

13. Burinska, Z., Runovski, K., Sehmeisser, H.-J. (2006). On the Approximation by Generalized Sampling Series in Lp-Metrics. *Sampling Theory in Signal and Image Processing*, 5 (1), 59–87. doi: <https://doi.org/10.1007/bf03549443>
14. Dengler, J., Wagner, V., Dembicz, I., Garc a-Mijangos, I., Naqinezhad, A., Boch, S. et. al. (2018). GrassPlot – a database of multi-scale plant diversity in Palaearctic grasslands. *Phytocoenologia*, 48 (3), 331–347. doi: <https://doi.org/10.1127/phyto/2018/0267>

The detection of weeds at the stages of cultivation is very important for detecting and preventing plant diseases and eliminating significant crop losses, and traditional methods of performing this process require large costs and human resources, in addition to exposing workers to the risk of contamination with harmful chemicals. To solve the above tasks, also in order to save herbicides and pesticides, to obtain environmentally friendly products, a program for detecting agricultural pests using the classical K-Nearest Neighbors, Random Forest and Decision Tree algorithms, as well as YOLOv5 neural network, is proposed. After analyzing the geographical areas of the country, from the images of the collected weeds, a proprietary database with more than 1000 images for each class was formed. A brief review of the researchers' scientific papers describing the methods they developed for identifying, classifying and discriminating weeds based on machine learning algorithms, convolutional neural networks and deep learning algorithms is given. As a result of the research, a weed detection system based on the YOLOv5 architecture was developed and quality estimates of the above algorithms were obtained. According to the results of the assessment, the accuracy of weed detection by the K-Nearest Neighbors, Random Forest and Decision Tree classifiers was 83.3 %, 87.5 %, and 80 %. Due to the fact that the images of weeds of each species differ in resolution and level of illumination, the results of the neural network have corresponding indicators in the intervals of 0.82–0.92 for each class. Quantitative results obtained on real data demonstrate that the proposed approach can provide good results in classifying low-resolution images of weeds

Keywords: agriculture, weeds, machine learning, YOLOv5, segmentation, Otsu's method, classification, algorithm evaluation

UDC 004.8

DOI: 10.15587/1729-4061.2021.246706

DEVELOPMENT OF A WEED DETECTION SYSTEM USING MACHINE LEARNING AND NEURAL NETWORK ALGORITHMS

Baydaulet Urmashiev
Professor*

Zholdas Buribayev
Corresponding author
Master*

E-mail: zholdas_87@mail.ru

Zhazira Amirgaliyeva

Doctor of Technical Sciences, Professor
Laboratory of Artificial Intelligence and Robotics
Institute of Information and Computational Technologies
Shevchenko str., 28, Almaty, Republic of Kazakhstan, 050010

Aisulu Ataniyazova
Masters Student*

Mukhtar Zhassuzak
PhD Student*

Amir Turegali
Bachelor Student*

*Department of Computer Science
Al-Farabi Kazakh National University
Al-Farabi ave., 71, Almaty, Republic of Kazakhstan, 050040

Received date 03.11.2021

Accepted date 05.12.2021

Published date 29.12.2021

How to Cite: Urmashiev, B., Buribayev, Z., Amirgaliyeva, Z., Ataniyazova, A., Zhassuzak, M., Turegali, A. (2021). Development of a weed detection system using machine learning and neural network algorithms. *Eastern-European Journal of Enterprise Technologies*, 6 (2 (114)), 70–85. doi: <https://doi.org/10.15587/1729-4061.2021.246706>

1. Introduction

The agricultural sector is one of the main branches of the economy of our country, since this industry annually provides 35–40 % of income to the state budget, and 15 % of the entire labor force of the country is employed in this sector. Weed control and monitoring of crop diseases have become an urgent task in the robotization of agriculture [1]. Monitoring of diseases and weeds at the stages of cultivation is very important for detecting and preventing diseases and eliminating significant crop losses, and traditional methods of performing this process require high costs and human resources, besides exposing workers to the risk of contamination with harmful chemicals. Therefore, the development of a pest control system that performs the detection and removal of weeds is the main area of research in the agricultural industry.

At the present time, the most optimal means for pest control is the large-scale use of herbicides, but the fact of uneven growth

of weeds is not taken into account. As a result, crops also come under treatment with chemicals used to kill weeds, which can harm the environment. Previously used technologies could only distinguish between the presence or absence of plants, they were not capable of dividing them into weeds and agricultural crops. New technologies allow for more efficient spraying of herbicides, using them only in the right areas to preserve crops and protect the environment [2, 3]. The introduction of intelligent weed detection systems will also solve the problem of saving herbicides and pesticides, which are in demand means to combat plant diseases, various weeds and vectors of dangerous diseases in industrial agricultural production.

The use of autonomous robots and automated systems in agriculture can lead to a significant minimization of human efforts required to perform several agricultural tasks. To solve these problems, new classification systems have been proposed that can identify agricultural crops, distinguishing them from undesirable harmful vegetation [4, 5].

Since agriculture is the second leading branch of material production, and more than 1 billion people are engaged in growing grain, vegetable and fruit crops, automation and robotization of some tasks in this industry will significantly increase efficiency and can replace heavy agricultural machinery, and systems for distinguishing weeds from agricultural crops can lead to significant savings in chemicals by applying them only to the leaves of weeds. Also, harvesting, sorting by quality and size of harvested crops increase expenditures and lead to an increase in the crop cost, so it is necessary to conduct scientific research in the field of weed detection and discrimination to organize fast and accurate work on agricultural fields.

The results of the study make it possible to use these robot systems, which will allow achieving productivity growth, reducing overspending of materials used and increasing the quality of the crop due to precision farming.

2. Literature review and problem statement

In the scientific work [5], segmentation based on the 3D Otsu's method was used to distinguish agricultural crops and weeds, and classification was performed by compressing three-dimensional image vectors using the Principal Component Analysis (PCA) method. Combining the two methods, the authors proposed a real-time weed detection program.

The authors of [6] conducted a review of computer vision methods for determining the location of weeds and crops. The paper analyzed the advantages and disadvantages of algorithms based on deep learning and considered aspects of using these methods to solve problems of weed detection. In the course of the study, it was determined that the AgroAVNET (a hybrid model of AlexNet and VGGNET), Graph Weeds Net and MTD (multiscale triangle descriptor) and LBP-HF (local binary pattern histogram Fourier) methods have high accuracy of weed detection compared to other algorithms. This paper presents a good comparative analysis of various architectures and algorithms within the same data set of weeds.

A similar work is [7], which offers deep learning using an image processing framework for classifying various crops and weeds. The result of the work achieved 95 % accuracy using a convolutional neural network and max pooling layers supported by a reduced frequency of incorrect classification of weeds.

In [8], a model for detecting the presence of weeds in the fields of agricultural crops based on deep learning is proposed. To extract informative features, the authors used a convolutional neural network, and the creation of additional images for the training sample was performed by various augmentation techniques, and Inception V3 was used as a function extractor. Classification of weeds was carried out using U-Net. According to the results of testing the system on 158 figures, the detection accuracy was 90 %. However, the authors did not compare the system proposed by them with the works of other authors, it would also be better if the studies were carried out on a large data set, taking into account different lighting conditions.

In the paper [9], the k-means algorithm was used in combination with a convolutional neural network in order to detect weeds associated with soybean seedlings. Using the method of training uncontrolled signs with k-means as a preliminary preparation, better parameter values were

obtained to increase the accuracy of weed identification. The proposed method has achieved 92.89 % accuracy and exceeds the detection results using convolutional neural networks with random initialization and two-layer networks without fine-tuning parameters. Thus, the authors have shown that fine-tuning the parameters significantly affects the increase in accuracy.

The research carried out in [10] is useful for the development of weed seed detection systems, since accurate classification of weed seeds is important for the elimination of crop pests. In the course of the study, six models of a deep convolutional neural network were compared to determine the best method for detecting weed species. As a result, the GoogLeNet architecture has shown high accuracy, and the SqueezeNet architecture is superior in terms of seed and background detection time. But it would not hurt the authors to improve the accuracy of identification by considering cases of fusion of several types of seeds and their close location.

The scientific work [11] is aimed at developing a method of segmentation and detection of weeds based on Mask R-CNN. The ResNet-101 network was used to extract a map of semantic and spatial information about weeds. The calculation of the loss of classification, regression and segmentation is done by the output modules. The average accuracy of the Mask R-CNN neural network was 0.853, which is better than the Sharp Mask and DeepMask algorithms.

The authors of the scientific work [12] have developed a program that performs the detection of weeds in crops, as well as the distinction of weeds of herbaceous and broad-leaved species. Using convolutional networks, the researchers obtained results with high accuracy in the classification of all classes of weeds. But the authors did not evaluate the algorithms they used to classify weeds, which would allow determining the best algorithm in terms of quality, performance and speed.

The paper [13] provides a generalized overview of the achievements in the field of weed detection using machine vision and image processing methods. A detailed description of procedures such as preprocessing, segmentation, feature extraction and classification is provided. The authors also discussed the problems that arise when detecting weeds and solutions that allow recognizing in different lighting conditions and at different stages of weed growth.

The proposed method based on a fully convolutional network in the paper [14] provides higher classification accuracy and can effectively classify pixels of rice seedlings, background and weeds in images of rice fields and determine the position of their areas. This approach has been compared with the classical semantic segmentation models of AST and G-Net and surpasses them in some parameters.

