Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 15th International Scientific and Practical Conference. Volume II, 29-33

Cherry Fruitlet Detection using YOLOv5 or YOLOv8?

Ilmārs Apeināns
Institute of Engineering
Rezekne Academy of Technologies
Rezekne, Latvia
ilmars.apeinans@rta.lv

Marks Sondors
Institute of Engineering
Rezekne Academy of Technologies
Rezekne, Latvia
ms20094@edu.rta.lv

Lienīte Litavniece
Research Institute for Business and Social Processes
Rezekne Academy of Technologies
Rezekne, Latvia
lienite.litavniece@rta.lv

Sergejs Kodors
Institute of Engineering
Rezekne Academy of Technologies
Rezekne, Latvia
sergejs.kodors@rta.lv

Imants Zarembo
Institute of Engineering
Rezekne Academy of Technologies
Rezekne, Latvia
imants.zarembo@rta.lv

Daina Feldmane
Institute of Horticulture (LatHort)
Dobele, Latvia
daina.feldmane@llu.lv

Abstract. **Agriculture 5.0 incorporates autonomous decision-making systems in order to make agriculture more productive. Our study is related to the development of the autonomous orchard monitoring system using unnamed aerial vehicles for automatic fruiting assessment and yield forecasting. Respectively, artificial intelligence must be developed to count fruits in an orchard. The modern solutions are mainly data-based. Therefore, we collected and annotated cherry dataset with natural images (CherryBBCH81) for neural network training. The goal of the experiment was to select the optimal "You Look Only Once" (YOLO) model for the rapid development of fruit detection. Our experiment showed that YOLOv5m provided better results for CherryBBCH81 – mean average precision (mAP) at 0.5 0.886 in comparison with YOLOv8m mAP@0.5 0.870. However, additional tests with dataset Pear640 showed that YOLOv8m can outperform YOLOv5m: 0.951 vs 0.943 (mAP@0.5).**

Keywords: Agriculture 5.0, artificial intelligence, deep learning, yield estimation.

I. Introduction

Sweet cherries (*Prunus avium* L.) are among the top 5 most sought after fruits in the world. According to industry information, the demand for fresh cherries will grow by 7.5% in the period from 2022 to 2027, reaching 84.3 billion dollars [1].

Analogous to all industries, agriculture has also evolved over the centuries from Agriculture 1.0, where economic activity was based on the physical strength of people and animals, to Agriculture 5.0, where the essence of economic activity is characterized by smart and more energy-efficient management [2]. The European Commission set the year 2021 as the official start of the "Industry 5.0" era [3]. Agriculture 5.0 can also be called "digital agriculture", which aims to maximize yields and other agricultural results by applying the latest methods and tools. Agriculture 5.0 is characterized by: the efficiency of data collection, accuracy of data, and timeliness of data acquisition in order to make correct and data-based decisions. Data-driven decision-making is essential because as the planet's population grows, it is necessary to produce more food while respecting the principles of sustainability.

The aim of the project lzp-2021/1-0134 is to develop an autonomous decision making smart fruit growing solution for apple (*Malus* × *domestica* (L.) Borkh), pear (*Pyrus communis* L.) and sweet cherry (*Prunus avium* L.) orchard management that could provide an accessible and low cost smart horticulture solution for commercial orchard owners. The system is aimed to enable an automatic and autonomous orchard monitoring capabilities using unnamed aerial vehicles (UAV) and allow automatic flowering and fruiting assessment. The digital twin paradigm is applied to orchard management in combination with an UAV and artificial intellect (AI) solution [4]. UAV autonomously collects orchard data. Transmits it back to the base station, which, in turn, sends the data to the server where the AI solution performs yield estimation. Finally, orchard managers can interact with the yield forecast via a web interface on their computer and make decisions accordingly.

Print ISSN 1691-5402
Online ISSN 2256-070X
https://doi.org/10.17770/etr2024vol2.8013

When it comes to implementation of a yield forecasting solution based on artificial intelligence in the autonomous orchard system, the yield forecasting is based on the yield estimation results. The orchard monitoring is completed by UAV, while the images are processed by object detection algorithms like "You Look Only Once" (YOLO) to estimate the visible amount of fruits. Meanwhile, the fruit load on trees is predicted by a post-processing algorithm using yield data from multiple images. For example, citrus yield prediction solution was presented, utilizing YOLO and four post-processing algorithms: the gradient boosting regression, random forest regression, linear regression and partial least squares regression; showing the following results, respectively: 41.12%, 41.47%, 35.59% and 35.84% mean average precision (mAP) [5]. Another example, the wild blueberry (*Vaccinium angustifolium* Ait.) yield prediction was implemented using YOLO and nonlinear regression model, which achieved a mean absolute error of 24.1% [6].

