Turkish Journal of Agriculture - Food Science and Technology, 10(8): 1441-1446, 2022
DOI: https://doi.org/10.24925/turjaf.v10i8.1441-1446.5183

Turkish Journal of Agriculture - Food Science and Technology

Available online, ISSN: 2148-127X | www.agrifoodscience.com | Turkish Science and Technology Publishing (TURSTEP)

# How to Affect the Number of Images on the Success Rate for Detection of Weeds with Deep Learning

**Guzel Mustafa[1,a,*], Turan Bulent[2,b], Kadioglu İzzet[3,c], Sin Bahadir[4,d], Basturk Alper[5,e], Khaled R. Ahmed[6,f]**

[1]*Department of Plant, Soil, and Agricultural Systems, Southern Illinois University, Carbondale, IL, United States*
[2]*Faculty of Engineering and Natural Science, Tokat Gaziosmanpaşa University, 60250 Tokat, Türkiye*
[3]*Anamur/Mersin, Türkiye*
[4]*Faculty of Agriculture, Sakarya University of Applied Science, 54050 Sakarya, Türkiye*
[5]*Faculty of Engineering, Erciyes University, 38039 Kayseri, Türkiye*
[6]*School of Computing, Southern Illinois University, Carbondale, IL, United States*
[*]*Corresponding author*

A R T I C L E  I N F O

*Research Article*

Received : 24/03/2022
Accepted : 06/05/2022

*Keywords:*
Weed detection
Deep learning
Real time detection
*Sinapis arvensis*
Confusion Matrix

A B S T R A C T

The detection of weeds with computer vision without the help of an expert is important for scientific studies and other purposes. The images used for the detection of weeds are recorded under controlled conditions and used in image processing-deep learning methods. In this study, the images of 3-4-leaf (true-leaf) periods of the wild mustard (Sinapis arvensis) plant, which is the critical process for chemical control, were recorded from its natural environment by a drone. The datasets were included 50-100-250-500 and 1 000 raw images and were augmented by image preprocessing methods. Totally 12 different augmentation methods used and datasets were examined for understand how to affects the numbers of images on training-validation performance. YOLOv5 was used as a deep learning method and results of the datasets were evaluated with the Confusion Matrix, Metrics-Precision, and Train-Object Loss. For results of Confusion Matrix where 1 000 images gave the highest results with TP (True Positive) 80% and FP (False Positive) 20%. The TP-FP ratios of 500, 250, 100 and 50 image numbers were respectively; 65%-35%, 43%-57%, 0%-100% and 0%-100%. With 100 and 50 images, the system did not show any TP success. The highest metrics-precision ratio was found 92.52% for 1 000 images set and for 500 and 250 image sets respectively; 88.34% and 79.87%. The 100 and 50 images datasets did not show any metrics-precision ratio. The minimum object loss ratio was 5% at 50th epochs in the 100 images dataset. This dataset was followed by other 50, 250, 500, and 1 000 images respectively; 5.4%, 6.14%, 6.16%, and 8.07%.

[a] mustafa.guzel@siu.edu https://orcid.org/0000-0002-6069-3561
[b] bulent.turan@gop.edu.tr https://orcid.org/0000-0003-0673-469X
[c] izzetkadioglu@gmail.com https://orcid.org/0000-0002-5080-4424
[d] bahadirsin@subu.edu.tr https://orcid.org/0000-0002-0109-3662
[e] ab@erciyes.edu.tr https://orcid.org/0000-0001-5810-0643
[f] kahmed@cs.siu.edu https://orcid.org/0000-0002-3707-4316