The methodology presented in the paper [15] consists of two stages. At the first stage, background segmentation is performed using maximum likelihood, and the second stage is devoted to manual marking of weeds. A comparative analysis of the deep learning architectures of SegNet and UNET was carried out, and the results and advantages of the architectures used were revealed by evaluating the methodology.

The researchers of the scientific work [16] carried out an analysis of morphological features for the classification of agricultural crops and weeds in agricultural production sectors. Based on descriptors of singular points, such as a histogram of directional gradients and local binary patterns, the authors presented a method for extracting features with the least computational complexity and with a higher resolution.

In the paper [17], the authors consider the problem of automating the process of weed removal using machine learning algorithms. The collected data set consists of 4 types of commercial crops and 2 types of weeds. In this paper, the performance of classification algorithms, artificial neural networks and convolutional neural networks is compared.

As a result of our review, the following factors and shortcomings were identified in the field of weed detection research:

- the effect of uneven lighting on color images of various weeds: in most cases, the images were generated in shades of gray or the Hue, Saturation, and Intensity color model was used for image processing. Changing lighting conditions significantly affect the accuracy and reliability of object detection;
- changes in morphology and texture at different stages of leaf growth also negatively affect detection, as the process of distinguishing weeds from crops becomes more difficult due to the same level of their growth;
- the complexity of the algorithm used also limits the speed of weed identification, so the methods used should be optimized for fast image processing.

All this suggests that it is advisable to conduct a study on the detection of weeds with high accuracy and the least errors based on computer vision, using images not only of good quality, but also low resolution and poor illumination.

3. The aim and objectives of the study

The aim of this study is to develop a system for detecting weeds and distinguishing them from agricultural crops based on machine learning algorithms and neural networks using its own data set. To achieve this goal, the following objectives were set:

- to explore different geographical areas to identify common types of weeds, analyze the main types of grasses that are often found in the fields;
- to collect data for processing and training neural networks such as YOLOv5, and machine learning algorithms KNN, Random Forest and Decision Tree;
- to improve the architecture of the YOLOv5 neural network with the supplementation of additional modules and check the work effectiveness using evaluation metrics.

4. Materials and methods

4. 1. Segmentation

For image segmentation, the Otsu’s method is selected [18], which is an adaptive algorithm based on binarization. The algorithm uses the maximum variance value, that is, deviations from the average brightness between the background and the selected image as the threshold selection rule. First, there is a process of dividing the image into foreground and background in accordance with its characteristics of the gray scale. If you select the best threshold value, the difference between the two parts will increase. The probability of incorrect classification is minimized when the difference between the background and the target image has the maximum value. Segmentation of images by the Otsu’s method is carried out as follows:

1. The original image is divided into $l=[0, 1, \dots, l-1]$ levels. The number of pixels at a certain level i is denoted by n_i ,

and the total number of pixels is denoted by the sum of all n_i , that is, $n=n_1+n_2+\dots+n_l$.

2. Image pixels are divided into classes with gray levels without a threshold and with a threshold. The probability of a gray level is calculated to distribute them into classes.

3. The average gray value is calculated using

$$u_i = \sum_{i=0}^l \frac{iP_i}{w_i}, \tag{1}$$

where P_i is the probability of the gray level, w_i is the probability of the distribution of gray levels into classes. The total average is denoted by ut and is determined by the sum of all u_i .

4. Finding the variance of each class, the interclass variance, the total variance of gray levels is a key step of the Otsu’s method, since by maximizing the interclass variance, the optimal threshold is selected and the ideal image segmentation is realized.

4. 2. Classification

Three classical machine learning algorithms were used to classify weeds, such as K-Nearest Neighbors, Random Forest, Decision Tree. Since the classifier must use a class-balanced training set to be effective [19], a random sampling process was implemented to select the same number of objects for each class in the entire image. Then the values of NIR/G (Near-Infrared/Green), average red, average green, average NIR, brightness, standard deviation of NIR were extracted for each object making up the training set and used as features to distinguish weeds, crops and bare soil using RF. 400 decision trees were used in the training process, as this value turned out to be acceptable when using an RF classifier. In order to avoid any misclassification of large weeds between or within rows, the height of the object was not included as a classification parameter.

4. 3. Evaluation of algorithms

Many metrics have been developed to assess the quality of machine learning algorithms. All metrics are calculated using combinations of the error matrix, which provide information about the numbers of true-positive, true-negative, false-positive and false-negative solutions of the classifier. The functions FR, FR, recall, precision, accuracy, and the Jacquard index use two or three combinations of the inaccuracy matrix and do not give an objective assessment of the classification results [20]. And the metrics $F1$, Cohen’s kappa, and the Matthews correlation coefficient, using all the elements of the error matrix, evaluate the results of the classifier with unbalanced data. Below are the formulas for finding these metrics:

1. The proportion of errors made by the classifier when assigning one or another object to the selected class is demonstrated by the FPR metric, a false-positive rate. The value of this metric depends on the number of false-positive and true-negative solutions.

$$FPR = \frac{FP}{FP + TN}. \tag{2}$$

2. A false-negative rate demonstrates a second kind of error when the machine learning model predicts a negative decision, but in fact it is an object of the selected class.

$$FNR = \frac{FN}{FN + TP}. \tag{3}$$

3. The recall metric shows how many examples of positive solutions were lost as a result of classification. It is responsible for the ability to detect objects of a certain class; therefore it is determined using true-positive and false-negative solutions.

$$\text{Recall} = \frac{TP}{TP + FN}. \quad (4)$$

4. Precision demonstrates the ability of the algorithm to distinguish the selected class from other classes, but unlike recall, it takes into account false-positive results.

$$\text{Precision} = \frac{TP}{TP + FP}. \quad (5)$$

5. Accuracy calculates the proportion of correct classifications and is defined as the ratio of all true results and the sum of all combinations of the confusion matrix.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN}. \quad (6)$$

6. *F1*-measure is a metric that reduces two main evaluation metrics to one number: precision and recall. It is needed for balancing when the maximum values of precision and recall are not achievable at the same time.

$$F1 = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}. \quad (7)$$

7. The Jaccard index is used to detect faces from an image, as it is able to quantify the similarity between the identification of computer faces with the identification of training data. Therefore, this index is important for semantic image segmentation.

$$\text{Jaccard index} = \frac{TP}{TP + FN + FP}. \quad (8)$$

8. Cohen's kappa, like other evaluation indicators, is calculated based on the confusion matrix. Unlike calculating the overall accuracy, Cohen's kappa takes into account the imbalance in the distribution of classes. Therefore, as shown in the formula, it excludes the possibility of a classifier and a random guess matching and measures the number of predictions it makes.

$$\begin{aligned} \text{Cohen's kappa} &= \\ &= \frac{(TP + FP)(TP + FN) + (FN + TN)(FP + TN)}{(TP + FP + FN + TN)^2}. \end{aligned} \quad (9)$$

9. The Matthews correlation coefficient (MCC) is used in cases where class sizes vary greatly, and the data set of negative and positive are unbalanced. The indicator of this metric depends on the product of the true and false solutions of the classifier separately, as well as on the subtraction of these two numbers.

$$\begin{aligned} \text{MCC} &= \\ &= \frac{(TP * TN - FP * FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}. \end{aligned} \quad (10)$$

10. Specificity or a true-negative indicator is responsible for the probability that the classifier does not correctly assign objects to the selected class, ignoring the error of the second kind, that is, the number of false-negative solutions.

$$\text{Specificity} = \frac{TN}{TN + FP}. \quad (11)$$

11. The Youden index is determined by the difference between the proportion of true-positive results and the proportion of false-positive decisions.

$$\text{Youden index} = \text{Recall} + \text{Specificity} - 1. \quad (12)$$

All these metrics are used to assess the quality of classifiers, determining how well they correctly predict which class the target object belongs to.

5. Research results of object recognition by neural network architecture and machine learning algorithms

5.1. Study of geographical locations to identify common types of weeds

To solve the problem of studying geographical places to identify varieties of weeds, trips were organized to the village of Koram in the Enbekshikazakh district, to the village of Saty in the Raiymbek district and to the village of Kyrgauyldy in the Karasay district. After analyzing the types of herbs that grow in these visited places, it was decided that there are 5 main types of herbs that are common to all these three places: ambrosia, amaranthus, bindweed, bromus and quinoa (Fig. 1).

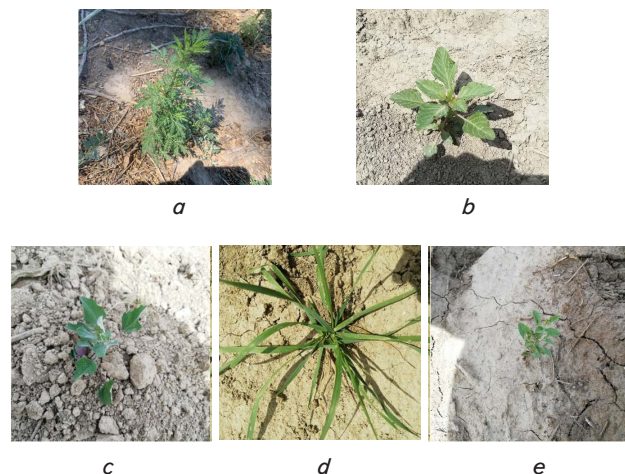


Fig. 1. Types of collected weeds: *a* – Ambrosia; *b* – Amaranthus; *c* – Quinoa; *d* – Bromus; *e* – Bindweed

Having selected the types of herbs to collect data, there were repeated trips to the above-mentioned areas and about 1,000 photos were photographed for each species.