YOLOv8 was released in 2023. Despite the existence of studies on YOLOv8, e.g. [7] and [8], the number of experiments in smart agriculture remain limited at the moment. The publications mostly cover the problems of fruit and vegetable quality assessment and disease detection. For example, YOLOv8 was used to inspect the quality of tomatoes (*Lycopersicon esculentum* Mill.) [8]. The trained convolutional neural network (CNN) achieved mAP of up to 99.5%, with the precision of 96.3% and the recall of 96.1% [8]. Meanwhile, the authors of "Tomato Maturity Detection and Counting Model Based on MHSA-YOLOv8" [7] improved YOLOv8 by adaptation of the multi-head self-attention mechanism (MHSA), which is used to enhance the network's ability to extract diverse features. The MHSA improved YOLOv8 results on recall, F1-score, and mAP@0.5 by 0.044, 0.003, and 0.004 compared to YOLOv8. The MHSA-YOLOv8 was compared with other YOLO family algorithms: YOLOv3, YOLOv4, YOLOv5, YOLOv7 and YOLOv8. The comparison results showed that the best algorithm for classical object detection is YOLOv8 with precision of 84.7% compared to runner up, YOLOv5 with 84.0% and mAP@0.5 of 0.859 for YOLOv8 compared to 0.778 for YOLOv5.The results of the comparison revealed that YOLOv8 outperforms in classical object detection, achieving a precision of 84.7%, slightly higher than its closest competitor, YOLOv5, which scored 84.0%. Additionally, YOLOv8 boasts a mAP@0.5 of 0.859, surpassing YOLOv5's 0.778.

Another modification of YOLOv8 was YOLOv8-Seg developed for tomato disease detection [9]. YOLOv8-Seg was used to detect tomato fruits, and classify them according to health status, and if they are infected with disease, then classify disease that is discovered. YOLOv8-Seg goes a step further than object detection and involves identifying individual objects in an image and segmenting them from the background. Then each object is classified into classes based on the health status of fruit. The publication describes how the algorithm was further improved to get better results. The improved YOLOv8s-Seg algorithm achieves precision, recall, F1-score, and segment mAP@0.5 of 91.9%, 85.8%, 88.7%, and 0.922, respectively. Compared to the YOLOv8s-Seg algorithm, the improvements were 1.6%, 0.4%, 1.0%, and 2.4%, respectively.

Our project team has already experimented with the different models of YOLOv5 and YOLOv7 to select the optimal solution for the rapid prototyping of fruit detection CNN. Our previous experiment showed that YOLOv5m is the most suitable model for fruit detection [10]. Now, we want to update our experiment results comparing the YOLOv5m with the new CNN models of YOLOv8 architecture, as well as, to present the new dataset called "CherriesBBCH81" under CC-BY 4.0 license.

The aim of study is to experimentally compare YOLOv5m with YOLOv8n, YOLOv8s and YOLOv8m to select the most suitable of them for the fruit detection tasks.

The novelty of publication:

- The new natural image dataset called "CherryBBCH81" is presented, which contains annotated images of cherry fruitlets BBCH81 prepared for YOLO model training.
- The best results were obtained by using YOLOv5m for CherryBBCH81 and by using YOLOv8m for Pear640.

II. MATERIALS AND METHODS

CherryBBCH81 collection and annotation:

The photo fixation of cherry fruitlets was done in the LatHort orchard in Dobele, at the beginning of fruit coloration (BBCH stage 81) [11]. Two photo images were taken for each tree – perpendicularly, in a tree-facing view and in an oblique view. To determine the number of fruits in the tree, the fruits were counted on sample branches and multiplied by the number of such type branches in the tree. The photo images were annotated (See Fig. 1), adding the information of tree identity (tree number) and basic parameters of the tree and orchard (cultivar, rootstock, canopy type, planting distances, tree dimensions). The images were taken from the cultivars 'Arthur', 'Artis', 'Bryansk 3-36', 'Doneckiy 42-37', 'Paula', 'Radica' and 'Techlovan' grown under the cover, and from the cultivars 'Aija', 'Aleksandrs', 'Elfrida', 'Bryansk 3-36', 'Bryanskaya Rozovaya', 'Kompaktnaya Venyaminova' and 'Paula' grown in open orchard.