## Introduction

There are many smart agriculture applications such as plant disease detection, crop yield estimation, plant species detection, weed detection, water, and soil protection are already performed with computer vision technology (Tian, et al., 2020) (Mavridou, et al., 2019) (Zhang, et al., 2021). The control of weed assets is one of the important ways to increase agricultural production efficiency. Barberi has suggested sensitive variable spraying methods to avoid the problems caused by the overuse and residues of herbicides used in conventional spraying methods (Bàrberi, 2002). To detect crop plants and weeds in real-time is an important problem to be able to apply the right amount of spraying to the right area. Guzel et al. carried out a study that detected Wild Mustard (*Sinapis arvensis*) by deep learning method for real-time detection of this weed (Güzel, et al., 2021). The detection of plants and weeds by using computer vision technology can be done with traditional image processing and deep learning methods (Wu, et al., 2021). Since the weed detection is done with traditional image processing technology, features such as color, texture, and shape of the image are distinguished and combined with traditional machine learning methods such as Random Forest or Support Vector Machine (SVM) algorithm (Sabzi, et al., 2020). The features of plants appearances used in these methods need to be determined manually. The success rate in these methods depends on the image acquisition process, preprocessing methods, and feature extraction quality. With the improvement in computing power and increase in data volume, deep learning algorithms can extract multi-scale and multidimensional spatial semantic feature information of weeds through Convolutional Neural Networks (CNNs) due to their enhanced data expression capabilities for images, avoiding the disadvantages of traditional methods. For this reason, deep learning methods attract the attention of researchers working on the detection of weeds.

In the literature, there are several reviews on the application of machine learning in agriculture (Liakos, et al., 2018) and some studies presented on the use of deep learning methods to accomplish agricultural application (Kamilaris & Prenafeta-Boldu, 2018). A general compilation study of artificial intelligence methods applied by researchers in all areas of agricultural production (Weng, et al., 2019) and a review that compiles studies currently carried out on a specific type of technology for a specific task (Su, 2020). Koirala et al. conducted a study on the application of deep learning in product detection and yield estimation, problems that hinder the display of products and their solutions (Koirala, et al., 2019). However, this study focused only on crop detection and yield estimation, ignoring other agricultural tasks involving multiple objects such as weed detection. Kamilaris et al. presented a review compiling the application areas of deep learning methods in agriculture, including many studies in the fields of weed identification, land cover classification, plant identification, crop counting, and crop type classification (Kamilaris & Prenafeta-Boldu, 2018). Yuan et al. conducted a study describing the research progress in the field of weed detection at home and abroad and the advantages and disadvantages of various segmentation, extraction, and diagnostic methods (Yuan, et al., 2020). However, there are different studies for new proposals on the use of deep learning methods to solve the problem of weed detection. Hasan et al. has provided a comprehensive review of weed detection and classification research and focused on deep learning studies within these studies (Hasan, et al., 2021).

When these studies are examined, there is no study on the importance of how the number of images used for weed detection affects the prediction rate and deep learning achievements. Therefore, different numbers of data sets were prepared in this study. These datasets have been tested on the same deep learning architecture (YOLOv5) and their results were evaluated.

The Wild Mustard (*Sinapis arvensis* L.) was chosen to determine the effect of image numbers on deep learning predictions. *Sinapis arvensis* is an important weed that is a member of the Brassicaceae family. It is a plant that is especially rich in nutrients, like the basic character, hummus, and clay soils. Although it originates from Mediterranean countries, it is a plant that is reported to be frequently seen in fields, gardens, and pastures (Uygur, 1986). The plant forms capsules after seed flower formation and the existing seeds are formed in these capsules. It is known that a healthy, properly grown plant gives approximately 1 200 seeds. If the seeds that have grown from the plant are not found in suitable conditions, they can remain without germination for a long time (up to about 35 years) (Uygur, 1986); (Şin, 2021). This plant is in the category of invasive plants seen in almost every part of the world. Although it is assumed that wild mustard came to North America from Europe in various studies (Mulligan, 1975) (Rollins, 1981), some archaeological excavations have encountered fossilized wild mustard seeds dating back to pre-Columbian times. Studies in Canada have determined that different wild mustard species are used as medicine and food by the indigenous people of Canada (Arnason, 1981). This plant can cause serious yield losses in field crops in Canadian meadows. A strong persistent seed bank, competitive growth habit, and high fertility all add to the nature of the weed and ensure that weed is an ongoing problem. Before the widespread use of phenoxy herbicides, *S. Arvensis* was the worst weed on grassland plantations (Mulligan, 1975).