5.2. Data collection for processing and training neural networks and classical algorithms

As shown in Fig. 2, the data set for classification using machine learning algorithms contains 4 types of weeds. Each class under consideration was segmented and stored in the dataset as an array.

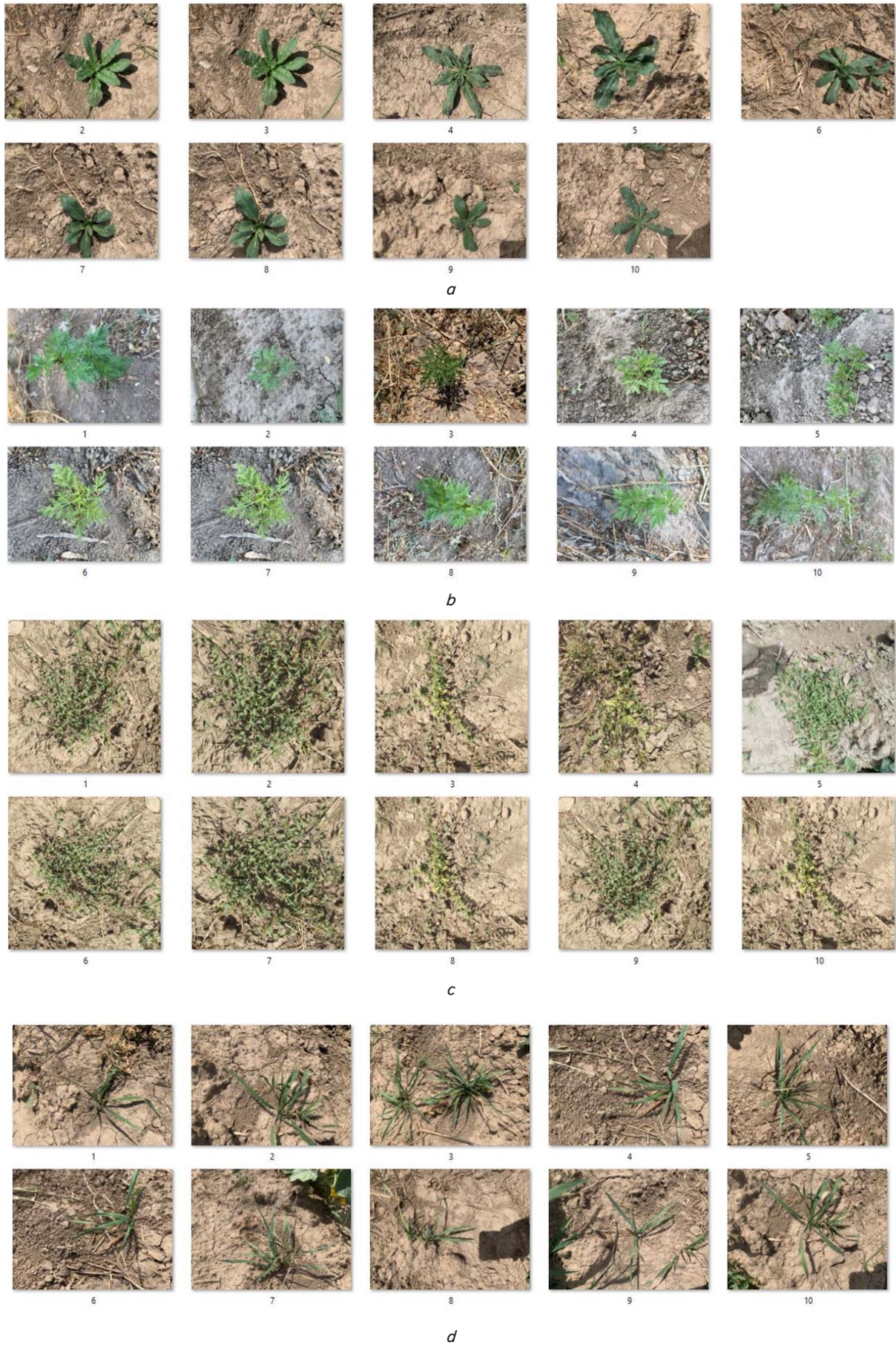


Fig. 2. Formation of the dataset: *a* – Amaranthus; *b* – Ambrosia; *c* – Bindweed; *d* – Bromus

One of the main tasks of image processing is the segmentation process, and it consists of several stages. The first stage involves converting the image (Fig. 3, *a*) to shades of gray, from 0 (black) to 255 (white). The image after processing looks as follows (Fig. 3, *b*).

The next step is binarization. Its main purpose is to reduce the amount of information in the image. A popular image binarization method, the Otsu's method, was used here. After binarization by the Otsu's method, a small noise can be seen in the image (Fig. 4, *a*). Therefore, the final stage of noise removal is performed. The processed image is shown in Fig. 4, *b*.

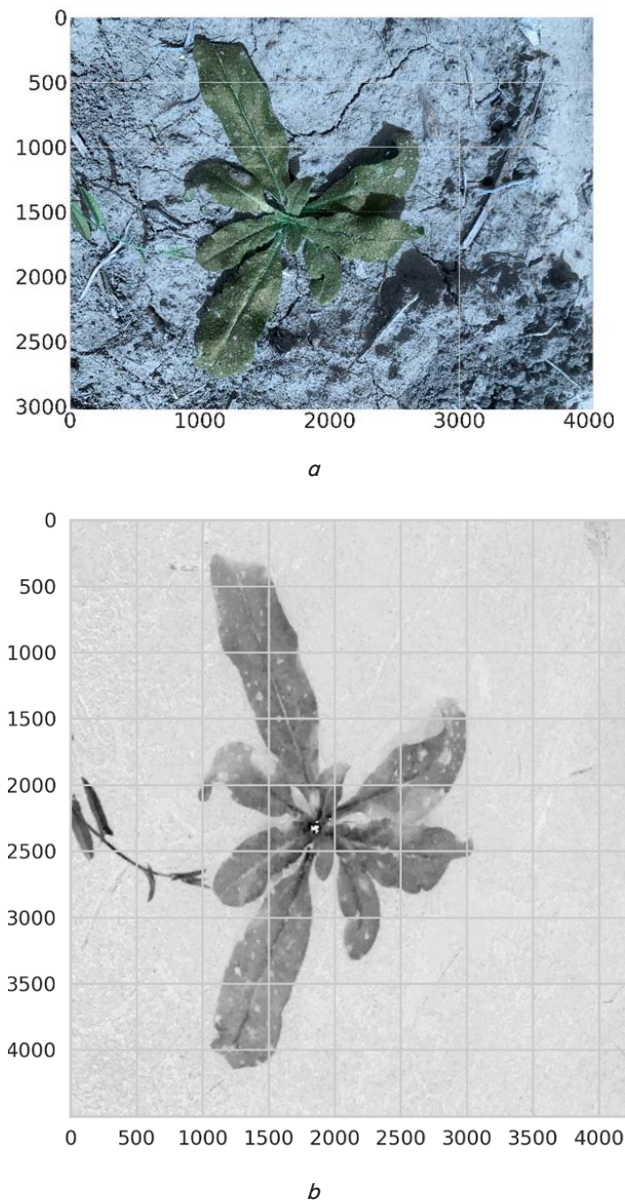


Fig. 3. Weed image: *a* – original image; *b* – conversion to grayscale

Reducing the influence of extraneous noise improves the image quality, thus, after preparing the data, segmenting them into the target object and background, classification of weeds is performed.

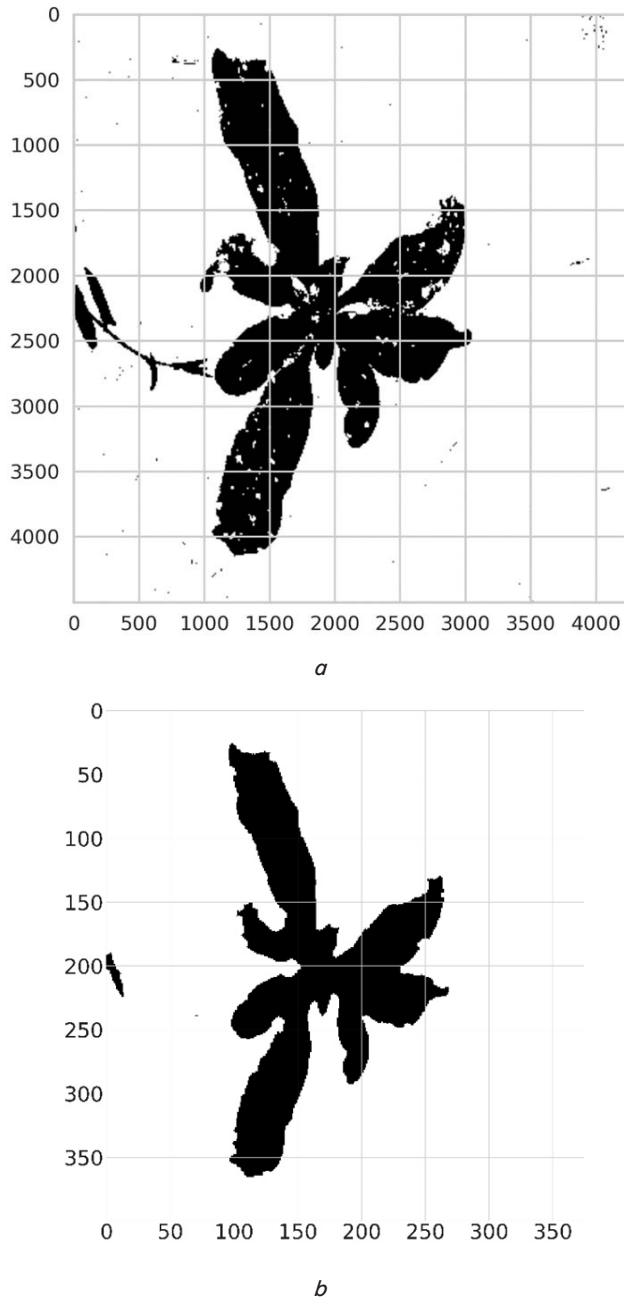


Fig. 4. Image segmentation: *a* – by the Otsu's method; *b* – with noise reduction

