E3S Web of Conferences **493**, 01006 (2024) https://doi.org/10.1051/e3sconf/202449301006
ABR 2024

Apple Flower Recognition Using Convolutional Neural Networks with Transfer Learning and Data Augmentation Technique

Alexey Kutyrev* and *Nikita* Andriyanov

Federal Scientific Agroengineering Center VIM, 1-st Institutsky proezd, 5, Moscow, 109428, Russia

Abstract. Automated monitoring of apple flowers using convolutional neural networks will enable informed decision-making for planning thinning and fruit set operations, optimizing crop load, preventing fruiting periodicity, and enhancing crop quality. The article presents the results of apple flower recognition quality on images using the YOLOv8 (You Only Look Once version 8) convolutional neural network model with the application of transfer learning and data augmentation technique. Pre-trained weights on the Common Objects in Context (COCO) dataset were utilized in the research. To expand the dataset and enhance model performance, the tools Flip, 90° Rotate, Crop, Rotation, Shear, Grayscale, Hue, Saturation, Brightness, Exposure, Blur, Noise, and Cutout were applied. The result showed that artificial augmentation of the training dataset significantly improves the quality of training for the YOLOv8 convolutional neural network model, increasing the average accuracy of detecting class features apple flowers. The analysis of the Precision-Recall curve allowed establishing a classification threshold (0.47) that provides the optimal balance between precision and recall in recognizing apple flowers at the flowering stage in images. The mAP metric for recognizing the «flower» class (flowers in the flowering stage) was 0.595. The analysis of the obtained results revealed an increase in the Precision metric by 2.1%, Recall metric by 10.13%, and mAP@0.5 metric by 5.31% when using the augmentation technique. The obtained results indicate a significant improvement in the performance of the model in recognizing apple flowers when applying the augmentation technique to the training dataset.

1 Introduction

Thinning apple flowers is a key factor in a crop management strategy aimed at improving fruit quality in the perspective of long-term production [1, 2]. Automated methods of flower recognition represent an efficient way to reduce labor costs in monitoring industrial apple orchards for making optimal managerial decisions [3-5]. Automated monitoring of apple flowers using convolutional neural networks will enable more precise adaptation of agronomic practices, optimizing crop load, preventing fruiting periodicity, and enhancing crop quality [6, 7]. Simulating human visual capabilities, utilizing real-time digital image

* Corresponding author: alexeykutyrev@gmail.com

processing with machine learning algorithms obtained in field conditions, is a crucial step towards automating the monitoring process of orchards. [8-10]. The aim of the research is to develop a convolutional neural network for monitoring apple flowers and analyze the performance of models using transfer learning and data augmentation technique [11, 12].

2 Materials and Methods

For systematic monitoring of orchard conditions, a methodology is employed involving the inspection of at least five control trees located diagonally in the orchard, along with regular collection of meteorological data conducted no less than once every 8-10 days. The obtained data allows for systematic tracking of flower bud formation, development of plans for thinning flowers and stimulating fruit set, as well as detecting focal points of pest and disease spread. This information serves as the foundation for making managerial decisions regarding the necessity of agricultural operations [9, 10].

Convolutional neural network models such as YOLO (You Only Look Once), SSD, Faster R-CNN, RetinaNet, EfficientDet, Mask R-CNN, as well as architectures like VGGNet, ResNet, Inception, and DenseNet, are widely used in computer vision for image recognition tasks. Such architectures has one great advantage in comparison with Vision Transformers, e.g. ViT, SWIN, DETR [13]. It is convolutional networks performance. These models demonstrate high efficiency in processing visual data, including images of biological objects. They have the ability to extract essential features and provide high accuracy in classification tasks [14, 15]. Among various neural network models, modern YOLO convolutional models stand out as the most efficient. Contemporary architectures such as YOLOv5, YOLOv7, and YOLOv8 exhibit a variety of structural features, ensuring high accuracy and efficiency in object detection with relatively low computational complexity [16, 17]. For research and monitoring of apple flowers, the YOLOv8 convolutional neural network model was utilized. The PyTorch 2.0 framework, Python libraries (tensorflow-gpu, Numpy, OpenCV, ultralyrics), Nvidia components (CUDA, CUDA toolkit, CuDNN), and the Windows 10 operating system were employed in the study.