The Wild Mustard causes serious economic losses by being found in various cultivated plant growing areas. If an example is given for these losses; It has been determined that there is a product loss of 20% in canola cultivation areas if there are 10 plants per square meter, and 36% in 20 plants $m^{-2}$ (Thomas, 1984) (Blackshaw, 1987) (McMullan, 1994). In wheat-growing areas, according to the researchers, if there are 11 wild mustards per square meter, it is stated that it causes a loss of 49.97%. In addition, various studies have reported that wild mustard populations have gained resistance to the herbicides used (Şin & Kadıoğlu, 2021).

## Material and Methods

The application of herbicides against broad-leaved weeds, especially in the young period with 2-4 leaves (true leaves), is important for successful control. In this study, deep learning classification performance was evaluated with a different number of pictures in the 2-4 leaf period of Wild Mustard and it was coded as YH-2. Images of the natural environment of the Wild Mustard plant used in the study were obtained in the true leaf period. These images were obtained from the wheat field in the province of Tokat, which was reviewed for the year 2020. The locations of the plants, which were determined during the period of cotyledonous leaves, were marked with the Magellan eXplrosit 310 Handheld GPS (Figure 1-a) device with 1m sensitivity, and videos recorded by a drone (the DJI Mavic 2 Pro (Figure 1-b) drone, which has a camera 4K: 3840×2160 30 FPS resolution and quality) when the true leafy period of the plants. Videos of the plants were taken by flying 1 m above the ground. A desktop computer with 11 core Intel ® i7 11700 KF CPU, NVIDIA GeForce RTX 2080 Ti GPU, 32 GB DDR4 3600 MHz RAM, 1 TB Samsung Pro 980 M.2 hardware was used for further processing by transferring the recorded videos to the computer environment.

To make photo frames from these captured videos and tagged them to create the data set, random photo frames must be extracted from the videos. Some distortions or loss of detail occur in the photo frames extracted from the videos because of the movements of the drone during the video recording, the shaking of the plants due to the naturally occurring wind events, and the air pressure created by the drone propellers. The photos should be having as possible as the highest resolution because the deep learning method automatically finds the attributes (edge, color, area, texture, etc.). To minimize these losses and make the images clearer, FrameGui program was used to improve the FPS (Frame per Second) feature of the videos. The frame rate of the videos recorded with the drone has been increased from 30 FPS to 60 FPS. The detail-quality difference between the photo frames taken from the improved videos is given in Figure 2.

a b

Figure 1. (a) 1m precision Handheld GPS device Magellan eXplorist 310 (b) DJI Mavic 2 Pro Drone

a b

Figure 2. (a) Random image in motion at 30fps, (b) Random image from 60 FPS enhanced video

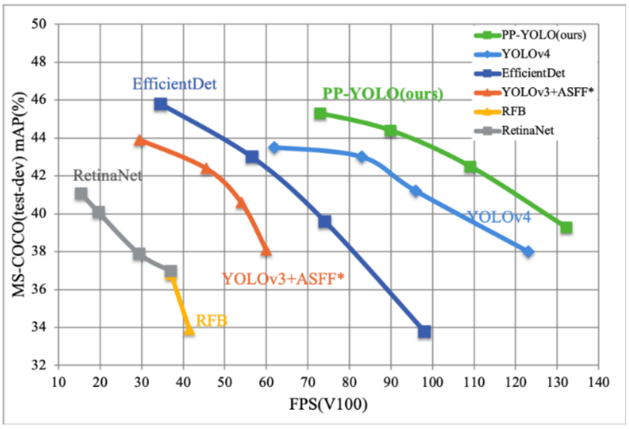



Figure 3. YOLOv5 (PP-YOLOv(ours)) Comparison of Deep Learning Architecture with Other Architectures