5. 3. Comparison of machine learning algorithms results and development of an improved YOLO architecture

5. 3. 1. Classification of weeds using machine learning algorithms

Fig. 5, *a* shows a general matrix of algorithm errors for all classes of weeds. The number of true-positive solutions (TP) predicted by the KNN classifier is 8, and the number of true-negative combinations (TN) that do not belong to the selected class and were classified as negative correctly is 32. The false-negative combination is responsible for the number of erroneously predicted solutions, that is, the classifier predicted them as negative objects, but in reality they are positive objects belonging to the selected class, the number of FN is 4.

The matrix of inaccuracies of the Random Forest classification model is shown in Fig. 5, *b*. The number of correctly predicted objects is 9, and the number of true-negative combinations is 33. The number of errors that the classifier made when recognizing weeds is 3.

According to the confusion matrix of the Decision Tree algorithm (Fig. 5, *c*), it can be seen that the number of true-positive decisions of the classifier is 7. The number of TN combinations is 31, and the number of errors of the first and second kind made by the machine learning model is 5. Since there are fewer correctly predicted objects than KNN and RF, accuracy has a low indicator.

Using combinations of the error matrix, metrics are calculated that determine the quality of the classifier. Nu-

merical indicators of the KNN algorithm estimates for each class are presented in Table 1.

Table 1

Estimates for each class of the K-Nearest Neighbors algorithm

Class/Metrics	Precision	Recall	F1	Accuracy
Amaranthus	0.67	0.50	0.57	0.75
Ambrosia	1.00	0.75	0.86	0.92
Bindweed	1.00	1.00	1.00	1.00
Bromus	0.40	0.67	0.50	0.67

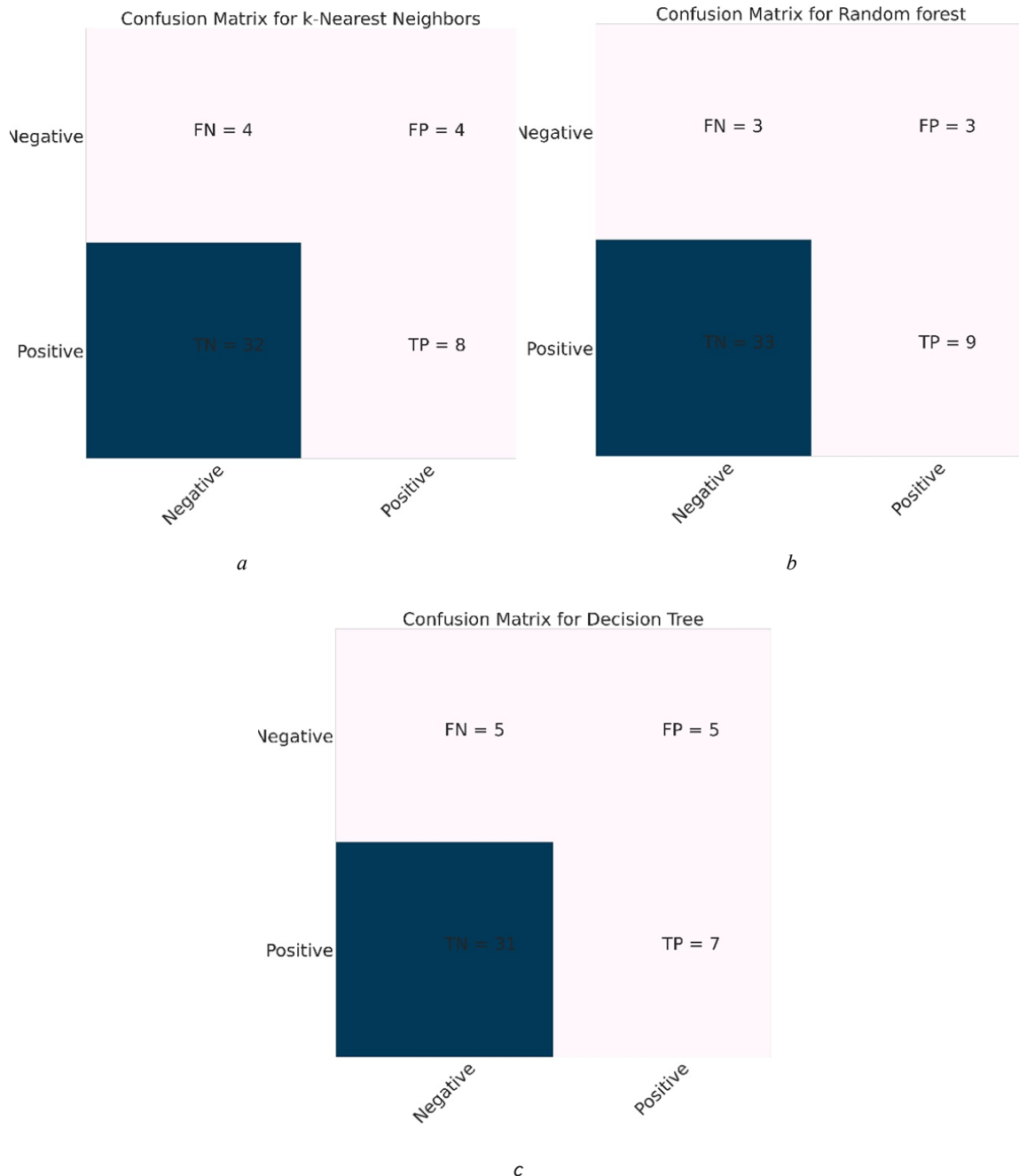


Fig. 5. Confusion matrix: *a* – for K-Nearest Neighbors; *b* – for Random Forest; *c* – for Decision Tree

According to (5), the precision of the plant class “Amaranthus” is 0.67, the classes “Ambrosia” and “Bindweed” are equal to 1, which shows the exact classification of objects of these classes, and for weeds “Bromus” is 0.4. The average precision value was 0.67, since the classifier mistakenly predicted objects of the “Bromus” class as objects of other classes, and this affected the overall quality. The next metric, recall, is calculated by (6). For objects of the “Amaranthus” class, the value of this metric is 0.5, for weeds of the “Ambrosia” type – 0.75, for the “Bindweed” class, the recall score has reached a maximum, that is, 1, for weeds of the “Bromus” class is 0.67. The overall score of the recall metric is also 0.67, as is precision. Therefore, the value of the F1 metric, which is the harmonic mean of these two metrics, will be 0.67. Based on the results of the assessment, it can be concluded that the KNN algorithm incorrectly classifies 1/3 of all objects.

Compared to the KNN algorithm, Random Forest performed the classification better, so the metrics of the evaluation are also higher than those of the previous algorithm (Table 2).

Table 2

Estimates for each class of the Random Forest algorithm

Class/Metrics	Precision	Recall	F1	Accuracy
Amaranthus	1.00	0.50	0.67	0.83
Ambrosia	1.00	0.75	0.86	0.92
Bindweed	0.33	1.00	0.50	0.83
Bromus	0.75	1.00	0.86	0.92

The precision metric value for “Amaranthus” and “Ambrosia” objects is 1, for “Bindweed” class objects – 0.33, for “Bromus” – 0.75, and the overall precision score was 0.75, this is due to the fact that many objects of the “Bindweed” type were mistakenly classified as “Bromus” class objects and lowered the average value of this metric. And the recall index of the “Bindweed” and “Bromus” classes is higher than the value of this metric of the “Amaranthus” class, since more correct predictions were lost when classifying objects of the first class. The accuracy of the Random Forest classifier for two classes, in addition to “Amaranthus” and “Bindweed” amounted to 0.92, due to low ratings recall class “Amaranthus” and precision class “Bindweed”, their recognition accuracy was only 0.83.

The most incorrect predictions were made in the classification of weed species “Amaranthus”, so the precision, recall and accuracy in this class have a lower value than other classes. Numerical indicators are presented in Table 3.

Table 3

Estimates for each class of the Decision Tree algorithm

Class/Metrics	Precision	Recall	F1	Accuracy
Amaranthus	0.00	0.00	0.00	0.67
Ambrosia	0.67	1.00	0.80	0.83
Bindweed	1.00	1.00	1.00	1.00
Bromus	0.40	0.67	0.50	0.67

Using the F1 metric, precision and recall are balanced, thus, for the Decision Tree classifier, their values were 0.8, 1.0 and 0.5, respectively, for “Ambrosia”, “Bindweed” and “Bromus”.

Table 4 presents numerical indicators of all metrics for evaluating algorithms that were used to classify weeds.

