

Weed detection on embedded systems using computer vision algorithms

D. Shadrin^{1,2}, S. Illarionova¹, R. Kasatov^{1,3}, M. Akimenkova¹, G. Rudensky¹, E. Erhan¹

¹Skolkovo Institute of Science and Technology, 143026, Bolshoy Bulvar 42, bldg. 1, Skolkovo, Moscow, Russia;

²Irkutsk National Research Technical University, 664074, Lermantov st. 83, Irkutsk, Russia;

³ITMO University, 197101, Kronverksky Pr. 49, bldg. A, St. Petersburg, Russia

Abstract

Agriculture is a vital component of a sustainable development of many states. It supports economic growth and ensures food security. Therefore, great attention is paid to increasing production efficiency and yields. One of the problems occurring in the agricultural section is weed spreading that can corrupt the quality and amount of yields. To achieve better harvest, weed control measures should be conducted in time. Currently, computer vision techniques are implemented in various areas of industry, in particular, in agriculture. They allow one to automate data analysis process and to make decisions faster. However, the weed detection task in agriculture requires not only high recognition accuracy, but also fast computations on portable devices with low memory availability that makes it possible to embed computer vision systems on unmanned aerial vehicles (UAVs). To address these challenges, we proposed a neural-based approach for real-time weed recognition that combines state-of-the-art detection architectures and optimization techniques for faster inference. To conduct a comprehensive study using real field data, we collected and labelled two unique datasets in Volgograd Region. The experiments involved YOLO, SSD, and Faster R-CNN architectures with inference on NVIDIA Jetson Nano. The highest results were achieved for YOLOv5 architecture with mAP of 0.668 for Carrot Dataset (two weeds classes) and 0.882 for Onion Dataset (one weed class), while inference prediction time equals to 29 FPS and 31 FPS respectively.

Keywords: weed detection, computer vision, deep learning, precision agriculture.

Citation: Shadrin D, Illarionova S, Kasatov R, Akimenkova M, Rudensky G, Erhan E. Weed detection on embedded systems using computer vision algorithms. *Computer Optics* 2025; 49(1): 103-111. DOI: 10.18287/2412-6179-CO-1454.

Introduction

Currently, Russia is the world leader in wheat export and is among ten countries for export of cereals. In 2021, the export of wheat amounted to 37.3 million tons while, in 2020, exports amounted to 40 million tons. However, crop yields are decreasing because of crop weeds, which impede their growth. Weeds in agricultural systems are a critical factor affecting the health of crops and the quality and quantity of crops [1]. In the absence of a mechanical solution, the vast majority of traditional agriculture either refrained from selective weed removal or adopted the use of herbicides in combination with herbicide-resistant crops [2].

The use of herbicides has raised a number of serious concerns [3], including:

- The use of herbicides is associated with environmental damage and adverse effects on human health;
- The use of herbicide-resistant crops, especially those created by genetic modification, has been resisted by environmentalists and the general public;
- The use of herbicides has led many weeds to develop resistance to herbicides, making this form of weed control less effective over time and posing a risk to natural ecosystems.

The environmental and health concerns associated with chemical weed control require viable alternatives or a significant reduction in its use. Weed control is already in use,

but it is labour-intensive and currently limited to organic farming systems. To automate the weed removal process, unwanted plants must first be successfully identified.

Volgograd Oblast plays a prominent role in the agricultural domain of country, with its extensive cultivation of a diverse array of crops. However, the propagation of weeds such as *Theophrastus* sp. and *Ambrosia artemisiifolia* is a growing problem in the agriculture of this region. *Abutilon* is widespread in the former Soviet republics, but it is most commonly found in the southwest and desert areas. It can interfere with the growth of crops such as cotton, flax, and other cultivated plants. *Ambrosia*, on the other hand, is a short-day plant that thrives in hot and relatively dry conditions [4]. Arable crops provide the most favorable conditions for this weed to grow. Unfortunately, *Ambrosia* has now contaminated almost all fields, especially those with field crops.

Traditionally, agronomists have manually examined fields and crops using various portable devices. Although these approaches may still be convenient for precision farming, there is a growing trend towards using autonomous systems for research [5]. The most commonly used tools for visual data gathering in agriculture are satellites, Unmanned Ground Vehicles (UGVs), and Unmanned Aerial Vehicles (UAVs). Precision site-specific weed control reduces pesticide consumption and environmental pollution, increases yields, simplifies and speeds up farmers' labor, resulting in economic profits.

To address these problems, it is necessary to automatically detect weeds and their locations for targeted spraying or removal, and to generate a map of weed coverage. Thanks to the advent of UAV technology, this detection can be achieved by collecting data from various heights using a drone.

UAVs have great potential in environmental studies, mainly because they are relatively easy to operate and can fly close to the ground, providing very high spatial resolution [6]. Although costs are relatively high at present, both acquisition and operating costs have been decreasing over time and are expected to continue to do so, making UAVs applicable for monitoring in agriculture. For large fields, dense vegetation, heterogeneous weed distribution, and regular monitoring of weeds or pests, the use of UAVs may be a better alternative to the use of UGVs. However, UGVs are much closer to the object under study and often have the ability to view the object from different angles. The advantage of this is that multiple measurements can be made, eliminating the effect of random errors in measurements. These conditions are difficult to achieve with drones. Weed detection studies with UGVs mainly focus on identifying weeds in real time. The aim of this strategy is the direct eradication of the weed once it has been identified. Research on weed detection by satellite and on-board vehicles often does not focus on real-time recognition. This is partly due to the fact that the vehicle should not perform the operation after weed identification. However, with the rapid development of UAV technologies, weed sensing becomes more effective and popular in practical applications [7]. This means that interactive systems in which land vehicles communicate with UAVs are now theoretically achievable. The use of drones for automated weed control opens up a new range of unexplored areas. Such problems range from which cameras to use with which resolution, which algorithms to run for detection, and how to include the collection of this information in the agricultural system. The study further explored the potential for UAV identification of specific plants.