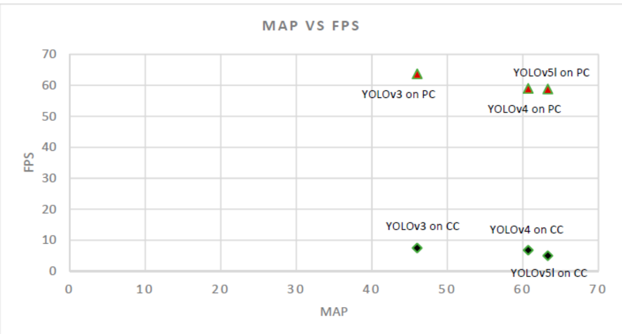



Figure 4. Comparison of YOLOv5 Libraries with Other YOLO Versions

Python 3.9 was used to collect random photo frames from videos with increased frames per second (DeepLabel) to label Wild Mustard (*Sinapis arvensis* L.) images on the obtained photos (LabelImageMaster), and to create training-validation sets. With this data set, the YOLO-v5 library, which has instant object detection, was evaluated. The advantages of the YOLO-v5 library over other deep learning architectures used in the literature are given in Figure 3 and Figure 4 (Nepal & Eslamiat, 2022).

Datasets consisting of 50-100-250-500-1 000 raw images were prepared to evaluate the effects of the number of images in the dataset on the correct detection and prediction rates. These datasets were then preprocessed with data augmentation methods, each of which was increased to 12 folds. The image reproduction methods applied are given in Table 1.

The created 5 different data sets (1 000, 500, 250, 100, 50 pieces) were evaluated by subjecting them to training and validation practices in the YOLOv5 library. To train the system with these datasets, the images in the datasets were randomly distributed as 90% training and 10% validation. For training, the batch size was 12, the number of epochs chosen as 50, YOLOv5s.pt picked for weight, and YOLOv5s.yaml deep learning algorithm were used.

## Results and Discussion

To compare the correct estimation and detection rates of the created data sets; The results of the Confusion Matrix (Figure 5), Metrics-Precision (Figure 6), and Train-Object Loss (Figure 7) were compared. For the Confusion Matrix, the success indicators are True Positive; which correctly classifies the given object, and False Positive means that the object that needs to be detected (YH-2) is estimated as one of the other plants existing in the background.

According to Figure 5, neural networks trained with 1 000 images gave the highest results with TP 80% FP 20%. The TP-FP ratios of 500, 250, 100 and 50 image numbers were respectively; 65%-35%, 43%-57%, 0%-100% and 0%-100%. With 100 and 50 images, the system did not show any TP success.

The test results of the datasets trained with different numbers of images on the validation dataset are given in Figure 6. Accordingly, the highest metrics-precision ratio was 92.52% for 1 000 image sets, while it was 88.34% and 79.87% for 500 and 250 image sets, respectively. Datasets containing 100 and 50 photos could not make any correct predictions and remained at 0%.

In Figure 7, different datasets were utilized from the training-not-evaluated object. Object loss during training was 5% at step 50 in the dataset with at least 100 images. This dataset was followed by others 50, 250, 500, and 1000 images respectively; 5.4%, 6.14%, 6.16%, and 8.07%.

The difference between the training and validation sets of the data sets whose training has been completed is given as an example in Figure 8, with 2 identical photo sets (a) that are tagged and uploaded to the system for validation and that is (b) requested to be labeled automatically by the trained system.

In this study, the results were evaluated by using data sets containing different image numbers, the same deep learning architecture and algorithm, and a computer with the same hardware features. The highest correct prediction rate among the datasets and the success of correctly classifying these predictions belonged to the 1 000 raw image datasets. When compared with other datasets, as the number and diversity of images in the dataset increases, although the data loss rate is higher during the validation phase, the highest and most accurate prediction number was obtained from the dataset containing the highest number of images.