The percentage of errors made when classifying objects to a certain class by the K-Nearest Neighbors classifier is 0.11, the Random Forest algorithm is 0.08, and Decision Tree mistakenly classified a 14 % share of all objects. The Jaccard index, which is an important indicator when distinguishing between the background and the target image, is 0.50 for K-Nearest Neighbors, 0.60 for Random Forest, and 0.41 for Decision Tree.

Table 4

Algorithm quality metrics

Metrics/Algorithms	KNN	RF	DT
Confusion matrix	[8, 4, 4, 32]	[9, 3, 3, 33]	[7, 5, 5, 31]
FPR	0.11	0.08	0.14
FNR	0.33	0.25	0.42
Recall	0.67	0.75	0.70
Precision	0.67	0.75	0.70
Accuracy	0.83	0.88	0.79
F1	0.67	0.75	0.70
Jaccard index	0.50	0.60	0.41
Cohen's kappa	0.63	0.63	0.63
MCC	0.56	0.67	0.44
Specificity	0.89	0.92	0.86
Youden's index	0.56	0.67	0.56
Cross-validation	0.58	0.68	0.45

The Matthews correlation coefficient is useful when working with unbalanced data, in cases where the number of objects of each class is different. The number of images of weed classes in our dataset is the same, so this coefficient has an average. The Youden's index and AUC functions depend only on the total percentage of errors in both classes and do not change with a different distribution of errors between classes, even in the case of an imbalance. According to the table, you can see that the Youden's index for K-Nearest Neighbors and Decision Tree has the same score, and the Random Forest classifier surpasses them by this index.

The quality assessment of the classifier was carried out in order to determine the algorithm with high accuracy of weed detection. Based on numerical indicators, we can conclude that Random Forest is the best suited for classifying objects.

5. 3. 3. Weed detection using YOLO architecture

After collecting the data, it was decided that the YOLOv5 deep learning algorithm would be the optimal solution for creating a model. The main difference between YOLO and other convolutional neural network (CNN) algorithms used for object detection is that it recognizes objects very quickly in real time. The principle of operation of YOLO implies the input of the entire image at once, which passes through the convolutional neural network only once.

The YOLOv5 network architecture consists of three parts: Backbone, Neck, Head (Output). First, after submitting all the data from the image, all the information is first entered into CSP (Cross Stage Partial Network) to extract features [21]. In the end, the Head part is used to output results such as class, grades, location, and size of the object [22].

As shown in Fig. 6, in the Backbone stage, the extraction of informative features is performed using the focus module.

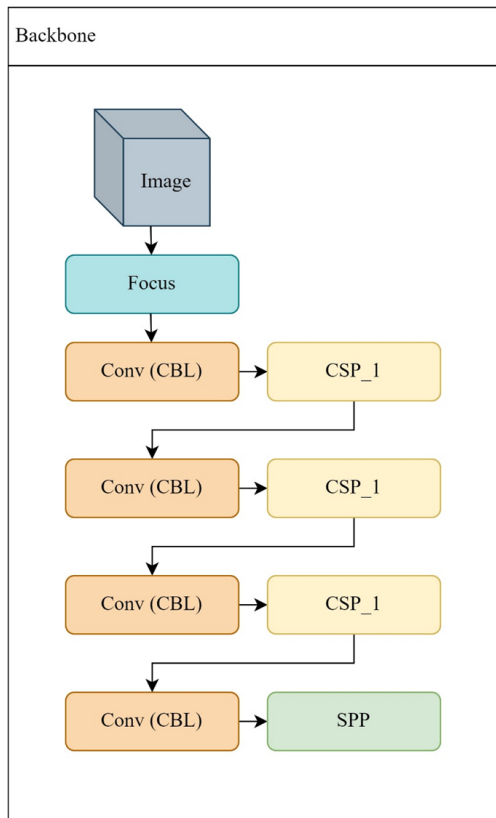


Fig. 6. The Backbone stage

In this part of the architecture, four layers of feature maps of different dimensions are created and combined to reduce data loss. And using CSP, the inference speed increases, and the complexity of the calculation also decreases. The principle of the CSP network is to separate and concatenate images without losing optimal speed and accuracy. The DarkNet-based CSP neural network was used to detect objects. This network divides the base layer into two parts, and using cross-level connectivity, these parts are concatenated. From the architecture, you can see that the last layer is replaced by the SPP (Spatial pyramid pooling) layer, and the CSP network is re-applied to its result to obtain convolution images.

The next stage (Fig. 7) takes the results of the last three layers of the feature map. The results of the CSP of the Backbone stage are transmitted to the Concat function, which is responsible for the operation of combining tensors. And the SPP block that replaced the last CSP layer is transferred to the cross-stage partial network of the second stage. Also at the Neck stage, the Upsample model is used, which performs a sampling operation to obtain an output image of the same size as the input image. After the Concat operation, the CSP network is applied once again to maintain accuracy and reduce the size of the model.

In the last stage, Head performs the final part of the detection (Fig. 8). It applies anchor blocks to objects and generates the final output vectors containing the predicted bounding box, coordinates (center, height, width), prediction confidence score and probability classes [21].

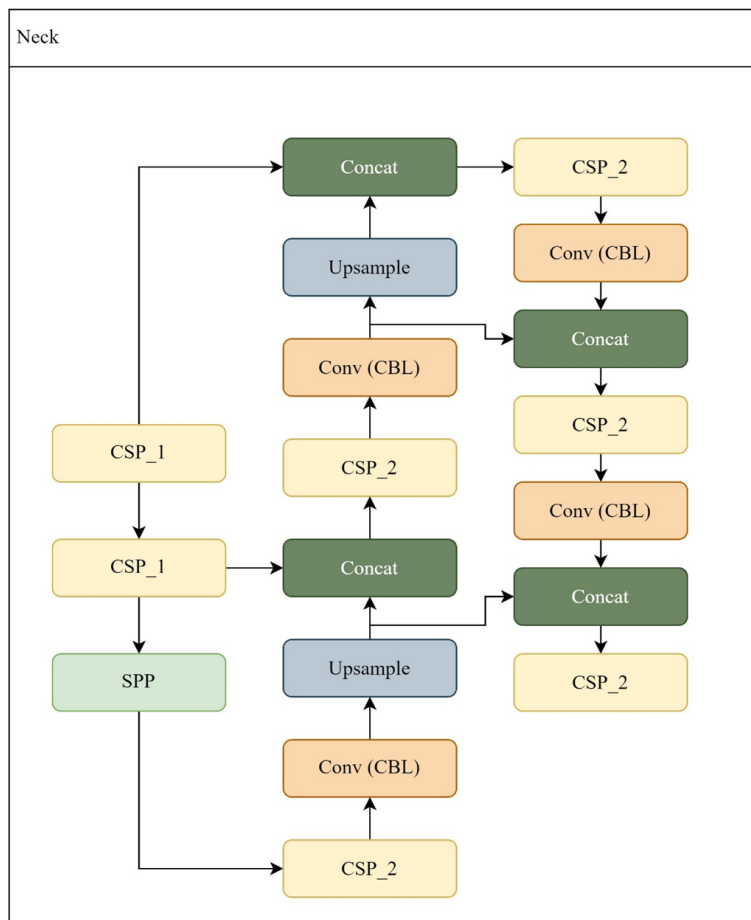


Fig. 7. The Neck stage

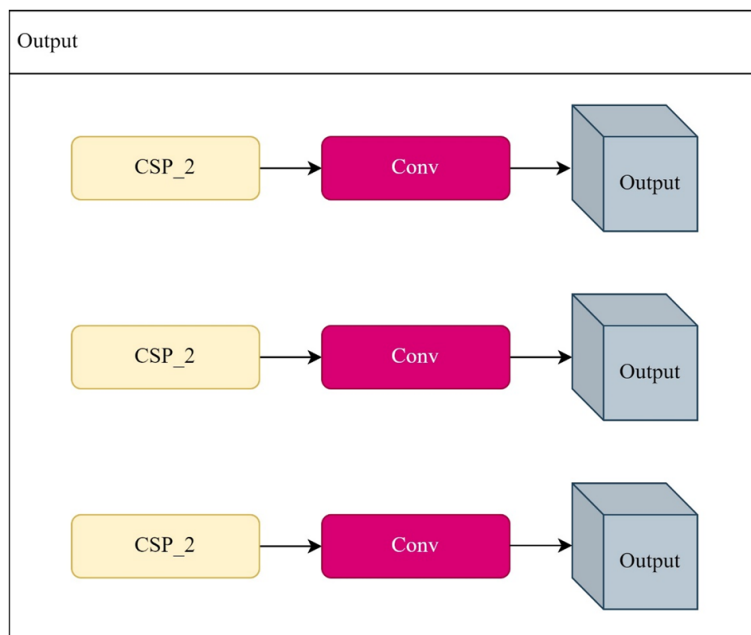


Fig. 8. The Head (Output) stage

As a research result, a contribution was made to improving the YOLOv5 neural network architecture by adding an attention module based on ECA-Net to achieve better performance of the neural network (Fig. 9). The attention module ECA was added between the Neck stage CSP blocks and

the Output stage convolution blocks. ECA-Net is an attention mechanism designed to balance complexity and performance parameters, which has previously been applied to neural network architectures such as ResNet and ImageNet [23].