Then the annotated images 6016x4000px were automatically cropped out on 640x640px images with overlap 30%. 640x640px images provide sufficient detail for object detection while still being viable in terms of computing resources, but image resizing was not possible due to small bounding boxes, which could achieve size until 25x25px. Once the annotation process for all the images was completed, they were saved in a YOLO format. The dataset is available in Kaggle repository under CC-BY 4.0 licence [12].

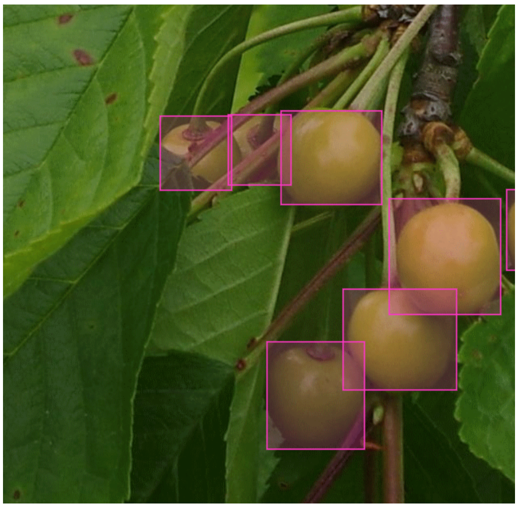

Fig. 1. CherryBBCH81 image example.

Comparison dataset Pear640:

Pear640 dataset [13] consists of 712 images (See Fig. 2.) containing 8340 pear objects. Digital images of pear fruits in this dataset were collected at the LatHort Institute's experimental site using 'Suvenirs' and 'Mramornaya' cultivars planted on 'Kazraushu' seedling rootstocks. The images were taken in field conditions, capturing the whole canopy as separate objects, around noon under clear sky conditions. This dataset provides a comprehensive collection of images taken under similar conditions as the CherryBBCH81 dataset so it makes for a reliable comparison dataset and it also was annotated using YOLO format, furthermore it is also in resolution of 640x640px.

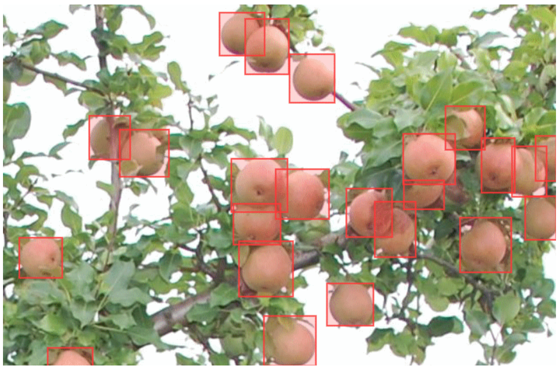

Fig. 2. Pear640 image example.

Experiment design:

In this experiment, we applied YOLOv5m [14] and YOLOv8 [15] models, specifically YOLOv8n, YOLOv8s, YOLOv8m. The experiment was conducted on an NVIDIA RTX 2070 GPU, which provided sufficient performance for training and testing the models.

The CherryBBCH81 dataset was randomly divided into training and validation folders using the random shuffle method in Python. This process was carried out five times to create five unique data splits. Each data split contained the same images, but they were located in different folders. In each split, 80% of the images were assigned to the training folder and 20% to the validation folder. The same procedure was also applied to the comparison dataset Pear640.

In the experiment, the default augmentation was modified. Mix-up was adjusted to zero, as was the mosaic and shearing. However, these augmentation modifications were only applied to the training of YOLOv5m. For YOLOv8, the optimizer was set to "auto" to highlight its potential since it demonstrated enhancements.

For every data split, both YOLOv5m and YOLOv8 models were separately trained, leading to the creation of five distinct trained models for each model type. Subsequently, the results were collected and examined.

The experiment was designed similarly "Rapid Prototyping of Pear Detection Neural Networks with YOLO Architecture in Photographs" [10] to get comparable results.

III. RESULTS AND DISCUSSION

If we analyze the results of each dataset separately, starting with the newly created dataset CherryBBCH81 that consists of images of cherry fruitlets. The best results were achieved with the YOLOv5m model with mAP@0.5 of 87.7% (Tab. 1, median). In comparison, YOLOv8 results were worse: 85.5%, 86.9%, 86.2% for YOLOv8n, YOLOv8s and YOLOv8m respectively. Important to note, the results improved as the size of YOLOv8 increased.

Importance of consistency in machine learning results is crucial as it indicates the model's reliability and robustness across different datasets or under varying conditions, ensuring that the insights derived are dependable. Furthermore, consistent performance facilitates the fine-tuning and generalization of models, making it easier to identify areas of improvement and build trust in the model's outputs for decision-making.