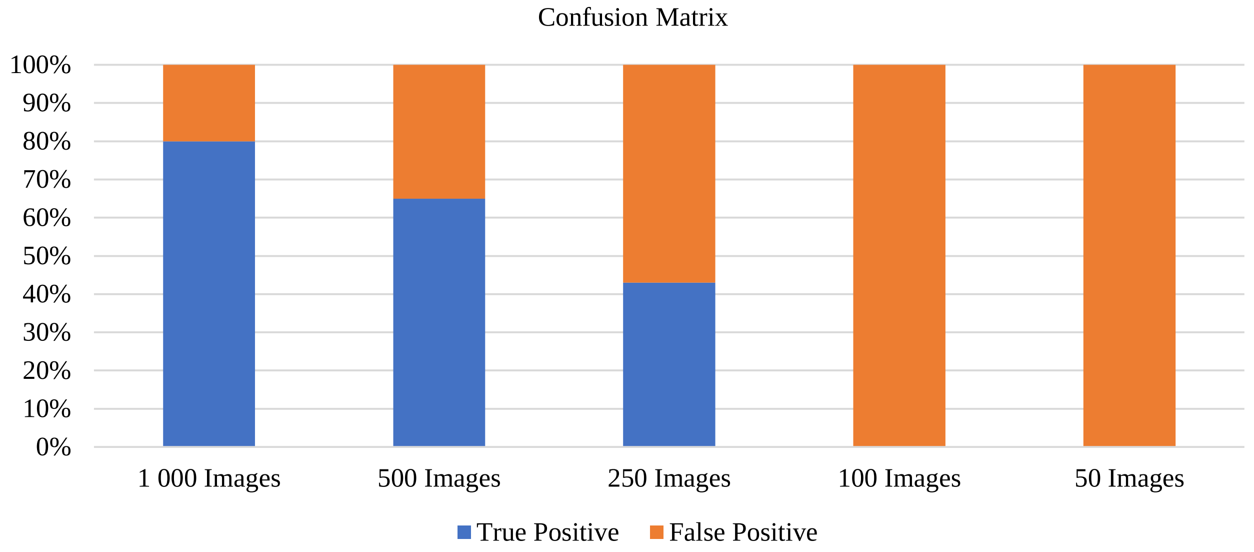



Figure 5. The Confusion Matrix Ratios of Data Sets Containing Different Image Numbers