In the training process of the convolutional neural network, the transfer learning method was applied, where a pre-trained model was utilized to address a new task. We used weights obtained from pre-training on the COCO (Common Objects in Context) dataset to optimize the model training process on our own dataset. The COCO dataset comprises an extensive collection of images with diverse annotations, covering various objects in different scenes (Fig. 1).

Fig. 1. Dataset COCO 2023 in the FityOne software environment

Adapting the model to new data allows the transfer of knowledge about object recognition from the general domain (COCO) to a more specific context. The adaptation process involves two stages: fine-tuning and retraining. During the initial fine-tuning, only the last layers are trained to adapt them to new data. Subsequently, retraining of deeper layers is performed on new data with a reduced learning rate.

To train the convolutional neural network model, a dataset of images was prepared. Industrial apple flower images with a resolution of 1280x1024 pixels were used in the research, totaling 1000 images. The image annotation process was conducted using the RoboFlow web service. A class «flower» was defined for the classification and recognition of objects representing flowers in the blooming stage. Object labeling involved outlining the regions of interest on the image with rectangular frames, followed by specifying the class (Fig. 2).

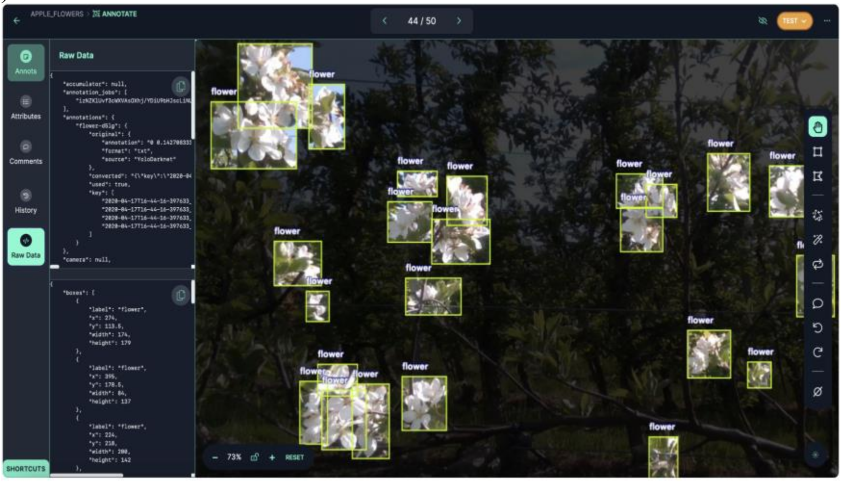


Fig. 2. Annotating apple flower images

To augment the dataset and enhance the model's performance, an augmentation technique was applied, resulting in the expansion of the dataset to 3000 images (x3). The tools used include Flip (Horizontal, Vertical), 90° Rotate (Clockwise, Counter-Clockwise), Crop (0% Minimum Zoom, 20% Maximum Zoom), Rotation (Between -15° and +15°), Shear (+15° Horizontal, +15° Vertical), Grayscale (Apply to 25% of images), Hue (Between -25° and +25°), Saturation (Between -25% and +25%), Brightness (Between -25% and +25%), Exposure (Between -25% and +25%), Blur (Up to 2.5px), Noise (Up to 5% of pixels), Cutout (20 boxes with 10% size each). Figure 3 shows different examples for used augmentations.