The attention module is applied to the results of the CSP network of the Neck stage before the final convolution. An important factor in learning the attention module is the dimension reduction, it was added after the second consecutive convolution. As shown in the presented architecture, an image of a weed with a size of 3.024×4.032 pixels was submitted to the input. Using the focus module, more informative features were extracted. After using the CBL module, which includes the functions of convolution, normalization and activation, the dimension decreased to 756×1.008 pixels. As a result of the next convolution, an image of size 378×504 was obtained. Using the third layer reduced the input data to 189×252 pixels. As a result, an image of 95×126 pixels was transferred to the Neck stage.

At the Head stage, the CSP results of the second stage are sequentially transmitted to the added attention module, thus, before the last convolution, the performance of the model increases, and the detection ability also improves. After convolution at the Output stage, three feature maps of different scales were obtained for the output, which are 378×504, 189×252 and 95×126 pixels, respectively, for each layer.

After the result of object detection, the presented architecture is evaluated.

As classes for training, there were 5 types of herbs that were collected in three localities and one unknown type of weed, and class markings were made on all these images. As a result, all these prepared data were submitted to the neural network for training. Thus, the neural network is trained,

and we have obtained a model of object detection, identification of the considered types of herbs using computer vision and machine learning. The result of the neural network has good indicators, and the neural network itself showed a good result (Fig. 10).

Also, the markup was made in such a way that for species with grasses with a long leaf stem, this algorithm recognizes both stems and leaves. This approach is very effective in our case since the ultimate goal is the destruction of this plant species. A good example of this type of plant is the grass Quinoa and the process of its recognition is shown in Fig. 11.

Fig. 12 shows the confusion matrix with normalization for multiclass classification (in our case, 6 classes). The diagonal shows the number of TP combinations for each class: 80 % of all objects of the Ambrosia class, 79 % of all objects of the Bromus class, 82 % of all objects of the Amaranthus class, 74 % of all objects of the Quinoa class, 65 % of all objects of the Bindweed class were classified correctly.

Also, 2 % of all objects of the Amaranthus class were mistakenly predicted as objects of the Quinoa class, another 2 % of objects of the Amaranthus class were found by YOLO as an unknown weed. The remaining 14 % were not assigned to any class by the classifier, so they are false negative solutions. And for the Bindweed class, 1 % of all objects are mistakenly classified as objects of the Bromus class.

Based on these combinations, the main metrics of the classification ability of the algorithm are calculated, such as precision, recall, accuracy, and *F1*-measure. Precision is responsible for the ability to distinguish a given class from all other classes; therefore, it depends only on positive results, that is, precision is the ratio between true-positive results (TP) and all positively classified objects (TP and FP).

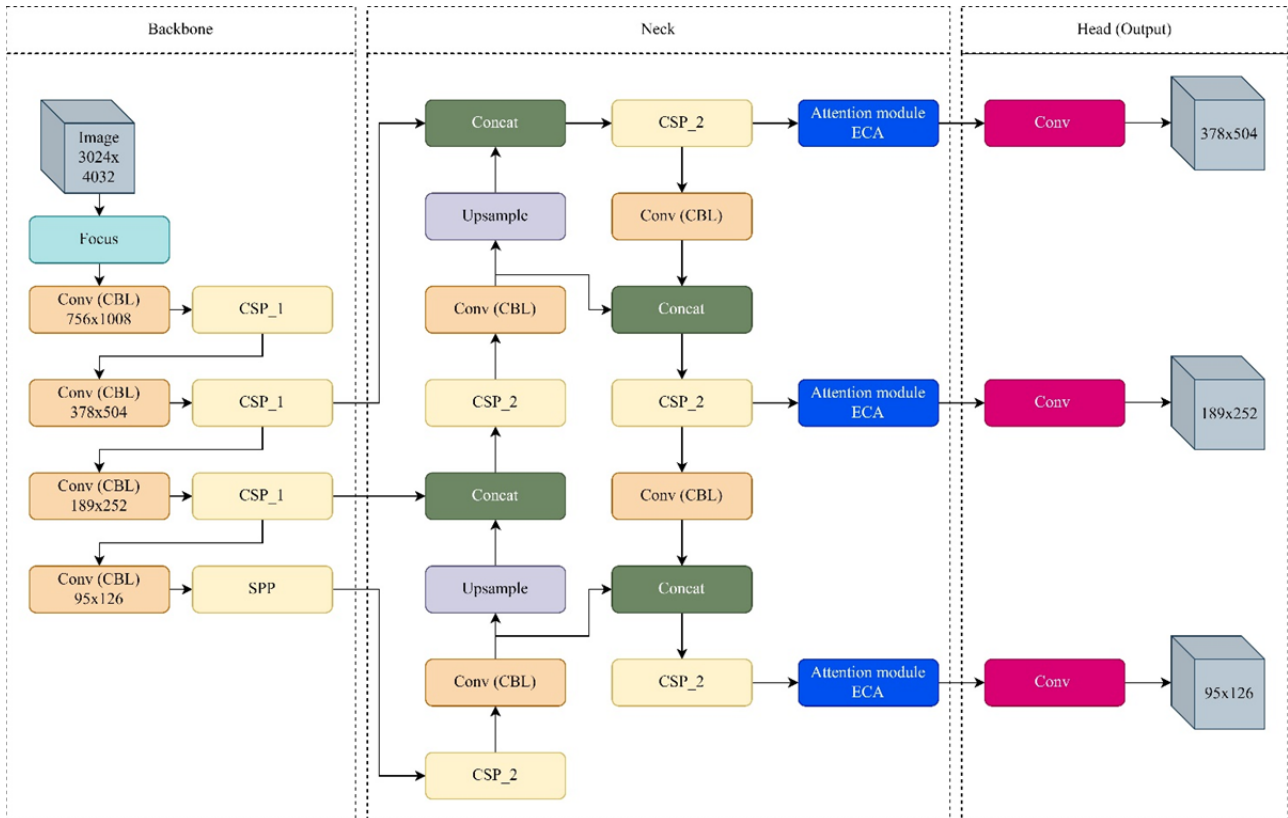


Fig. 9. Improved YOLO architecture



Fig. 10. The result of YOLO’s work: Koster – Bromus; Vyunok’ – Bindweed; Shirica – Amaranthus; Lebeda – Quinoa

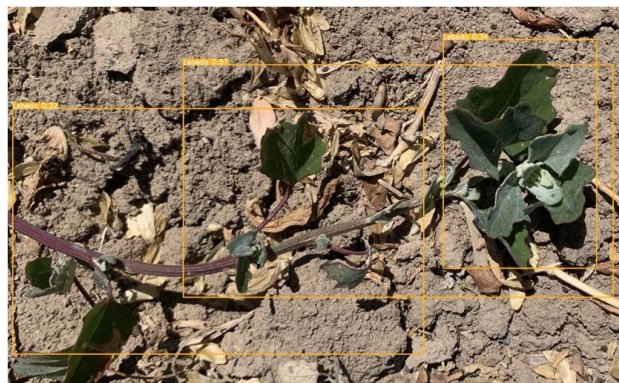


Fig. 11. Recognition of Quinoa grass (Lebeda – Quinoa)

Precision and recall demonstrate a precision-recall curve, which is an important metric when working with imbalanced data. High values under the curve for both metrics show that the classifier makes correct predictions. 0.5 was chosen as the threshold value. Fig. 13 shows that the area under the PR curve for the Ambrosia class is 0.82, for the Bromus class – 0.73, for the Amaranthus class is 0.85, for the Quinoa class – 0.75, and for the Bindweed class – 0.62, the average value of this curve for all classes is 0.78.

In YOLO, one of the significant metrics is confidence, which provides information about the reliability of clas-

sifier forecasts. If you increase the confidence threshold, the value of the precision metric increases, and recall will decrease. Fig. 14 shows that when the reliability threshold was 0.945, all classes achieved perfect accuracy, that is, a value of 1.00.

Fig. 15 shows how the value of the recall estimate decreases with an increase in the reliability threshold. The maximum result with a score of 0.92 was achieved at a threshold of 0.00. If you select 1.00 as the confidence threshold, recall will be equal to 0.00, which is the lowest metric indicator.