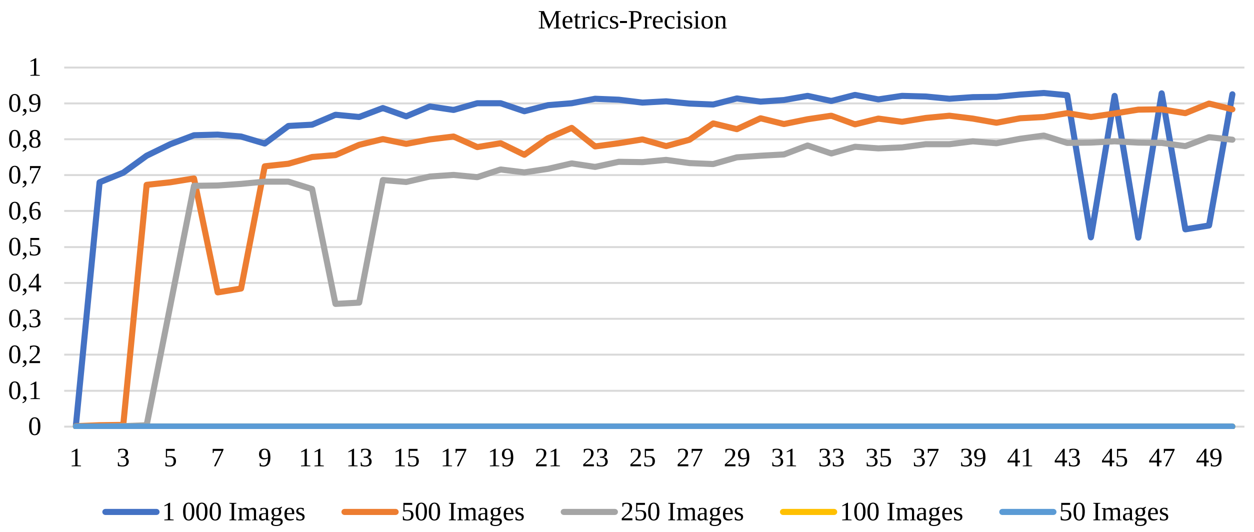



Figure 6. Metrics-Precision Rates of Datasets Containing Different Image Numbers

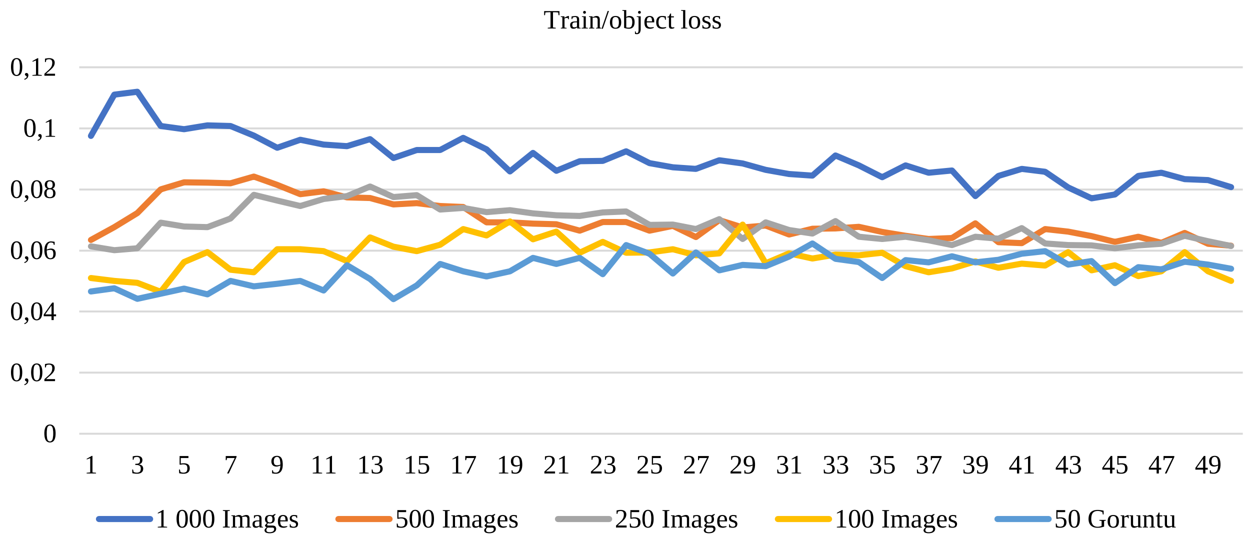
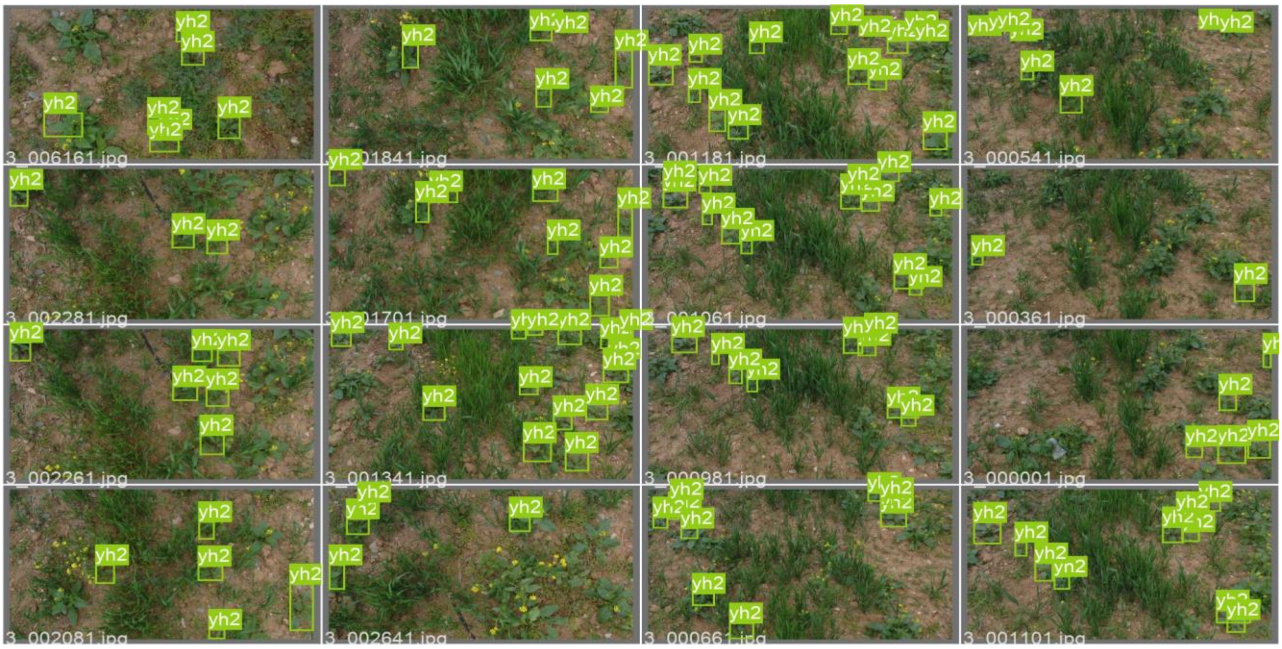