The recognition quality of the YOLOv8 neural network model was evaluated during training with and without the use of additional datasets. A heatmap in the form of a two-dimensional array was developed, where each image pixel determines intensity according to a color scale. Areas that the model considers most likely to contain the specified classes of apple flowers were highlighted. Figure 4 presents the results of annotation and obtained heatmap.

 https://doi.org/10.1051/e3sconf/202449301006

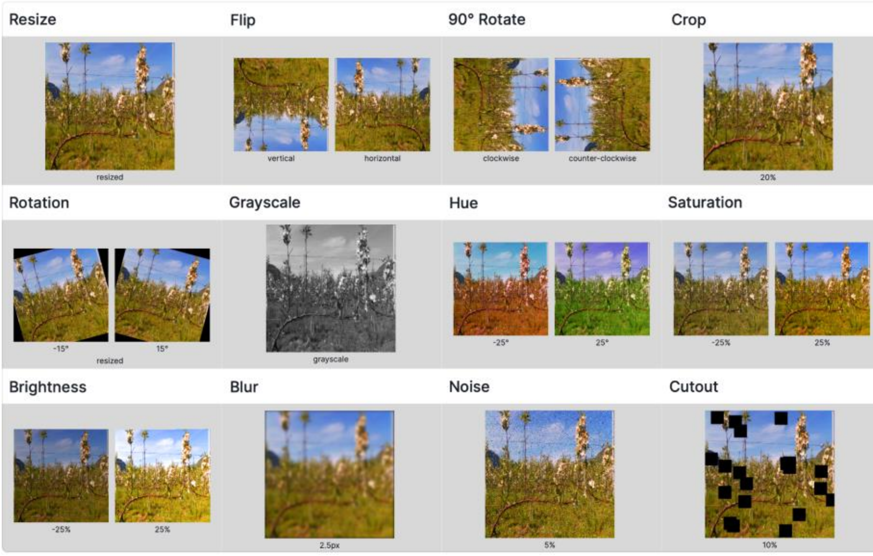


Fig. 3. The utilized augmentation techniques for apple flowers images

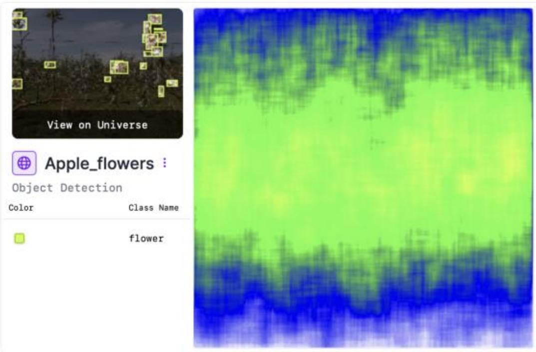


Fig. 4. Annotations hitmap apple flowers images

As a result of dataset augmentation, the dataset was expanded to 3000 images (Fig.5).

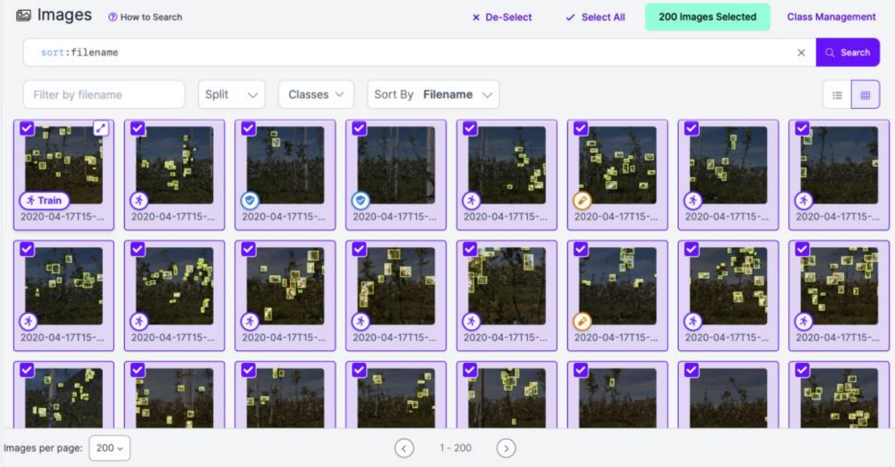


Fig. 5. The obtained annotated dataset of apple flower images

To assess the performance of the models, metrics such as Precision, Recall, F1-score, and mean Average Precision (mAP) were employed (Table 1).