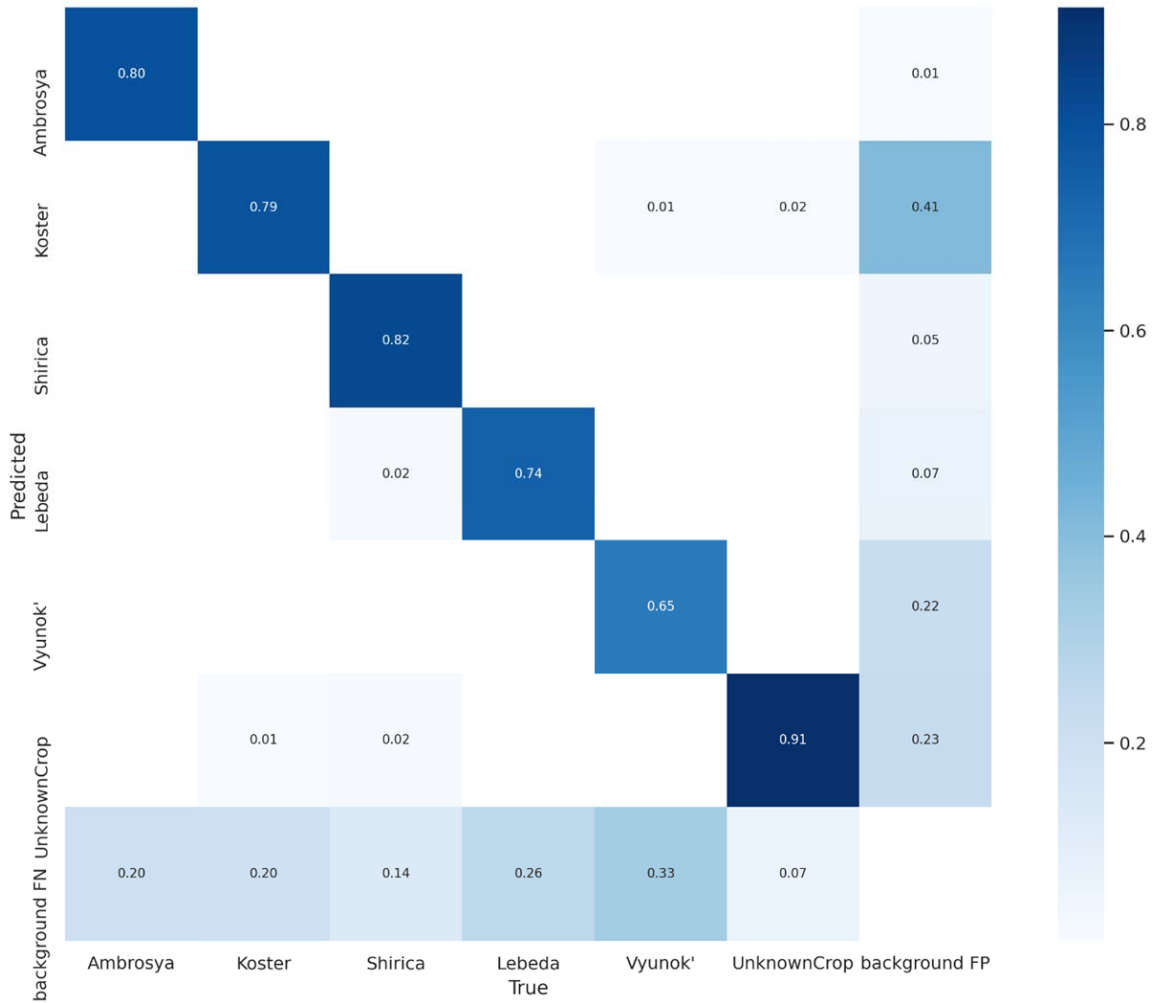


Fig. 12. YOLO confusion matrix with normalization: Ambrosya – Ambrosia; Koster – Bromus; Shirica – Amaranthus; Lebeda – Quinoa; Vyunok’ – Bindweed

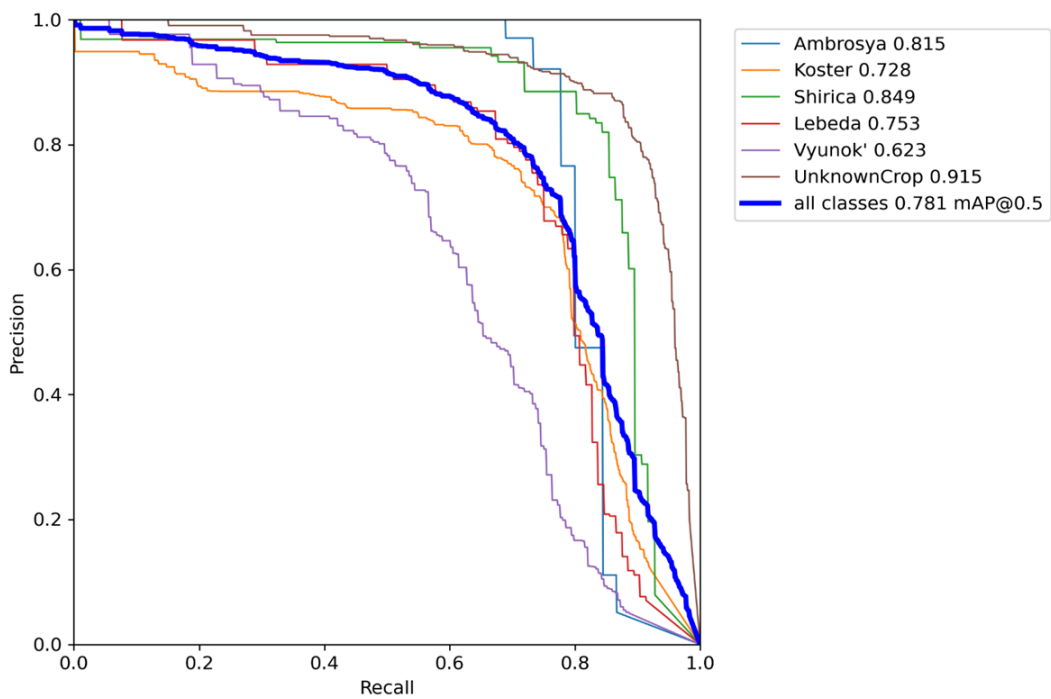


Fig. 13. Calculation of the area under the precision-recall curve for each class: Ambrosya – Ambrosia; Koster – Bromus; Shirica – Amaranthus; Lebeda – Quinoa; Vyunok’ – Bindweed

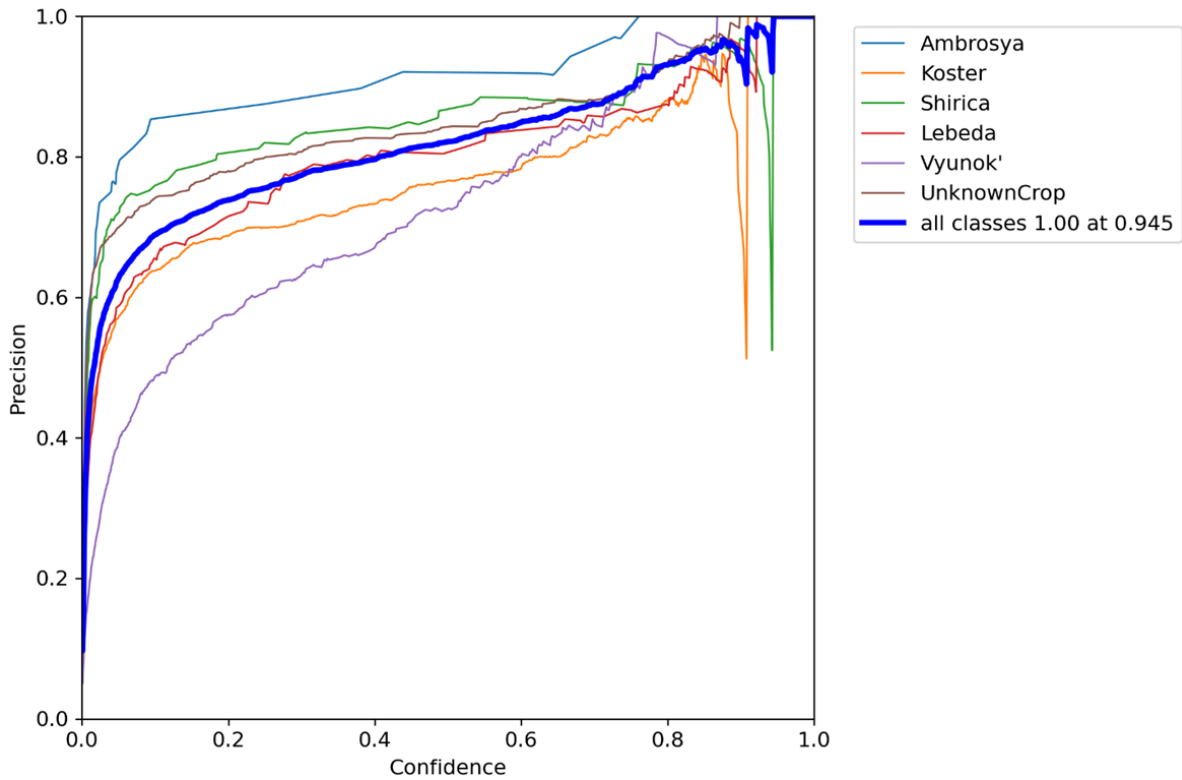


Fig. 14. Precision metrics indicators: Ambrosya – Ambrosia; Koster – Bromus; Shirica – Amaranthus; Lebeda – Quinoa; Vyunok’ – Bindweed