Figure 7. Train-Object Loss Rates of Datasets Containing Different Image Elements



a

b

Figure 8. (a) Ground Truth labeled Wild Mustard plants sample, (b) Predicted boxes around Wild Mustard plants and confidence scores generated by model on the same images in (a).

Table 1. Data augmentation methods and explanations used

| Data Augmentation Methods | Description |
|---|---|
| hsv_h: 0.015 | image HSV-Hue augmentation (fraction) |
| hsv_s: 0.7 | image HSV-Saturation augmentation (fraction) |
| hsv_v: 0.4 | image HSV-Value augmentation (fraction) |
| degrees: 0.0 | image rotation (+/- deg) |
| translate: 0.1 | image translation (+/- fraction) |
| scale: 0.5 | image scale (+/- gain) |
| shear: 0.0 | image shear (+/- deg) |
| perspective: 0.0 range 0-0.001 | image perspective (+/- fraction), range 0-0.001 |
| flipud: 0.0 | image flip up-down (probability) |
| fliplr: 0.5 | image flip left-right (probability) |
| mosaic: 1.0 | image mosaic (probability) |
| mixup: 0.0 | image mixup (probability) |

While the datasets consisted of raw 1 000, 500, 250, 100, and 50 labeled images, the number of images has increased to 12 000, 6 000, 3 000, 1 200, and 600 using data replication methods. With the increase in the number of images in the data set, each repetitive training epoch of the system has taken longer than the dataset with fewer images. The training time was for the datasets of 1 000, 500, 250, 100, and 50 has took for each epoch 18.6, 13.4, 8.2, 5.1 and 2.7 mins respectively. 90% of the CPU and GPU capacities used in the study were used and thus the number of images (batch size) examined at once was 48. Since the training times per epoch are inversely proportional to the batch size, weights with faster and higher prediction rates for larger image sets obtained naturally and/or by data replication methods, on an equipped computer with a higher-capacity graphics card and processor can create.

A significant correlation was found between the increase in the number of images in the datasets and the increase in the prediction precision and correct classification rates of the generated neural networks. The satisfying point of the system should be examined with different numbers of images, epochs, and batches at the training stage reach. To carry out this analysis, datasets containing more images should be obtained and tested on computers with higher capabilities. In this way, the optimum number of raw-amplified images will be known by other researchers at the ten study and planning stages to perform the best estimation process of deep learning applications, which are increasingly used in agricultural areas. To realize the correct and clear use of artificial intelligence algorithms in agricultural areas, the features that reveal the differences can show great changes depending on the climate and environmental conditions, the natural growth periods of the plant, or the occurrence of weather events where image acquisition becomes impossible and may prevent the system from being properly trained. For this reason, it is an important issue for the health of the studies to be taken before the laboratory studies, on time and in the correct number. Studies such as this study will reveal the relationship between the number of images in the data set and the result should be tried on different objects, as they may differ according to the variety of the item used.

## Acknowledgment

The related article is a part of the doctoral thesis named "Detection of Some Weeds in Wheat Production in Tokat Region by Deep Learning Method" and was supported by project number 120O888 within the scope of TUBITAK 3501 Career Development Program and supported by TUBITAK 2214-a Research Scholarship During the PhD Program.

## References

Arnason THRJaJT. 1981. Use of plants for food and medicine by Native Peoples of eastern Canada. Can. J. Bot., Issue 59: 2189-2325.

Bàrberi P. 2002. Weed management in organic agriculture: Are we addressing the right issues? Weed Res., Issue 42: 177-193.

Başaran BKYKİKDTHAM. 2016. The Effect of 2,4-D Acid Dimethylamin Against Broadleaf Weeds Applied at Different Phenological Periods on Grain Yield and Some Yield Components of Common Wheat (Triticum aestivum L.). Turk J Weed Sci, Issue 19(2): 1-9.