This study addresses the problem of detecting Theophrastus and Ambrosia using a drone equipped with an embedded system capable of processing real-time images using convolutional neural networks. The experiments were conducted for agriculture crops with onion and carrot in Volgograd Oblast. The selected study region and plant types are representative of agriculture management in southeast territories of Russia (and regions with similar climate) and help shed light on their specificities. We examined three popular object detection algorithms (namely, Faster R-CNN, YOLOv5, and SSD) and compare them in terms of both performance quality and speed. We additionally evaluated inference on NVIDIA Jetson Nano to assess applicability of developed approach for real-time applications. The contribution of this work can be summarized as follows:

1. An effective pipeline for real-time detection of Theophrastus and Ambrosia is proposed. The approach utilizes an UAV integrated with a neural network, which operates on an embedded system, ensuring high-speed and high-quality performance.
2. A comparative study of the performance of three detection algorithms was conducted.
3. Two datasets for Theophrastus and Ambrosia detection were collected and annotated.

1. Related work

Weed detection using computer vision approaches is a challenging task in the agriculture domain. Complex environmental conditions, diverse lightning, overlapping vegetation, and highly unbalanced target classes require advanced algorithms. In terms of computer vision, the weed recognition task can be considered both as object detection, semantic segmentation task, or classification. Object detection problem statement assumes that the bounding box and corresponding class for each individual weed is determined, while in case of semantic segmentation, a weed's mask is defined. In case of classification, each image with crop or weeds is ascribed with one target class. However, weed localization is more important in practice. Labels for object detection are easier to obtain than segmentation masks, and such problem statement meets most of the practical demands. Both classical machine learning and deep learning approaches are considered in weed detection tasks [8]. Currently, deep learning algorithms are capable of extracting more complex features that leads to better performance [9]. Among the most commonly used architectures are R-CNN and Yolo-family architectures. We provide a detailed description of these detection algorithms in the following section. The mentioned architectures are widely applicable in various agriculture tasks. In [10], they consider the task of tobacco crop and weed recognition using Faster R-CNN and YOLOv5-based models. A unique field dataset was collected with various lightning conditions and crop growth stages. However, target classes objects have minimum overlapping that can differ from other crop types and weed plants and allow to achieve higher detection results. The best performance has been achieved for the R-CNN-based model with mAP of 0.94. More complex field conditions were considered in [11]. To collect a dataset, a portable cart was used that supports fixed height above the ground. The authors achieved a mAP of 0.55 using the Faster R-CNN model. A lack of available annotated datasets for weed detection also affects the current state of practical applications [8]. In [12], to address challenges caused by environment diversity and training data limitations, the authors extended the dataset artificially. New field images were created by pasting target objects to new locations, also, images were adjusted to reduce various weather conditions effects. They achieved a mAP of 0.829.

UAV imagery is a vital data source for field analysis. It allows one to conduct observations of large territories and effectively recognize target vegetation or its state. In [13], they used data from UAV to localize weed plants in real-time field application. In [14], different detection architectures were compared to process UAV field imagery from broccoli crops. However, inference time evaluation of the proposed approach was beyond the scope of the study. Real-time inference is another challenge in practical applications of weed detection. To reduce inference computing time, one can leverage techniques such as neural network compression or smaller architectures creation. In [15], they proposed a neural network approach for weed localization in real-time application using UAV imagery.

2. Methods and data

The main objective of the study is to develop an efficient pipeline for weed detection in agricultural crops. The limitation of available datasets poses additional challenges for a comprehensive study. The requirement for real-time performance necessitates the measurement of frames per second and conducting analysis. Therefore, to address all these challenges, we have outlined the following main stages of the study:

1. Field data collection and annotation. Manual annotation was performed for two classes, Theophrastus and Ambrosia. Additionally, data augmentation techniques were applied to enlarge the dataset.
2. Neural network training. Several Convolutional Neural Networks (CNN) were examined to solve the detection tasks, and their performance was compared.
3. Optimization of CNN for operation on the embedded, energy-efficient Jetson Nano system. The performance evaluation of the optimized CNN models included the following metrics:
 - Frames Per Second (FPS)
 - Mean Average Precision (mAP)
 - Model Size

The proposed scheme is presented in Fig. 1.

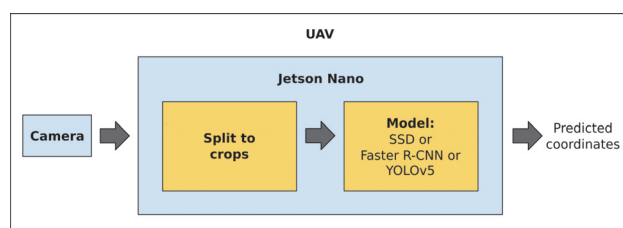


Fig. 1. Scheme of the final device for weed detection using UAV

2.1. Dataset collection and annotation

In this study, we collected and analyzed two datasets. For both of them, we captured the images using a UAV equipped with a 20 MP camera (5456×3632 dots) during the last week of May 2022 in Volgograd Oblast. The UAV was flown at an altitude of 5–8 meters, resulting in a resolution of 1.7–2.5 mm/pixel.

The first dataset contains 119 images of a carrot field. There are two target classes in the images: Ambrosia and Theophrastus (Fig. 2). The classes are imbalanced, with the number of instances of Ambrosia being much larger than that of Theophrastus. Statistics for the dataset are shown in Tab. 1.