If we analyze results on the basis of consistency, difference between min and max results of CherryBBCH81 dataset then it can be seen mAP@0.5 88.6%, mAP@0.5:0.95 40.1%, precision of 0.85 and recall of 0.82 that even though YOLOv5m produced the best results of training, consistency is worse with result variance of 1.9% between the best and the worst results achieved. In comparison results of YOLOv8m were the best with deviation of 1% between the best and the worst results. Deviation for YOLOv8s is 1.2% and for YOLOv8n is 1.7%. Based on the results achieved, in terms of consistency of training, YOLOv8m is better than the rest of the models in this experiment.

By looking at the previous experiment with Pear640 dataset [10], results of YOLOv5 showed 4.1% better mAP@0.5 compared to YOLOv7 model versions. During experiments with YOLOv5 and YOLOv8, the difference in results is not as impressive as it was in the previous experiments. Improvement that was achieved by YOLOv8 is only 0.9% (Tab. 2, median). YOLOv5m model mAP@0.5 is 93.8%, while YOLOv8m provides the best results with mAP@0.5 of 94.7%. However, the previous experiment showed the maximal YOLOv5m mAP@0.5 equal to 0.951 [10], which is equal to the best result of YOLOv8m in this experiment.

In scope of model versions of YOLOv8, the best results were with YOLOv8m with mAP@0.5 of 94.7%, however difference to other versions is insignificant, as YOLOv8s results were 94.4% and YOLOv8n 94.6%. It shows that all results of YOLOv8 at mAP@0.5 outperformed YOLOv5 in the case of.

Results of YOLOv8 were, in general, better than YOLOv5, but another aspect that can be seen in results is consistency of results (See Fig. 3). If we look at the results of YOLOv5m, the difference between min and max values achieved is 1.6%. Close second is YOLOv8n with a difference of 1.5%. The best results were achieved with YOLOv8s and YOLOv8m with results of 0.9%. Considering potential usage of trained models and amount of work that will be assigned to it, consistency of results is an important factor in decision making.

YOLOv8 models resulted in better recognition percentage with the Pear640, in comparison with YOLOv5, but at same time YOLOv5 resulted in better object recognition then YOLOv8 using CherryBBCH81 dataset. If results are examined in scope of consistency, then best results were achieved by YOLOv8m.

TABLE 1 EXPERIMENT RESULTS WITH CHERRYBBCH81

YOLO	Test Dataset CherryBBCH81 (mAP@0.5)			
	v5m	*v8n*	*v8s*	*v8m*
min	0.867	0.847	0.853	0.860
mean	0.878	0.857	0.867	0.864
median	0.877	0.855	0.869	0.862
max	0.886	0.864	0.875	0.870

TABLE 2 EXPERIMENT RESULTS WITH PEAR640

YOLO	Test Dataset Pear640 (mAP@0.5)			
	v5m	*v8n*	*v8s*	*v8m*
min	0.927	0.932	0.938	0.942
mean	0.935	0.941	0.939	0.947
median	0.938	0.946	0.944	0.947
max	0.943	0.947	0.947	0.951

IV. CONCLUSIONS

In this article we presented our public dataset CherryBBCH81 (which is available in Kaggle under CC-BY 4.0 licence), and the YOLO model comparison experiment results.

The objective of the experiments conducted during the writing of this article was to develop a yield estimation solution. To further the realization of the goal we wanted to identify the optimal YOLO model for the rapid development of fruit detection neural networks. The experiments were done using our own datasets Pear640 and CherryBBCH81. Our experiment showed that YOLOv8m provided best consistency of training, yet looking at training results, the best results were obtained by using YOLOv5m with CherryBBCH81 dataset: mAP@0.5 88%, mAP@0.5:0.95 42% and YOLOv8m for Pear640: mAP@0.5 95%, mAP@0.5:0.95 56%.

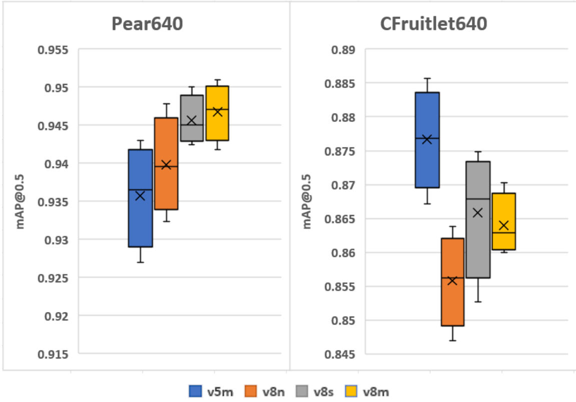

Fig. 3. Box-plot diagram of YOLO model accuracy (mAP@0.5).

