. The paper presents a near-infrared data-driven machine learning model for real-time weed detection in wide-row cultivated maize (*Zea mays*) fields. The basis of the model is a dataset of 5 120 objects including 18 species of weeds significant in the context of wide-row crop production in the Czech Republic. The custom model was subsequently compared with a state-of-the-art machine learning tool You only look once (version 3). The custom model achieved 94.5 % identification accuracy while highlighting the practical limitations of the dataset.

Keywords: computer vision; NIR images; machine learning; visual analysis; neural networks

In a historical perspective, the strong economic and security pressures have been instrumental in the development of chemical crop protection. Today, however, the supply role is enhanced by the growing need for sustainable production methods (Pretty 2008). A potential solution is to make more efficient use of the available chemical inputs in precision agriculture (Gebbers & Adamchuk 2010). The current refinement of applications includes the use of aerial and satellite photography for spectral reflection-based weed detection, as well as precise navigation systems for the implementation of output application maps (López-Granados 2011). The emerging phenomenon of machine learning driven object detection can potentially achieve a higher success rate, while being able to detect earlier growth stages of weeds. The core of the technology lies in the use of convolutional layers, which, though analysis of pixel-based shapes, detect pretrained objects (O'Mahony et al. 2020).

The initial decade of machine learning driven weed detection development offered the first commercially applied projects (Gerhards et al. 2002). Broader success in the implementation of the technology followed the publication by Dyrmann et al. (2016) who produced a model with an identification accuracy of 86.2%, while being able to successfully distinguish between 22 species of weeds. In the following years the discipline split into two major groups, focusing on either land-based or aerial platforms. Liakos et al. (2018) highlighted the aerial approach that excelled in large scale data collection, which was counterbalanced by the higher precision and detail of the land-based platforms (Wang et al. 2019). The use of multispectral, or hyperspectral images allowed for an increase in the model accuracy in terms of the identification of plants from the surrounding environment (Potena et al. 2017; Fawakherji et al. 2021), it also paved the way for the broader inclusion of alternative detection ap-

https://doi.org/10.17221/131/2022-PPS

proaches based on both the shape and the spectral reflection of the plants (Farooq et al. 2019). The latest models generally have an output accuracy near the 95 % benchmark, such as in the case of Farooq et al. (2019), Alam et al. (2020) or Junior and Ulson (2021).

The main objective of the paper is the development and evaluation of an analytical tool capable of detecting weeds under field conditions based on near-infrared visual data. The core of the analysis consists of a subgoals, including the collection of suitable data and the subsequent building and training of a machine learning algorithm, which is to be later compared with a state-of-the-art computer vision tool the You Only Look Once (version 3) (YOLOv3) model by Redmon and Farhadi (2018), trained on identical data.

MATERIAL AND METHODS

The cornerstone of the analysis is the acquisition of a relevant and sufficiently large dataset. In this case, the near-infrared (NIR) spectral plane was chosen, the selection of which was primarily motivated by the ease of any background elimination as well as the price of the device. Conceptually, the analysis worked with a ground-based platform, allowing for the higher resolution imaging of target objects due to their close proximity (Wang et al. 2019). The majority of the images were taken at the experimental plots of the Czech University of Life Sciences in Prague, during April and May 2021 to fit the growth phase ranging between the first and the fifth true leaf. In total, 18 species of weeds were included in the dataset (Table 1 and Table 2) including a single instance of a monocot species. The weed species were selected by their relevance to the local production of *Zea mays*, being the single crop in the dataset. The taxonomic nomenclature of the plants follows Kaplan et al. (2019).

The imaging itself was carried out using a modified Panasonic G5 camera, which underwent the removal of the internal IR-blocking filter and replacing it with an external Hoya R72 filter that blocks visible radiation up to a wavelength of 720 nanometres. The camera was equipped with a Panasonic G Vario 14–45 mm f/3.5–5.6 lens. Data collection was carried out in the form of images taken from a distance of 50 cm perpendicular to the soil surface at a lens focal length of 45 mm.