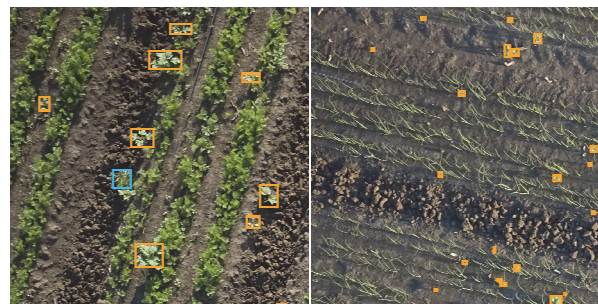


Fig. 2. Annotated image showcase: Ambrosia marked with orange bounding boxes and Theophrastus with blue. Displayed are 1364×1364 pixel fragments from the original photos for clarity. Left is from carrot dataset, right is from onion dataset

The second dataset contains 45 images of an onion field with only one target class of Ambrosia. The statistics of the dataset are presented in Tab. 2.

Each image was split into 6 non-intersecting parts due to its high original resolution. These parts were used as independent images. The dataset was split into training, validation, and test sets in an 80/10/10 ratio, respectively. Subsequently, the quality and speed of models performance were evaluated on these image segments.

Tab. 1. Number of objects in the dataset with carrot

class	train	test	val
Ambrosia	3589	445	463
Theophrastus	258	51	49

Tab. 2. Number of objects in the dataset with onion

class	train	test	val
Ambrosia	5028	583	578

2.2. Object detection models

Our task requires fast and accurate detection algorithms that can work in real-time. Modern object detection models are typically categorized into two types: single-stage and two-stage detectors. A two-stage detector first generates region proposals and then conducts classification only for the proposed regions. A single-stage detector skips the region proposal stage and performs detection directly over a dense sampling of possible locations of target objects. This approach is faster, but might lead to a particular degradation in prediction accuracy. As a result, two-stage detection models are generally more accurate than one-stage models, but they also have a higher computational cost and run slower on the same device.

In this paper, we consider the two-stage detector Faster R-CNN [16] and the single-stage YOLO [17], SSD [18].

The batch size was individually optimized for each model during training to enhance metrics.

2.2.1. Faster R-CNN

Faster R-CNN [16] is an object detection model that improves Faster R-CNN [19] by using a Region Proposal Network (RPN). RPN is a fully convolutional network that offers potential regions with various scales and aspect ratios. A classification network is applied to suggested regions next. Both networks share the same layers which decreases computational cost. This algorithm has been already applied in the object detection task for agriculture [20].

In this paper, in order to solve the problem of network degradation, ResNet is used as convolutional layers instead of the original VGG16.

Faster R-CNN model was trained in 100 epochs with batch size of 64.

2.2.2. YOLO

YOLO (You Only Look Once) is a family of models proposed in [17]. YOLO models are known for their high performance and small size, that make them ideal candidates for real-time launch and adaptation on embedded energy-efficient systems. YOLO-family algorithms have already proven to be highly effective in agricultural tasks with UAV observations [21].

The YOLO model contains convolutional layers for feature extraction and fully connected ones for predicting probabilities and coordinates of bounding boxes.

In this work, we use the small modification of YOLOv5 (YOLOv5s). That version was chosen because of its smaller size and faster image processing which is essential for the task.

YOLOv5s model was trained in 100 epochs with batch size of 16 and the optimizer SGD with momentum.

2.2.3. SSD detection algorithm

SSD is a popular single-stage object detection method that discretizes the output space of bounding boxes into a set of default boxes with different aspect ratios and scales per feature map location. During prediction, the network generates scores for the presence of each object category in each default box and refines the box coordinates to better match the object shape. The network also combines predictions from multiple feature maps with different resolutions to naturally handle objects of various sizes. SSD has been already implemented in agriculture tasks and shown high results [22].

For this task, the MobileNet [23] network was used instead of VGG16. MobileNet was designed specifically for mobile and embedded applications that require high speed. MobileNetV2 [24] features an inverted residual structure that helps eliminate non-linearity in narrow layers, resulting in advanced performance for the above applications. The SSD algorithm has been designed to integrate with various networks such as YOLO, MobileNet [24], and VGG [25]. Thus, MobileNet was integrated with SSD to improve performance and was called MobileNet-SSD.

SSD model was trained in 100 epochs with batch size of 4.

2.2.4. Processing of plants at the boundary

Some plants of target classes both in training and validation dataset were partially hidden due to being on the edge of the image. To assess its impact on the final metrics, such weeds were painted over with black rectangles and the corresponding bounding boxes were also excluded. Condition for painting over a plant is that the bounding box touches the edge of the image and its aspect ratio is greater than 2. Experimental results are presented both for original images and images with excluded plants on boundaries. The second case is marked with "_pb" (e.g., YOLOv5_pb).

2.3. Inference on embedded system

To address the challenge of weed detection, a complex platform is needed that is capable of autonomous performance in real-time. This requires a lightweight network that can operate at high speeds and a powerful, low-voltage embedded processor to support the network. In this work, the NVIDIA Jetson Nano was used since it allows the user to run a model written in any framework that supports CUDA. It does not require any neural network transformations and supports deep learning frameworks without limitations. Additionally, it includes the TensorRT [26] graph compiler, which is built to speed up inference.

2.4. Models performance evaluation

In this study, the quality of detection is evaluated using the Mean Average Precision metric with an Intersection over Union threshold of 0.5 (mAP0.5).

Object detection involves finding objects, classifying them, and localizing them by drawing a bounding box around them. Intersection over Union (IoU) is a metric that measures the quality of the match between the predicted bounding box and the true one:

$$IoU = \frac{\text{area_of_overlap}}{\text{area_of_union}} \in [0;1]. \quad (1)$$

In Figure 3 the black bounding box represents the true location of the object in the image, while the red bounding box represents the model's prediction. To calculate the Intersection over Union one needs to divide the green area by the blue area.