Blackshaw REAGWaDJ. 1987. Interference of Sinapis arvensis L. and Chenopodium album L. in spring rapeseed (Brassica napus L.). Weed Res., Issue 27: 207-213.

Güzel M, Şin B, Turan B, Kadıoğlu I. 2021. Real-Time Detection of Wild Mustard (Sinapis arvensis L.) With Deep Learning (YOLO-v3). Fresenius Environmental Bulletin, Issue 30 11/A20201, pp. 12197-12203.

Hasan A. et al. 2021. A survey of deep learning techniques for weed detection from images. Comput. Electron. Agric., Issue 184.

Kamilaris A, Prenafeta-Boldu F. 2018. Deep learning in agriculture: A survey. Comput. Electron. Agric., Issue 147: 70-90.

Koirala A, Walsh K, Wang Z, McCarthy C. 2019. Deep learning Method overview and review of use for fruit detection and yield estimation. Comput. Electron. Agric., Issue 162: 219-234.

Liakos K. et al. 2018. Machine Learning in Agriculture: A Review. Sensors, Issue 18: 2674.

Mavridou E. et al. 2019. Machine Vision Systems in Precision Agriculture for Crop Farming. J. Imaging, Issue 5:89.

McMullan PMDJKaDDR. 1994. Effect of wild mustard (Brassica kaber) competition on yield and quality of triazine-tolerant and triazine-susceptible canola (Brassica napus and Brassica rapa). Can. J. Plant Sci., Issue 74: 369-374.

Mulligan GAaBLG. 1975. The biology of Canadian weeds: Sinapis arvensis L. Can. J. Plant Sci, Issue 55: 171-183.

Nepal U, Eslamiat H. 2022. Comparing YOLOv3, YOLOv4 and YOLOv5 for Autonomous Landing Spot Detection in Faulty UAVs. Sensors, Issue 22: p. 464.

Rollins RC. 1981. Weeds of the Cruciferae (Brassicaceae) in North America. J. Arnold Arbor, Issue 62: 517-540.

Sabzi S, Abbaspour-Gilandeh Y, Arribas J. 2020. An automatic visible-range video weed detection, segmentation and classification. Heliyon, Issue 6.

Şin B. 2021. The determination of tribenuron methyl resistance of wild mustard (Sinapis arvensis L.) collected from wheat fields located in Amasya, Çorum, Tokat and Yozgat province. Tokat: Tokat Gaziosmanpaşa University, Institute of Science and Technology, Ph.D. Thesis.

Şin B, Kadıoğlu İ. 2021. Trp-574-Leu mutation in wild mustard (Sinapis arvensis L.) as a result of als inhibiting herbicide applications. PeerJ, Issue 9: e.

Su W. 2020. Advanced Machine Learning in Point Spectroscopy, RGB- and Hyperspectral-Imaging for Automatic Discriminations of Crops and Weeds: A Review. Smart Cities, Issue 3: 767-792.

Thomas AGaWRF. 1984. Weed survey of Manitoba cereal and oilseed crops 1978, 1979 and 1981. Weed Survey Series Publ., Issue 84-1: 230.

Tian H. et al. 2020. Computer vision technology in agricultural automation—A review. Inf. Process. Agric., Issue 7: 1-19.

Uygur FKWWH. 1986. Definition of Important Weeds in Wheat-Cotton Planting System of Çukurova Region. PLITS, Issue 1986/4(1): 169.

Weng Y. et al. 2019. A survey on deep-learning-based plant phenotype research in agriculture. Scientia Sinica Vitae, Issue 49: 698-716.

Wu Z, Chen Y, Zhao BKX, Ding Y. 2021. Review of Weed Detection Methods Based on Computer Vision. Sensors, Issue 21.

Yuan H, Zhao N, Cheng M. 2020. Review of Weeds Recognition Based on Image Processing. Trans. Chin. Soc. Agric. Mach., Issue 51: 323-334.

Zhang S, Huang W, Wang Z. 2021. Combing modified Grabcut, K-means clustering and sparse representation classification for weed. Neurocomputing, Issue 452: 665-674.