During the manual selection of the collected data, out-of-focus and otherwise poor-quality images

Table 1. The overview of species and the number of samples with identification success and error rates per weed species in custom model

Weed species	No. of samples	Succesful classification (%)	Type I error (%)	Type II error (%)	Missidentification (%)
Amaranthus retroflexus	170	98.3			1.7
Capsella bursa-pastoris	250	97.5			2.5
Chenopodium album	440	98.9		1.1	
Echinochloa crus-galli	231	89.6	10.4		
Fumaria officinalis	133	84.9		6.0	9.1
Galium aparine	420	98.6		1.4	
Galinsoga parviflora	260	97.4		2.6	
Lamium amplexicaule	480	99.2		0.8	
Tripleurospermum inodorum	130	80.8	3.8	7.7	7.7
Mercurialis annua	260	97.4			2.6
Papaver rhoeas	200	97.5			2.5
Polygonum aviculare	134	98.5			1.5
Fallopia convolvulus	360	98.8	1.2		
Stellaria media	385	98.7			1.3
Thlaspi arvense	145	65.2	8.7	8.7	17.4
Veronica persica	154	97.4	2.6		
Viola arvensis	417	98.8			1.2
Persicaria maculosa	286	97.9		2.1	
Zea mays	265	100.0			

https://doi.org/10.17221/131/2022-PPS

Table 2. The overview of species and the number of samples with identification success and error rates per weed species in You only look once (version 3) model

Weed species	No. of samples	Succesful classification (%)	Type I error (%)	Type II error (%)	Missidentification (%)
Amaranthus retroflexus	170	98.9	1.1		
Capsella bursa-pastoris	250	97.2		2.8	
Chenopodium album	440	99.2			0.8
Echinochloa crus-galli	231	98.2	1.8		
Fumaria officinalis	133	97.1	1.5	1.5	
Galium aparine	420	98.6		0.7	0.7
Galinsoga parviflora	260	98.4		1.6	
Lamium amplexicaule	480	99.0			1.0
Tripleurospermum inodorum	130	97.7	0.5		1.8
Mercurialis annua	260	98.4		0.8	0.8
Papaver rhoeas	200	95.2	1.6		3.2
Polygonum aviculare	134	97.8		1.1	1.1
Fallopia convolvulus	360	99.5		0.5	
Stellaria media	385	98.7		1.3	
Thlaspi arvense	145	92.3	1.5		6.2
Veronica persica	154	98.2		0.9	0.9
Viola arvensis	417	98.5	1.5		
Persicaria maculosa	286	97.6	2.4		
Zea mays	265	100.0			

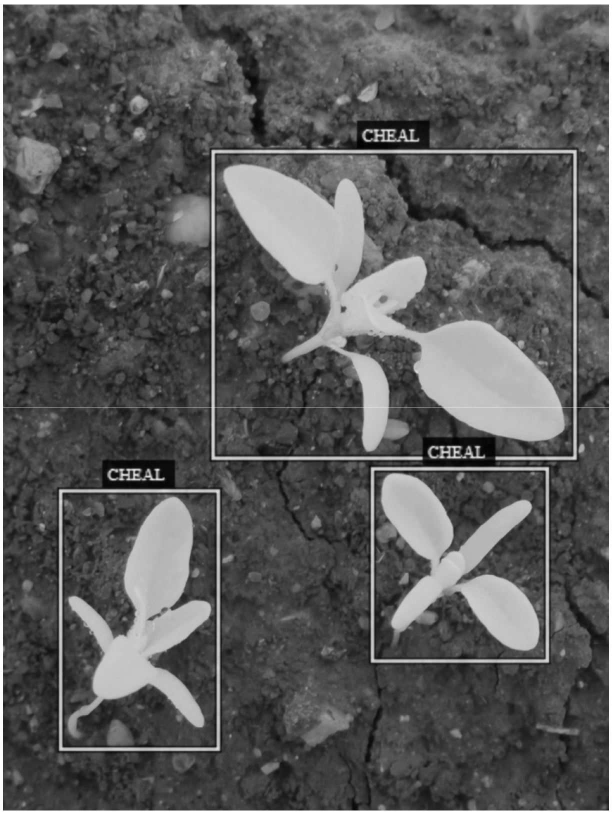

Figure 1. Example of annotated image of *Chenopodium album*

were eliminated. Annotation of the objects was performed using the "open source" computer software VIA from the Visual Geometry group (Figure 1). The individual annotations were formatted into a JavaScript Object Notation (JSON) or alternatively YOLO file, designated for each model.