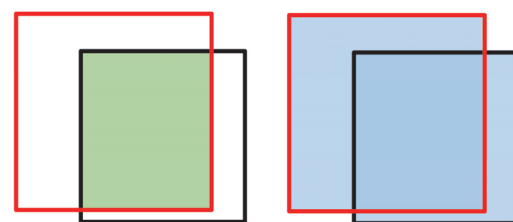


Fig. 3. Intersection over union

The higher the IoU , the better the detection quality is. IoU works well for any object size and shape. This metric is calculated for each object and forms the basis for the general object detection metric, Mean Average Precision (mAP). mAP is a popular metric in object detection task as it allows taking into account both errors in object localization and classification [27, 28]. Therefore, this metric considers the following types of errors: following types of errors: false positives, false negatives, localization errors, misclassification errors. False positives and false negatives correspond to the cases when the model incorrectly classifies an object that is not present or fails to detect an object that is present in the image, respectively. Localization error is connected with shifts in actual object position. Misclassification errors are connected with ascribing objects with wrong categories.

We consider the prediction of the object to be positive if $IoU > 0.5$. If several predictions of the same object are found, the first is considered positive and the rest are negative. We calculate the AP value at the IoU 0.5 threshold because it is important to us that the weed is detected, not its exact position (exact pixels), so in our case some bounding box deviation is allowed. Therefore, $mAP.50$ implies using IoU threshold 0.5. IoU is not a direct variable that is included in the calculation formula of mAP , but its boundary values influence this calculation.

mAP is the value of the average accuracy averaged over all categories (average precision, AP)

$$AP = \int_0^1 p(r)dr, \quad (2)$$

where p is the precision, r is the recall under the assumption that the bounding box is well defined if $IoU \geq 0.5$. Since precision and recall are between 0 and 1, AP , and hence mAP , is also between 0 and 1.

Besides model prediction quality, we evaluated its performance on Jetson Nano. The detection speed was measured in frames per second (FPS). FPS determines how fast the model processes the photo stream and returns the result.

3. Results and discussion

3.1. YOLOv5 experimentations and results

The results for weed detection using various algorithms in Carrot Dataset and Onion Dataset are presented in Tab. 3 and 4, respectively. YOLOv5 and Faster R-CNN outperform SSD in the Carrot Dataset with $mAP.50$ metrics of 0.668 and 0.682, respectively. SSD only achieves an $mAP.50$ of 0.526. It is worth noting that while Ambrosia is more widely represented in the Carrot Dataset, it is more challenging to correctly recognize this weed compared to Theophrastus. For example, YOLOv5 detects Ambrosia with an $mAP.50$ of 0.616, while Theophrastus is detected with the same metric at 0.719. The possible reason for this difference is the high similarity between Ambrosia weed and carrot plants at a particular growth stage. However, a $mAP.50$ score of 0.621 (such as for Faster R-CNN) is competitively high for

this type of computer vision task using UAV images. Examples with predictions of each model in the Carrot dataset are presented in Fig. 4, 6, 8.

Tab. 3. Experimental results on the test set for Carrot Dataset

Model	mAP.50	
	Ambrosia	Theophrastus
YOLOv5	0.616	0.719
YOLOv5 pb	0.624	0.720
SSD	0.464	0.588
SSD pb	0.480	0.602
Faster R-CNN	0.621	0.742
Faster R-CNN pb	0.636	0.758

Tab. 4. Experimental results on the test set for Onion Dataset

Model	mAP.50
YOLOv5	0.882
YOLOv5 pb	0.895
SSD	0.711
SSD pb	0.724
Faster R-CNN	0.842
Faster R-CNN pb	0.853

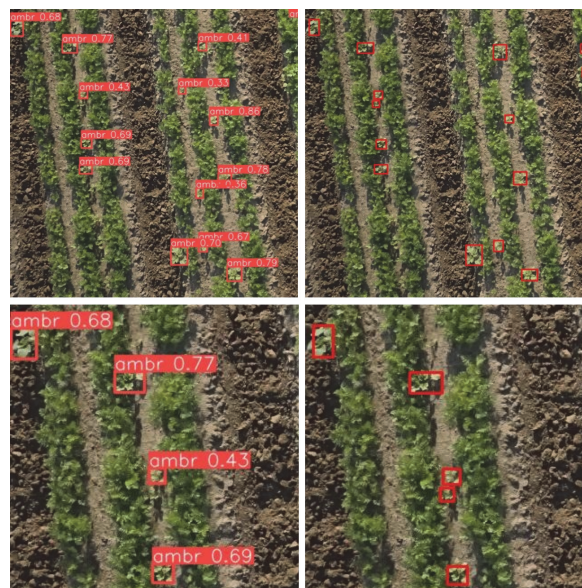


Fig. 4. Experimental results for YOLOv5 model on test image from Carrot Dataset. Predicted boxes are on the left, ground truth are on the right. The bottom patches depict increased areas from the top images

Although Faster R-CNN outperforms YOLOv5 model in the Carrot Dataset, we observe opposite results in the Onion Dataset. Here, YOLOv5 shows $mAP.50$ 0.882, while Faster R-CNN reaches only 0.842. The same as for the Carrot Dataset, SSD has the lowest results with $mAP.50$ equal to 0.711. Interestingly, that for this task the quality of Ambrosia detection is significantly higher than in the case of the Carrot dataset. From this, we can conclude that model performance highly depends on particular target classes and other objects that occur in dataset. In the considered example, onion plants and Ambrosia are more discernible than the same weed and carrot plants in another dataset. Examples with predictions for each model are presented in Fig. 5, 7, 9.