Table 1. Metrics used to analyze the quality YOLOv8 models when recognizing apple flowers

Metric	Formula
Precision	$\frac{TP_i}{TP_i + FP_i}$
Recall	$\frac{TP_i}{TP_i + FN_i}$
F1-score	$\frac{1}{C}\sum_{i=1}^{C}\frac{2 \cdot \mathrm{Precision}_i \cdot \mathrm{Recall}_i}{\mathrm{Precision}_i + \mathrm{Recall}_i}$
mAP	$\frac{1}{C}\sum_{i=1}^{C} AP_i$

In these formulas, TP_i, FP_i and FN_i refer to true positives, false positives, and false negatives for class i, respectively.

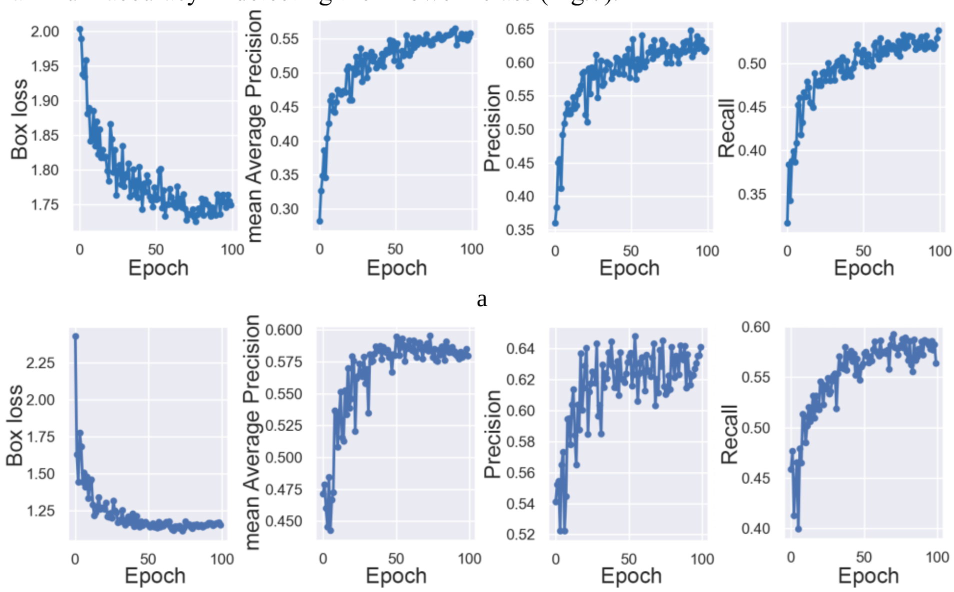
The research was conducted using a computational system with an Intel Core i9-10900X processor (20 virtual threads), NVIDIA GeForce RTX 2080 Ti graphics cards (2 units), GIGABYTE X299 UD4 Pro motherboard, Intel PCI-E SSD with 1 TB storage for data, and 32 GB Kingston DDR4 DIMM RAM.

3 Results and Discussion

Examples of detection and recognizing the «flower» class in the images of the test dataset using the trained YOLOv8 model with object bounding boxes highlighted are presented in Figure 6.

Fig. 6. Recognition results model YOLOv8

Analysis of the Box Loss-Epoch graphs during training, with and without augmentation, allowed determining the optimal number of epochs (74) for the highest quality detection of object bounding box coordinates. Precision-Epoch and Recall-Epoch curves were used to assess changes in precision and recall during training. The mAP-Epoch curve was plotted to evaluate changes in the model's average precision. Analysis of these curves enabled the selection of the optimal number of epochs (86) for achieving the best performance and maximum accuracy in detecting the «flower» class (Fig.7).