The multiplication of the total number of images was undertaken through the axis rotation duplication method, where randomly selected images were adjusted by rotating them by ninety degrees (Farooq et al. 2019). The partitioning of the dataset into training, evaluation and test parts was oriented towards the dominance of the training set, thus maximising the identification success rate (Miikkulainen et al. 2019). Therefore, 70% of the annotated images were randomly selected and assigned for training, 10% for the model evaluation, and the remaining 20% of the annotated images were used for testing the detection accuracy.

Individual convolutional layers in the architecture were complemented by output maximisation layers (Maxpool), whose output data contain only the highest value of the input parameters being monitored (Dyrmann et al. 2016). To effectively keep the range of inputs to each layer constant while avoiding the effect of the previous output updates,

https://doi.org/10.17221/131/2022-PPS

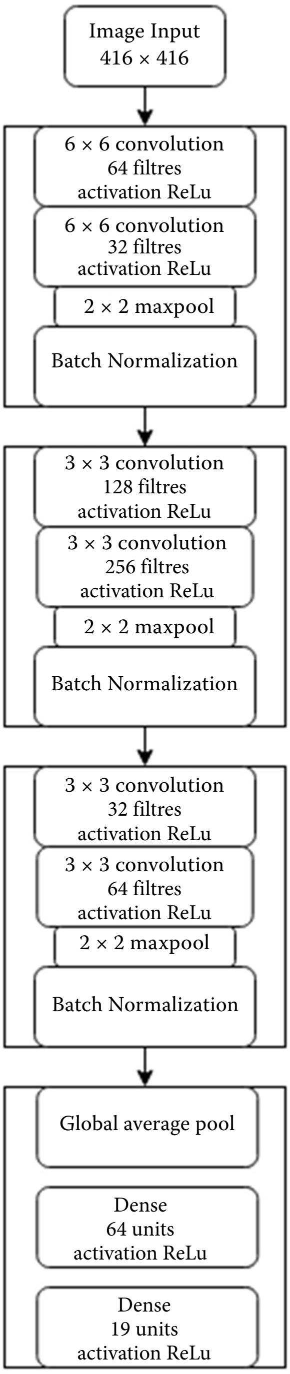

Figure 2. Architecture of the custom neural network

the model contained batch normalisation layers (Ioffe & Szegedy 2015). The global averaging (Global Average pool) served the model to average the output data from the previous segment and direct it to the pair of output layers (Figure 2). A final pair of "matrix-vector multiplication" layers completed the work of the previous neurons by assigning the data outputs to the corresponding category vectors.

The output results of the testing on 1 024 images containing 5 120 objects were subject to a statistical analysis based on the conflict or correspondence of the outputs with the annotations. While the core element of the identification success rate was the percentage of the correctly identified objects from the total plant objects present in the images consisting of the two primary factors (detection and classification), the unsuccessful elements included type I (false detection) and type II (non-detection of the object) errors as well as the misidentification of the classes (Table 1 and Table 2).

RESULTS

Custom model. The training achieved an identification success rate of 94.96% at epoch 150, which was subsequently verified to 94.5% by testing on the rest of the prepared data. Overall, the training development was able to obtain a relatively high object identification accuracy on the data. However, in terms of the performance, the actual model was slightly slower than expected, taking 28 ms to process an image.

From the individual species, *Lamium amplexicaule* was correctly detected in 99.2% of the cases. *Chenopodium album* came in second at 98.9% and, in a close third place, *Fallopia convolvulus* was detected in 98.8% of the cases. At the other end of the rankings, *Thlaspi arvense* dominated the error rate, with the model being correct only 65.2% of the time. The case of *Echinochloa crus-galli* was also responsible for a rather specific result within the model. On the one hand, the model managed to detect individuals quite accurately with an identification success rate of 89%, but, on the other hand, the species was responsible for virtually all the cases of the type I error in both the evaluation and test datasets due to it being mistakenly detected in the examples of the post-harvest residue. The identification accuracy and identification errors for all the weed species are listed in Table 1.