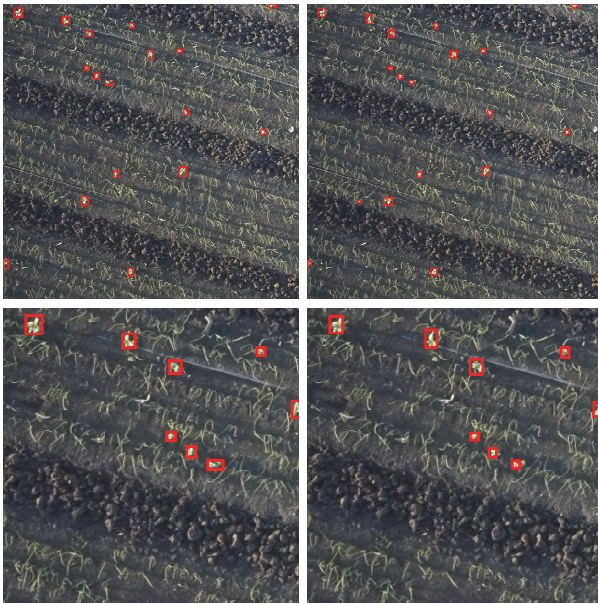


Fig. 5. Experimental results for YOLOv5 model on test images from Onion Dataset. Predicted boxes are on the left, ground truth are on the right. The bottom patches depict increased areas from the top images

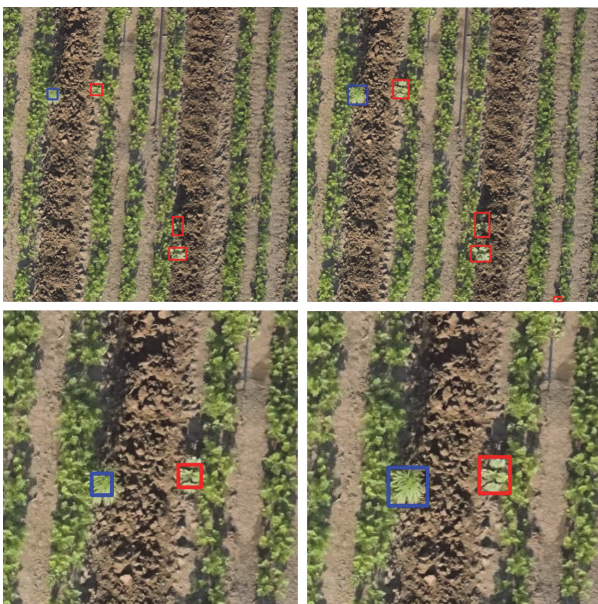


Fig. 6. Experimental results for SSD model on test image from Carrot Dataset. Predicted boxes are on the left, ground truth are on the right. The bottom patches depict increased areas from the top images

We also calculated metrics that excluded plants on the boundaries of the images. This was done to ensure a more accurate assessment of the developed models. For both datasets and all trained models, we observed that a common metric was slightly lower than the metric computed on samples with excluded weeds on the boundaries. For example, for the YOLOv5 model trained on the Carrot Dataset, we achieved an mAP.50 of 0.668 and 0.672 for the basic metric calculation and the proposed one, respectively. For the Onion Dataset, Faster R-CNN achieved an mAP.50 of 0.842, while another metric calculation resulted

in an mAP of 0.853. Although this difference is not significant, we consider this approach to be more "honest" when processing remote sensing data. It helps to avoid edge effects during model assessment on validation or test patches. We also suggest applying the same approach during model training, which should be further explored in future studies. In terms of real-life application and inference time, the flight plan for UAV should be designed with slight overlapping to ensure complete coverage of the study area without missing any weeds due to edge effects.

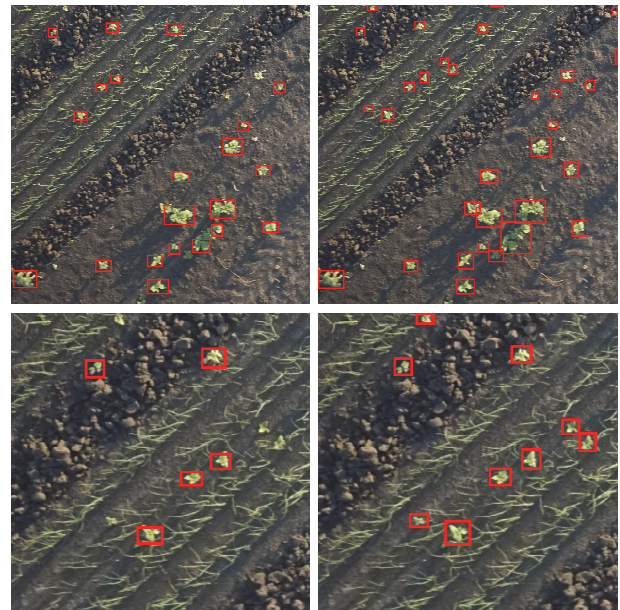


Fig. 7. Experimental results for SSD model on test images from Onion Dataset. Predicted boxes are on the left, ground truth are on the right. The bottom patches depict increased areas from the top images

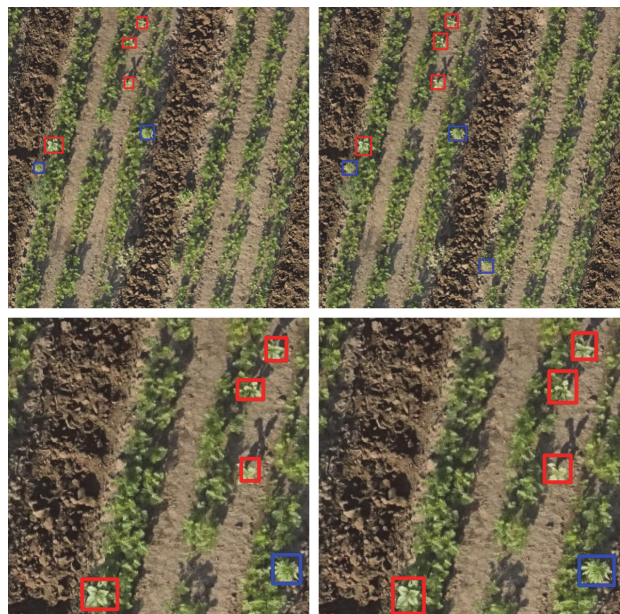


Fig. 8. Experimental results for Faster R-CNN model on test image from Carrot Dataset. Predicted boxes are on the left, ground truth are on the right. The bottom patches depict increased areas from the top images