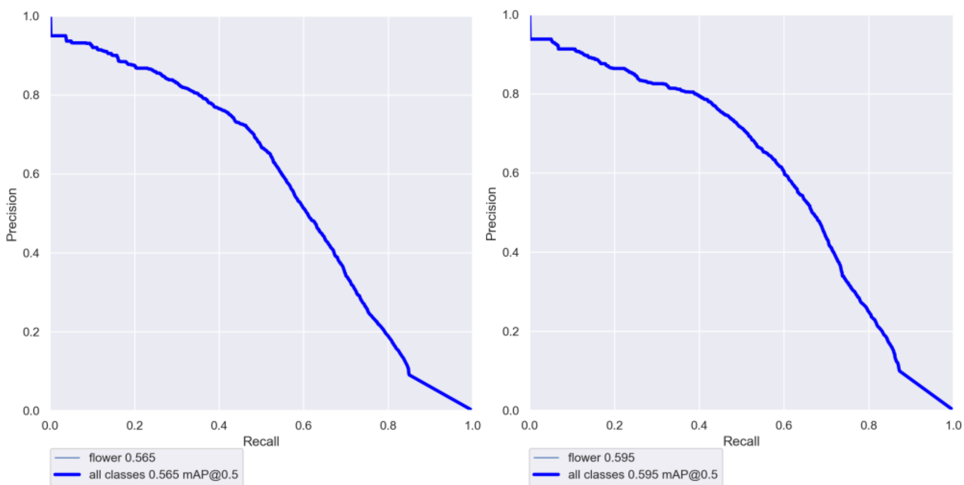


a

b

Fig. 7. Training results of YOLOv8 models: a – without using augmentation technique, b – with the use of augmentation technique

To evaluate the accuracy and recall values with the change in the decision threshold in the classification task, Precision-Recall and Recall-Confidence curves were plotted (Fig.8). The Confidence metric is defined as the maximum probability of an object belonging to one of the classes.



a b

Fig. 8. Precision-Recall Curves: a – without using augmentation technique, b – with the use of augmentation technique

Analysis of the Precision-Recall curve allowed to determine the classification threshold (0.47) providing the best balance between precision and recall in recognizing apple flowers. The obtained F1-Confidence graph allowed evaluating the influence of the model's confidence level on the combined metrics of precision and recall, as well as selecting the optimal threshold (0.32) for decision-making in classification (Fig.9).

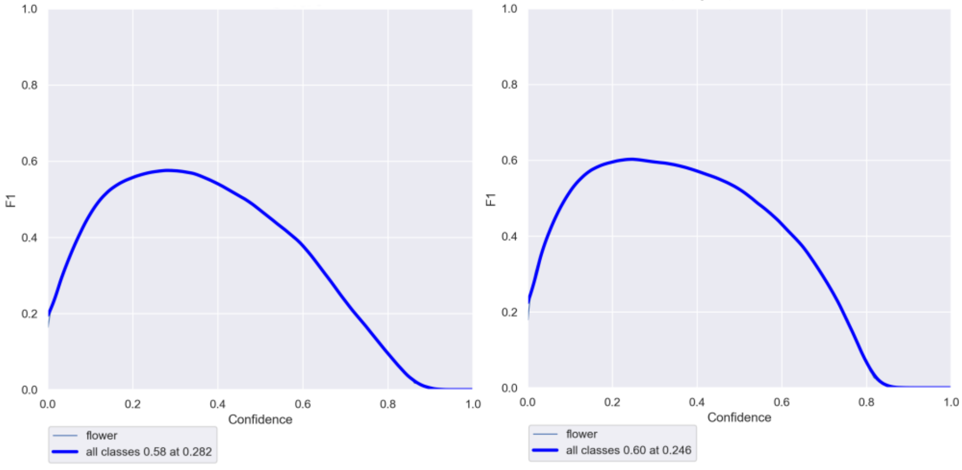


Fig. 9. Curves of F1-score – Confidence: a – without using augmentation technique, b – with the use of augmentation technique

The values of metrics for YOLOv8 models in recognizing the «flower» class with and without the use of augmentation technique on the overall dataset are presented in Table 2.