YOLOv3. In the case of the YOLOv3 model, a range of 150 training epochs was set for the training in parallel to the custom model. The identification success rate of the last training epoch of the model was equal to 98.45%. The final identification accuracy level was validated on the test dataset, where the model reached a 97.92% identification success rate (Table 2). During the testing process, similar to the

https://doi.org/10.17221/131/2022-PPS

custom model, type I error cases occurred, but the proportion of occurrences was lower. In terms of the speed, the YOLOv3 model performed very well indeed, achieving a process time of 22.8 ms per frame.

Images containing *F. convolvulus* were the most accurately detected in the model, with an average identification success rate of over 99.5%. The model also noticeably exceeded the 99% identification success rate threshold for the *Ch. album* images. Most of the remaining classes had between 98% and 99% an identification success rate, with the only exception being the combination of *Papaver rhoeas* and *T. arvense*. For both species, there was the relatively frequent confusion and misidentification in the test dataset, resulting in the overall identification success rates for both species being below 90% (Table 2). In addition to the occurrence of type I and type II errors, species confusion occurred in the test dataset when the model assigned an object to the wrong class, which, despite the smaller number of cases, highlights the limitations of the training dataset.

DISCUSSION

The primary goal of training the YOLOv3 model on the identical data was to validate the applicability of the gathered data for the weed detection. A second and complementary goal served to validate the performance of the custom model in competition with a contemporary model in practical use. The basic comparable benchmark of both models remains their architecture and scope. The initial element that comes directly from the architecture is the speed in the data processing, where YOLOv3 clearly wins, as it is the top contemporary model in terms of speed. When comparing 22.7 ms with the 28 ms of the actual model, the result is quite clear, but in the overall context, a difference of 5 ms can be considered as a relative success.

Despite the fact that, in absolute terms, the resulting difference of 3.32% is not staggering, in terms of the identification accuracy of the machine learning outputs, it is quite a significant difference (Wang et al. 2019). In practical terms, 95% can be considered as a crucial threshold of accuracy (Nawara 2010), which unfortunately was not surpassed by the custom model. The YOLOv3 model, on the other hand, reached the desired threshold with a relatively large margin. Overall, however, already exceeding the 90% identification success rate threshold can be considered an achievement due to the demonstration of the stability of the model for further development and improvement.

A relatively specific element of the comparison is the segmentation of the classification success rate in terms of the classes of the observed objects. Although the exact reason for the different results for the individual elements of the dataset is very difficult to trace, it is possible to argue that, within the analysis, it primarily occurs by the differences in the used filters. The similarity when *F. convolvulus* and *T. arvense* formed practical counterparts in the identification success rate in both models can be explained by two reasons: The first is the disparity in the volume of the training data, with individuals of *T. arvense* forming a minority. Secondly, the general suitability of the shape features of both species can be questioned. There are clearly more or less suitable shapes for the identification by machine vision techniques (Dyrmann et al. 2016). The pixel-scale detection of some species performed better with the smaller filters used in YOLOv3, while, for others, the combination of the larger filters of the custom model was a better fit. In addition to the pixel-scale detection, the size of the filters of the custom model could limit its use in highly weed-infested fields due to the overlay of multiple objects and the resulting detection of the one on top.

While the current iteration of the model could not be widely applied due to its relatively weak dataset base, the potential scaling of the technology by both means of the price and data gathering suggest a potential for practical use. With the restraints on the gathered data, the model can only be used in the predefined environment, but, given additional training, it might potentially work even in narrow-row cultivated crops. Generally, the model has the potential to be implemented in the form of autonomous robotic weeding platforms and a retrofitted sensing platform for the target control of conventionally used sprayers.

CONCLUSION

In spite of the number of problematic factors, it is possible to conclusively confirm the practical success of the primary objective. The results were used to verify the applicability of the infrared spectrum dataset for weed detection, as well as to verify the capabilities

https://doi.org/10.17221/131/2022-PPS

of the custom model in direct comparison with the YOLOv3 architecture. Despite a relatively promising level of identification accuracy, the algorithm remains limited by the scope of the data, which decides the reliability of the model. To successfully incorporate the technology into the reality of the agrarian sphere, the model would need to include more weed species and a larger source of visual data.