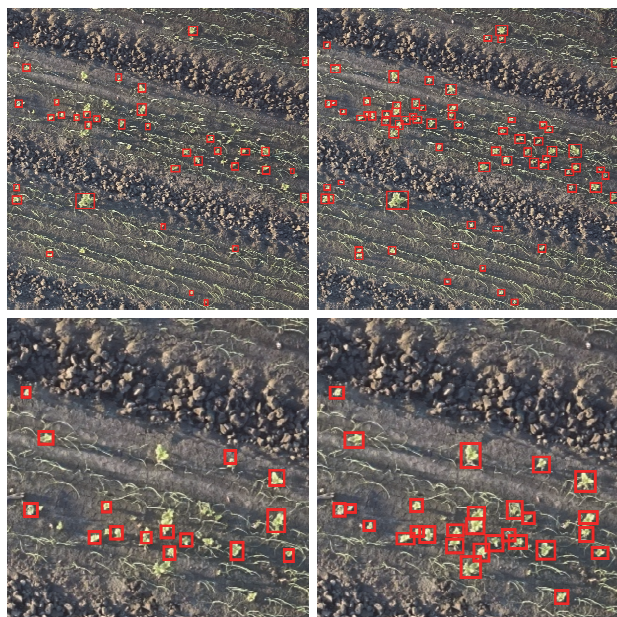


Fig. 9. Experimental results for Faster RCNN model on test images from Onion Dataset. Predicted boxes are on the left, ground truth are on the right. The bottom patches depict increased areas from the top images

As a promising direction for future study, we can consider advanced augmentation and synthetic samples generation to extend training datasets. Object-based augmentation approach has already shown high results in the remote sensing domain for vegetation mapping when objects masks and new backgrounds are used for creating new samples [29]. It may partially address the challenge of limited annotated field observations. Another perspective avenue is a few-shot learning techniques adaptation for new types of weeds to unsure fast and accurate object detection [30].

In this study, we considered a single vegetation period in May. For these crop types, this period corresponds to the beginning of growth and is the most relevant for this applied task. The reason is that it is easier to control weeds in the earliest growing stage as much as possible. Therefore, we currently focus exactly on this period. However, other observation periods might also be valuable for additional analysis and further practical implementations. Plant size and shape might differ in later stages of growth. To tackle the challenge of diverse plant variations, one can synthesize artificial training samples to represent different growing stages. For instance, by varying the relative size of agricultural plants and weeds on training images. To extend original field datasets, pseudo-labeling techniques can also be implemented to facilitate the labeling process [31]. A single or individual neural network can be trained for each stage. However, a single model might be more robust for new observation periods. All these approaches will allow farmers to take repeated measurements to control and destruct weeds during various growing stages.

If datasets with new plant types become available, transfer learning approaches are preferable to fine-tune

the model on new samples. It is less time-consuming and allows achieving higher performance even with a limited amount of training samples. It has been shown that such pre-training on data from the same agriculture domain is preferable than pre-training on a large scope of annotated data from another computer vision domain [21].

3.2. Inference on Jetson Nano and optimization using TensorRT

For this study, we utilized a C920s PRO HD WEBCAM camera connected to the Jetson Nano via USB. We measured the inference speed of each of the three trained models on the camera. The initial results are displayed in Table 5 under the column labeled FPS. We observed the highest FPS for the YOLOv5 algorithm, which was 13. The SSD and Faster R-CNN models achieved lower FPS values of 9 and 2, respectively. These FPS values may not be adequate for real-life applications, so we proceeded with additional adjustments to improve the speed.

To enable real-time object detection, we needed to reduce the inference time. When we ran images from our dataset on three models (SSD, YOLOv5, and Faster R-CNN) on an NVIDIA Jetson Nano, we obtained unsatisfactory results. The initial results indicated that image processing took a significant amount of time. To optimize the inference process, we utilized specialized software called TensorRT, which optimized the network and resulted in an increased FPS value. The results of inference before and after optimization are presented in Tab. 5. The comparison between three models for two datasets are shown in Fig. 10. After optimization, the best algorithm in terms of speed, YOLOv5, achieved an increase in FPS from 13 to 31. These results are considered acceptable for real-time performance on embedded systems.

Tab. 5. Optimized Inference speed on Jetson Nano, input image size 416×416

Model	FPS	FPS, opt
YOLOv5, onion dataset	13	31
SSD, onion dataset	9	17
Faster R-CNN, onion dataset	2	11
YOLOv5, carrot dataset	11	29
SSD, carrot dataset	8	17
Faster R-CNN, carrot dataset	2	10

The developed approach can be further integrated into systems of weeds management with laser treatments such as was proposed in [32]. The previous studies investigated diode lasers in their effectiveness in killing or stunting weeds in agriculture crops. Such systems require high accuracy of weed localization that can be provided by the considered in our study neural networks.

In conclusion, our findings highlight the importance of model optimization for real-time performance in computer vision applications. It is crucial to pay attention to both model quality and performance. The conducted study demonstrates the effectiveness of the YOLOv5 algorithm in terms of both detection quality and speed.

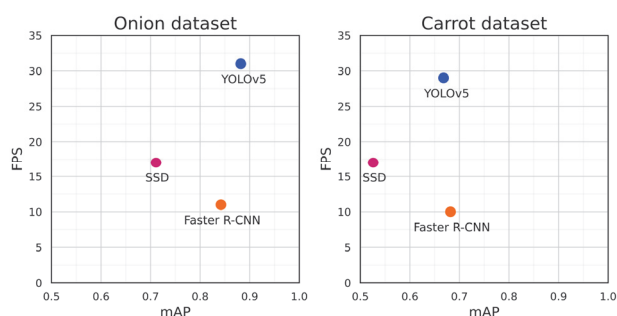


Fig. 10. Optimized FPS and mAP for Onion and Carrot datasets

Conclusion

Weed detection plays a crucial role in promoting sustainable agriculture and ensuring optimal plant growth. The implementation of real-time computer vision algorithms offers a promising solution for addressing the challenges associated with infestation control and precision farming in the near future. In this study, we specifically focus on the analysis of advanced methods for object detection and propose our own effective pipeline for weed recognition.