ACKNOWLEDGMENTS

This research is funded by the Latvian Council of Science, project "Development of autonomous unmanned aerial vehicles based decision-making system for smart fruit growing", project No. lzp-2021/1-0134.

REFERENCES

[1] Fresh Cherries Market – Forecast (2023 - 2028). 2021. [Online]. Available: https://www.industryarc.com/Research/Fresh-Cherries-Market-Research-511079 [Accessed: Dec. 12, 2023]

[2] K. Ragazou, A. Garefalakis, E. Zafeiriou and I. Passas, "Agriculture 5.0: A new strategic management mode for a cut cost and an energy efficient agriculture sector," *Energies*, vol. 15, no. 9, p. 3113, Apr. 2022, https://doi.org/10.3390/en15093113.

[3] A. Ahmad, R. Damaševičius, Agriculture 5.0 and Remote Sensing. *Encyclopedia*. 2023. [Online]. Available: https://encyclopedia.pub/entry/11655 [Accessed: Nov. 6, 2023]

[4] I. Zarembo, S. Kodors, I. Apeināns, G. Lācis, D. Feldmane and E. Rubauskis, "DIGITAL TWIN: ORCHARD MANAGEMENT USING UAV", *ETR*, vol. 1, pp. 247–251, Jun. 2023, https://doi.org/10.17770/etr2023vol1.7290.

[5] V. Vijayakumar, Y. Ampatzidis and L. Costa, 'Tree-level citrus yield prediction utilizing ground and aerial machine vision and machine learning', *Smart Agricultural Technology*, vol. 3, p. 100077, Feb. 2023. https://doi.org/10.1016/j.atech.2022.100077

[6] C. B. MacEachern, T. J. Esau, A. W. Schumann, P. J. Hennessy and Q. U. Zaman, 'Detection of fruit maturity stage and yield estimation in wild blueberry using deep learning convolutional neural networks', *Smart Agricultural Technology*, vol. 3, p. 100099, Feb. 2023. https://doi.org/10.1016/j.atech.2022.100099

[7] P. Li, J. Zheng, P. Li, H. Long, M. Li and L. Gao, "Tomato Maturity Detection and Counting Model Based on MHSA-YOLOv8," *Sensors*, vol. 23, no. 15, p. 6701, Jul. 2023, https://doi.org/10.3390/s23156701.

[8] T. Dao, N. Nguyen and V. Nguyen, "CNN-YOLOV8 - Based Tomato Quality Inspection System - a case study in Vietnam", *SSRG International Journal of Electrical and Electronics Engineering*, 10(7), 31–40, 2023, https://doi.org/10.14445/23488379/ijeee-v10i7p103

[9] X. Yue, K. Qi, X. Na, Y. Zhang, Y. Liu and C. Liu, "Improved YOLOv8-Seg Network for Instance Segmentation of Healthy and Diseased Tomato Plants in the Growth Stage," *Agriculture*, vol. 13, no. 8, p. 1643, Aug. 2023, https://doi.org/10.3390/agriculture13081643.

[10] S. Kodors, M. Sondors, G. Lācis, E. Rubauskis, I. Apeināns and I. Zarembo, "RAPID PROTOTYPING OF PEAR DETECTION NEURAL NETWORK WITH YOLO ARCHITECTURE IN PHOTOGRAPHS", *ETR*, vol. 1, pp. 81–85, Jun. 2023, https://doi.org/10.17770/etr2023vol1.7293.

[11] G. Zhao, Y. Gao, S. Gao, Y. Xu, J. Liu, C. Sun, Y. Gao, S. Liu, Z. Chen and L. Jia, "The Phenological Growth Stages of Sapindus mukorossi According to BBCH Scale," *Forests*, vol. 10, no. 6, p. 462, May 2019, https://doi.org/10.3390/f10060462.

[12] lzp-2021/1-0134, Cfruitlets81-640, 2023. [Online]. Available: https://www.kaggle.com/datasets/projectlzp201910094/cfruitlets81-640 [Accessed: Nov 12, 2023]

[13] lzp-2021/1-0134, Pear640, 2023. [Online]. Available: https://www.kaggle.com/datasets/projectlzp201910094/pear640 [Accessed: Nov 12, 2023]

[14] Ultralytics, YOLOv5-7.0 GitHub repository. [Online] Available: https://github.com/ultralytics/yolov5 [Accessed: Nov 15, 2023]

[15] Ultralytics, YOLOv8 GitHub repository. [Online] Available: https://github.com/ultralytics/ultralytics [Accessed: Nov 15, 2023]