Table 2. Results of binary classification metrics calculation for YOLOv8 models in recognizing the «flower» class with and without using augmentation technique

Classifier	Precision	Recall	mAP@0,5
Without using augmentation technique			
Class «flower»	0,621	0,523	0,565
With using augmentation technique			
Class «flower»	0,634	0,576	0,595

The analysis of the obtained results revealed an increase in Precision metric by 2.1%, Recall metric by 10.13%, and mAP@0.5 metric by 5.31% when utilizing the augmentation technique. These findings indicate a significant improvement in the model's performance in recognizing apple flowers when applying the augmentation technique to the training dataset.

4 Conclusions

The settings for the machine learning algorithm for the YOLOv8 model, designed for apple flowers recognition at the flowering stage, are as follows: learning rate – 0.01 LR, number of epochs – 100, batch size – 8. The use of artificial augmentation in the training dataset, including tools such as Flip, 90° Rotate, Crop, Rotation, Shear, Grayscale, Hue, Saturation, Brightness, Exposure, Blur, Noise, Cutout, significantly improves the model's training quality. This approach contributes to the adaptation of the model to real conditions, resulting in a 5.31% increase in the average accuracy of detecting class features compared to the dataset without volume augmentation. The conducted research indicates the prospects of integrating convolutional neural networks into decision support systems for monitoring and planning the technological operation of thinning apple flowers and forming fruit sets. The ability of CNN models to accurately recognize flowers in the early stages of development dynamically enables informed decisions to optimize agricultural processes even before the onset of active flowering, considering the uneven development of orchards.

References

1. H. Zhou, J. Ou, P. Meng, J. Tong, H. Ye, Z. Li, Horticulturae, **9**, 400 (2023)
2. P.A. Dias, A. Tabb, H. Medeiros, Comput. Ind, **99**, 17–28 (2018)
3. M. Tian, H. Chen, Q. J. Phys. Conf. Ser., 1237, 032045 (2019)
4. B.V. Biradar, S.P. Shrikhande, Int. J. Comput. Sci. Inform. Technol, **6**, 2498–2501 (2015)
5. B. Shurygin, I.G. Smirnov, A. Chilikin, D.O. Khort., A.I. Kutyrev, S. Zhukovskaya, A. Solovchenko, Horticulturae, **8**(12), 1111 (2022)
6. G. Farjon, O. Krikeb, A.B. Hillel, V. Alchanatis, Precis. Agric, **21**, 503–521 (2020)
7. A.I. Kutyrev, R.A. Filippov, Agricultural Science Euro-North-East, **24**(4), 685–696 (2023)
8. K. Sun, X. Wang, S. Liu, C. Liu, Comput. Electron. Agric., **185**, 106150 (2021)
9. J. Lee, S.A. Gadsden, M. Biglarbegian, J.A. Cline, Appl. Sci, **12**, 11420 (2022)
10. A. Kutyrev, K. Nikolay, M. Jewiarz, K. Dmitriy, I. Smirnov, V. Zubina, T. Hutsol, M. Tomasik, M. Biliuk, Sensors., **22**(22), 8901 (2022)
11. A. Buslaev, V. Iglovikov, E. Khvedchenya, A. Parinov, M. Druzhinin, A. Kalinin, Information, **11**, 125 (2020)
12. V.E. Dementyiev, N.A. Andriyanov, K.K. Vasilyiev, 2020 Systems of Signal Synchronization, Generating and Processing in Telecommunications (SYNCHROINFO), 1-4 (2020)
13. N.A. Andriyanov, V.E. Dementiev, A.G. Tashlinskiy, Computer Optics, **46**(1), 139-159 (2022)
14. D. Khort, A. Kutyrev, R. Filippov, N. Kiktev, D. Komarchuk 2019 IEEE International Scientific-Practical Conference: Problems of Infocommunications Science and Technology, 869-872 (2019)
15. I. Smirnov, A. Kutyrev, N. Kiktev, E3S Web of Conferences, **270**, 01021 (2021)
16. S. Lyu, Y. Zhao, R. Li, Z. Li, R. Fan, Q. Li, Sensors, **22**, 1255 (2022)
17. D. Wu, S. Lv, M. Jiang, H. Song, Comput. Electron. Agric., **178**, 105742 (2020)