We evaluate the performance of three popular algorithms, namely YOLOv5, Faster R-CNN, and SSD, in distinguishing weeds in agriculture crops such as onions and carrots. To conduct our assessment, we collect and utilize two datasets. Our results demonstrate the high accuracy of Theophrastus and Ambrosia weeds detection on UAV images. Additionally, we optimize the network on NVIDIA Jetson Nano to enhance its efficiency.

Among the examined algorithms, YOLOv5 outperforms the others in terms of frame rate for both datasets. It is three times faster than Faster R-CNN. However, when considering a dataset with two classes, Faster R-CNN achieves the best mean Average Precision (mAP). On the other hand, YOLOv5 outperforms Faster R-CNN in terms of quality for a dataset with only one class. SSD performs slightly slower than YOLOv5, with a difference of 3–4 frames per second, but it is inferior to both networks in terms of quality.

During the work, the following conclusions were made:

- The best detection result for the carrot dataset is provided by the Faster R-CNN model, but it consumes too much memory on the NVIDIA Jetson Nano and runs at a speed insufficient for real-time execution.
- The MobileNet-V2 SSD model works at a fairly high speed, but is inferior in quality to the other two models.
- The YOLOv5 model has the highest speed and gives sufficient accuracy for both datasets.
- Network optimization with TensorRT software is an important step when deploying on energy efficient systems.

Overall, our study emphasizes the potential of computer vision algorithms for weed detection in agriculture and highlights the comparative advantages of YOLOv5 among the considered algorithms in terms of speed and quality. It is

worth noting that the same approach can be applied to various tasks such as plant disease detection, forest monitoring, and plant phenotyping on board UAVs.

Acknowledgements

The work of Svetlana Illarionova was supported by the Russian Science Foundation (Project No. 23-71-01122).

References

- [1] Korres NE, Burgos NR, Travlos I, et al. New directions for integrated weed management: Modern technologies, tools and knowledge discovery. *Adv Agron* 2019; 155: 243-319. DOI: 10.1016/bs.agron.2019.01.006.
- [2] Gage KL, Krausz RF, Walters SA. Emerging challenges for weed management in herbicide-resistant crops. *Agriculture* 2019; 9(8): 180. DOI: 10.3390/agriculture9080180.
- [3] Wang A, Zhang W, Wei X. A review on weed detection using ground-based machine vision and image processing techniques. *Comput Electron Agr* 2019; 158: 226-240. DOI: 10.1016/j.compag.2019.02.005.
- [4] Lavoie C, Jodoin Y, De Merlis AG. How did common ragweed (*Ambrosia artemisiifolia* L.) spread in Québec? A historical analysis using herbarium records. *J Biogeogr* 2007; 34(10): 1751-1761. DOI: 10.1111/j.1365-2699.2007.01730.x.
- [5] Mahmud MSA, Abidin MSZ, Emmanuel AA, Hasan HS. Robotics and automation in agriculture: present and future applications. *Applications of Modelling and Simulation* 2020; 4: 130-140.
- [6] Illarionova S, Shadrin D, Tregubova P, Ignatiev V, Efimov A, Oseledets I, Burnaev E. a survey of computer vision techniques for forest characterization and carbon monitoring tasks. *Remote Sens* 2022; 14(22): 5861. DOI: 10.3390/rs14225861.
- [7] Menshchikov A, Shadrin D, Prutyantov V, Lopatkin D, Sosnin S, Tsykunov E, Iakovlev E, Somov A. Real-time detection of hogweed: UAV platform empowered by deep learning. *IEEE Trans Comput* 2021; 70(8): 1175-1188. DOI: 10.1109/TC.2021.3059819.
- [8] Wu Z, Chen Y, Zhao B, Kang X, Ding Y. Review of weed detection methods based on computer vision. *Sensors* 2021; 21(11): 3647. DOI: 10.3390/s21113647.
- [9] Hasan ASMM, Soheli F, Diepeveen D, Laga H, Jones MGK. A survey of deep learning techniques for weed detection from images. *Comput Electron Agr* 2021; 184: 106067. DOI: 10.1016/j.compag.2021.106067.
- [10] Alam MS, Alam M, Tufail M, Khan MU, Güneş A, Salah B, Nasir FE, Saleem W, Khan MT. TobSet: A new tobacco crop and weeds image dataset and its utilization for vision-based spraying by agricultural robots. *Appl Sci* 2022; 12(3): 1308. DOI: 10.3390/app12031308.
- [11] Le VNT, Truong G, Alameh K. Detecting weeds from crops under complex field environments based on Faster RCNN. *2020 IEEE Eighth Int Conf on Communications and Electronics (ICCE)* 2021: 350-355. DOI: 10.1109/ICCE48956.2021.9352073.
- [12] Gao J, French AP, Pound MP, He Y, Pridmore TP, Pieters JG. Deep convolutional neural networks for image-based *Convolvulus sepium* detection in sugar beet fields. *Plant Methods* 2020; 16: 29. DOI: 10.1186/s13007-020-00570-z.
- [13] Revanasiddappa B, Arvind C, Swamy S, et al. Real-time early detection of weed plants in pulse crop field using drone with IoT. *Technology* 2020; 16: 1227-1242. DOI: 10.13140/RG.2.2.27656.03845.
- [14] Psiroukis V, Espejo-Garcia B, Chitos A, Dedousis A, Karantzalos K, Fountas S. Assessment of different object