Acknowledgement: We thank to Ing. Jakub Lev Ph.D. for his valuable help as well as the Czech University of Life Sciences for tools and space necessary for the completion of the analysis.

REFERENCES

Alam M., Alam M.S., Roman M., Tufail M., Khan M.U., Khan M.T. (2020): Real-time machine-learning based crop/weed detection and classification for variable-rate spraying in precision agriculture. In: 7th International Conference on Electrical and Electronics Engineering, April 14–16, 2020, Antalya, Turkey: 273–280.

Dyrmann M., Karstoft H., Midtiby H.S. (2016): Plant species classification using deep convolutional neural network. Biosystems Engineering, 151: 72–80.

Farooq A., Jia X., Hu J., Zhou J. (2019): Multi-resolution weed classification via convolutional neural network and superpixel based local binary pattern using remote sensing images. Remote Sensing, 11: 1692. doi: 10.3390/rs11141692

Fawakherji M., Potena C., Pretto A., Bloisi D.D., Nardi D. (2021): Multi-spectral image synthesis for crop/weed segmentation in precision farming. Robotics and Autonomous Systems, 146: 103861. doi: 10.1016/j.robot.2021.103861

Gebbers R., Adamchuk V.I. (2010): Precision agriculture and food security. Science, 327: 828–831.

Gerhards R., Sökefeld M., Nabout A., Therburg R.D., Kühbauch W. (2002): Online weed control using digital image analysis. Journal of plant diseasesand protection, 18: 421–427.

Ioffe S., Szegedy C. (2015): Batch normalization: Accelerating deep network training by reducing internal covariate shift. In: International Conference on Machine Learning. PMLR, 37: 448–456.

Junior L.C.M., Ulson J.A.C. (2021): Real time weed detection using computer vision and deep learning. In: 2021 14th IEEE International Conference on Industry Applications, September 10, 2021, Sao Paulo, Brasil: 1131–1137.

Kaplan Z., Danihelka J., Chrtek J., Kirschner J., Kubát K., Štech M., Štěpánek J. (eds.), 2019. Klíč ke květeně České republiky [Key to the flora of the Czech Republic], Academia, Praha.

Liakos K.G., Busato P., Moshou D., Pearson S., Bochtis D. (2018): Machine learning in agriculture: A review. Sensors, 18: 2674. doi: 10.3390/s18082674

López-Granados F. (2011): Weed detection for site-specific weed management: mapping and real-time approaches. Weed Research, 51: 1–11.

Miikkulainen R., Liang J., Meyerson E., Rawal A., Fink D., Francon O., Raju B., Shahrzad H., Navruzyan A., Duffy N., Hodjat B. (2019): Evolving deep neural networks. Artificial intelligence in the age of neural networks and brain computing. Elsevier Academic Press: 293–312.

Nawara J. (2010): Machine learning: face recognition technology evidence in criminal trials. University of Louisville Law Review, 49: 601.

O'Mahony N., Campbell S., Carvalho A., Harapanahalli S., Hernandez G.V., Krpalkova L., Riordan D., Walsh J. (2020): Deep learning vs. traditional computer vision. In: Advances in Computer Vision: Proceedings of the 2019 Computer Vision Conference, April 25–26, 2019, Las Vegas, USA: 128–144.

Potena C., Nardi D., Pretto A. (2017): Fast and accurate crop and weed identification with summarized train sets for precision agriculture. In: International Conference on Intelligent Autonomous Systems. Springer, Cham: 105–121.

Pretty J. (2008): Agricultural sustainability: concepts, principles and evidence. Philosophical Transactions of the Royal Society B: Biological Sciences, 363: 447–465.

Redmon J., Farhadi A. (2018). Yolov3: An incremental improvement. ArXiv preprint. arXiv:1804.02767

Wang K., Gao X., Zhao Y., Li X., Dou D., Xu C.Z. (2019): Pay attention to features, transfer learn faster CNNs. International conference on learning representations. doi: 10.48550/arXiv.1804.02767

Recieved: December 12, 2022
Accepted: June 6, 2023
Published online: August 2, 2023