- detectors for the maturity level classification of broccoli crops using UAV imagery. *Remote Sens* 2022; 14(3): 731. DOI: 10.3390/rs14030731.
- [15] Deng J, Zhong Z, Huang H, Lan Y, Han Y, Zhang Y. Lightweight semantic segmentation network for real-time weed mapping using unmanned aerial vehicles. *Appl Sci* 2020; 10(20): 7132. DOI: 10.3390/app10207132.
- [16] Ren S, He K, Girshick R, Sun J. Faster R-CNN: Towards real-time object detection with region proposal networks. *NIPS'15: Proc 28th Int Conf on Neural Information Processing Systems* 2015; 1: 91-99.
- [17] Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: Unified, real-time object detection. 2016 IEEE Conf on Computer Vision and Pattern Recognition (CVPR) 2016: 779-788. DOI: 10.1109/CVPR.2016.91.
- [18] Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C-Y, Berg AC. SSD: Single shot multibox detector. In Book: Leibe B, Matas J, Sebe N, Welling M, eds. *Computer Vision – ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I*. Cham, Switzerland: Springer International Publishing AG; 2016: 21-37. DOI: 10.1007/978-3-319-46448-0_2.
- [19] Girshick R. Fast R-CNN. 2015 IEEE Int Conf on Computer Vision (ICCV) 2015: 1440-1448. DOI: 10.1109/ICCV.2015.169.
- [20] Wan S, Goudos S. Faster R-CNN for multi-class fruit detection using a robotic vision system. *Computer Networks* 2020; 168: 107036. DOI: 10.1016/j.comnet.2019.107036.
- [21] Dolgaia L, Illarionova S, Nesteruk S, Krivolapov I, Baldycheva A, Somov A, Shadrin D. Apple tree health recognition through the application of transfer learning for UAV imagery. 2023 IEEE 28th Int Conf on Emerging Technologies and Factory Automation (ETFA) 2023: 1-8. DOI: 10.1109/ETFA54631.2023.10275369.
- [22] Wang Y, Xing Z, Ma L, Qu A, Xue J. Object detection algorithm for lingwu long jujubes based on the improved SSD. *Agriculture* 2022; 12(9): 1456. DOI: 10.3390/agriculture12091456.
- [23] Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M, Adam H. MobileNets: Efficient convolutional neural networks for mobile vision applications. *arXiv Preprint*. 2017. Source: <<https://arxiv.org/abs/1704.04861>>. DOI: 10.48550/arXiv.1704.04861.
- [24] Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L-C. MobileNetV2: Inverted residuals and linear bottlenecks. *Proc IEEE Conf on Computer Vision and Pattern Recognition* 2018: 4510-4520. DOI: 10.1109/CVPR.2018.00474.
- [25] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv Preprint*. 2014. Source: <<https://arxiv.org/abs/1409.1556>>. DOI: 10.48550/arXiv.1409.1556.
- [26] Jeong EJ, Kim J, Ha S. TensorRT-based framework and optimization methodology for deep learning inference on Jetson boards. *ACM Transactions on Embedded Computing Systems (TECS)* 2022; 21(5): 51. DOI: 10.1145/3508391.
- [27] Padilla R, Passos WL, Dias TLB, Netto SL, Da Silva EAB. A comparative analysis of object detection metrics with a companion open-source toolkit. *Electronics* 2021; 10(3): 279. DOI: 10.3390/electronics10030279.
- [28] Zhu H, Wei H, Li B, Yuan X, Kehtarnavaz N. A review of video object detection: Datasets, metrics and methods. *Appl Sci* 2020; 10(21): 7834. DOI: 10.3390/app10217834.
- [29] Illarionova S, Shadrin D, Ignatiev V, Shayakhmetov S, Trekin A, Oseledets I. Augmentation-based methodology for enhancement of trees map detalization on a large scale. *Remote Sens* 2022; 14(9): 2281. DOI: 10.3390/rs14092281.
- [30] Huang G, Laradji I, Vazquez D, Lacoste-Julien S, Rodriguez P. A survey of self-supervised and few-shot object detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2022; 45(4): 4071-4089. DOI: 10.1109/TPAMI.2022.3199617.
- [31] Nesteruk S, Illarionova S, Zhrebzov I, Traweck C, Mikhailova N, Somov A, Oseledets I. PseudoAugment: Enabling smart checkout adoption for new classes without human annotation. *IEEE Access* 2023; 11: 76869-76882. DOI: 10.1109/ACCESS.2023.3296854.
- [32] Mwitwa C, Rains GC, Prostko E. Evaluation of diode laser treatments to manage weeds in row crops. *Agronomy* 2022; 12(11): 2681. DOI: 10.3390/agronomy12112681.

Authors' information

Dmitrii Shadrin, PhD in Data Science. He works as a Head of Direction at Skolkovo Institute of Science and Technology (Skoltech), Russia. E-mail: d.shadrin@skoltech.ru

Svetlana Illarionova, PhD in Data Science. She works as a Head of research group at Skolkovo Institute of Science and Technology (Skoltech), Russia. E-mail: s.illarionova@skoltech.ru

Roman Kasatov, student at ITMO University and works as a research intern at Skolkovo Institute of Science and Technology (Skoltech), Russia. His areas of scientific interest include machine learning, computer vision and programming. E-mail: r.kasatov@skoltech.ru

Maria Akimenkova, graduated from the Moscow Institute of Physics and Technology in 2022, works as a research intern at Skolkovo Institute of Science and Technology (Skoltech), Russia. Research interests: computer vision, programming, time series forecasting. E-mail: M.Akimenkova@skoltech.ru

Gennady Rudensky works as a Chief Design Engineer at Skolkovo Institute of Science and Technology (Skoltech) in Laboratory of cyber-physical systems, Russia. E-mail: g.rudenskiy@skoltech.ru

Eugene Erhan works as a Senior Engineer at Skolkovo Institute of Science and Technology (Skoltech) in Center of Excellence of the National Technology Initiative "Wireless Technologies and Internet of Things", Russia. E-mail: E.Erhan@skoltech.ru

Received November 06, 2023. The final version – March 22, 2024.