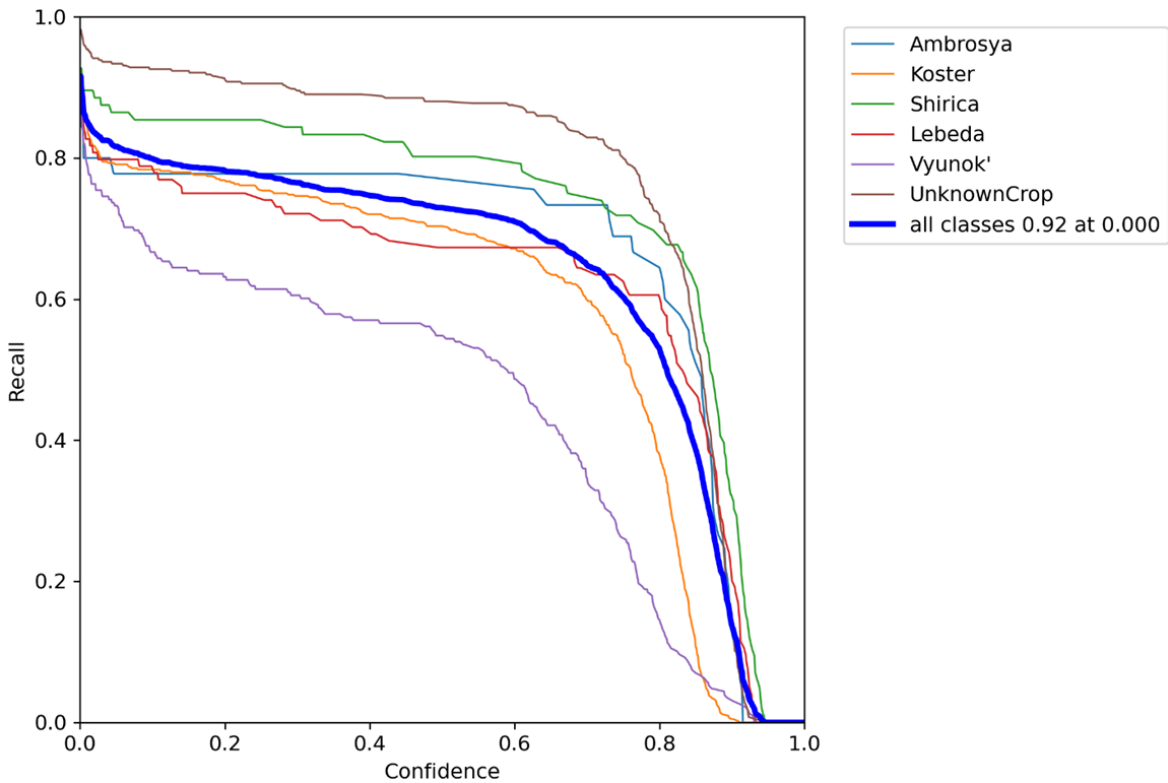


Fig. 15. Recall metrics indicators: Ambrosya – Ambrosia; Koster – Bromus; Shirica – Amaranthus; Lebeda – Quinoa; Vyunok’ – Bindweed

Therefore, to balance precision and recall, you will need the *F1*-measure metric, which combines information about these metrics, and is defined as their average harmonic value. As shown in Fig. 16, the confidence

value from the *F1* curve that balances precision and recall is 0.552. With this confidence value, according to the graph, you can see that the *F1*-measure metric is 0.77.

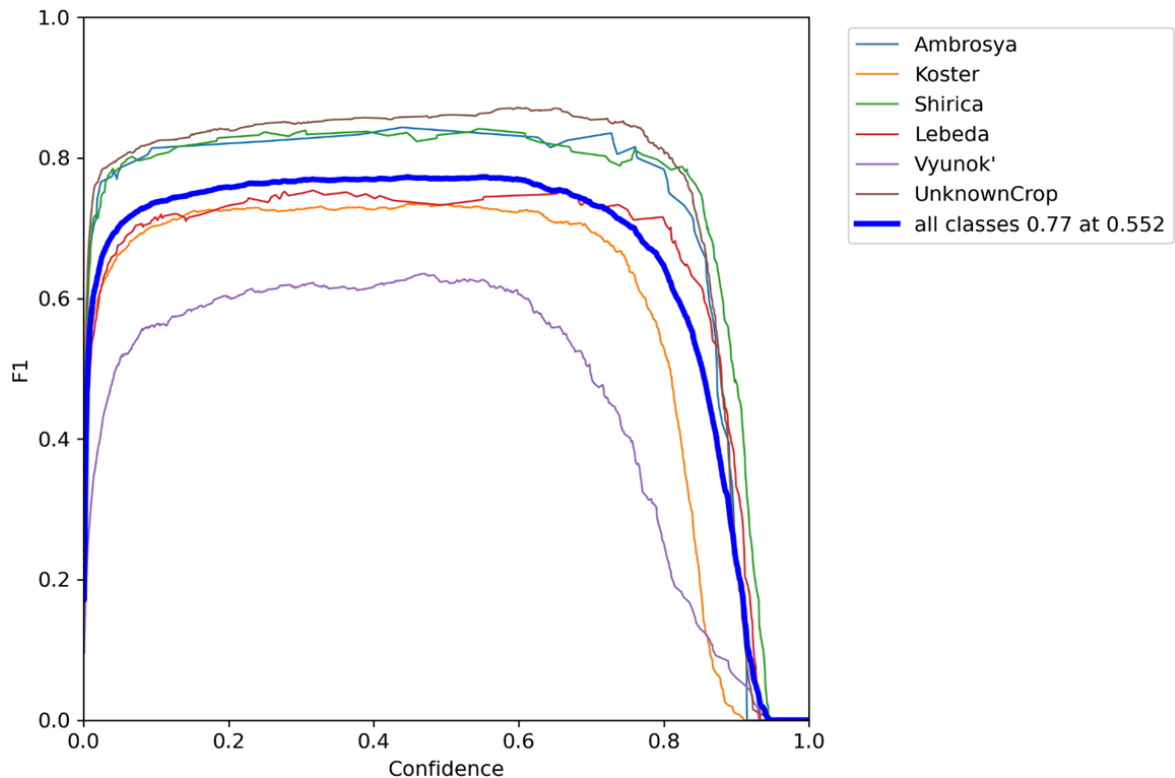


Fig. 16. Graphical interpretation of the *F1* score: Ambrosya – Ambrosia; Koster – Bromus; Shirica – Amaranthus; Lebeda – Quinoa; Vyunok’ – Bindweed

The results of the evaluation confirm that the YOLOv5 algorithm has good accuracy in detecting objects.

6. Results discussion of the weed detection investigation obtained by computer vision algorithms

The development of systems for detecting weeds and agricultural crops in modern conditions of minimizing labor and material costs is an urgent task. A weed identification program was proposed using the YOLOv5 neural network architecture. First, the task was to collect a data set from images of widespread weeds. The trips were organized in the country’s localities, which were transformed into many rural farms. After a visit to the villages of Koram, Saty and Kyrgauyldy of the Almaty region, it was discovered that in the country from late spring plants – Amaranthus, from winter weeds – Bromus, from root weeds – Bindweed, from rod-root plants – Ambrosia are common weeds in the fields of agriculture (Fig. 1). The dataset is collected from more than 5,000 photos of the aforementioned weeds with different levels of illumination, saturation, with different stages of plant leaf growth, only with the leaves of the weed, as well as with the stem, so that the algorithm can recognize weeds in various cases. After data collection, 80 % of all images were distributed to the training sample, and 20 % to the validation sample. The images were segmented by the Otsu’s method to highlight objects and borders on the images. During segmentation, each pixel is assigned a sign to identify pixels that have similar visual characteristics. The removal of interference that occurs during segmentation was carried out by the noise reduction function. After data pre-processing, the images were classified using the KNN, Ran-

dom Forest and Decision Tree algorithms. The sensitivity of the KNN algorithm was 0.89, the Random Forest classifier was 0.92, and the Decision Tree method showed 0.86. And the weed detection accuracy showed results of 0.83, 0.88 and 0.79, respectively. Table 4 presented all the metrics for evaluating machine learning algorithms to demonstrate the determination of the best algorithm in object detection. In order to improve the accuracy results, convolutional neural networks were investigated. The YOLOv5 architecture deserves special attention, which surpasses other neural networks in its lightness and quality of predictions. In size, YOLOv5 is 88 % smaller than its previous version. Also, the speed of this architecture is 140 frames per second, which is 3 times more than others. And the high detection accuracy of YOLOv5 allows using YOLOv5 for real-time object recognition tasks. In our research paper, the idea of changing the architecture was proposed to achieve high performance and reduce complexity. An attention module based on the ECA-Net algorithm has been added.

To check the quality of the proposed architecture, the accuracy indicators were compared with the results of similar work by other researchers. The work in [24] considers the same methods as proposed in our work, which consist in detecting weeds using computer vision. From classical algorithms, in [24] the Support Vector Machine method was chosen, and YOLOv3 was used as the architecture of the neural network. However, the recognition accuracy of our system is relatively higher than in [24]. Their accuracy result was 79 % for the machine learning algorithm, and neural network architectures showed a result of 89 % accuracy. The accuracy of detecting weeds by our system using classical algorithms was 87.5 %, and with the use of YOLOv5 – 92 %. The PR curve values for the Ambrosia class were 0.82, for

Bromus – 0.73, for Amaranthus – 0.85, for Bindweed – 0.75. In the recognition of objects of the Quinoa class, the neural network in many cases mistakenly predicted them as other class weeds, therefore it has an indicator of 0.62. The results were also compared with [25], which have an average accuracy of 65.6 % and 58.7 %. In the data set collected by us, most of the images of weeds are of low resolution, so the ability to identify objects in such conditions is considered one of the advantages of our system.

As a limitation of the proposed method, the small number of classes should be noted. 4 classes were selected for machine learning algorithms training, and 5 classes were selected for neural network training. To eliminate this drawback, further research plans to expand the data set by adding images of other weed species with different growth stages and increase the number of classes, as well as improve the results of detection accuracy by optimizing the proposed methods.

In the future, it is considered to jointly use the proposed tomato recognition system in [26], with this weed detection system, in order to obtain a full-fledged robotic complex capable of distinguishing crop pests from cultivated plants. Earlier, we proposed a mechanism of an agricultural robot with a new tool – a continuous manipulator for harvesting tomatoes [26, 27].

The difficulty of this study lies in the fact that to implement a detection system based on YOLOv5, powerful high-speed computers will be required. Therefore, choosing good performance, speed and computational complexity parameters is important for the implementation of this weed identification program.

7. Conclusions

1. The country's localities were studied to identify weeds that are more common in agricultural fields. To solve this problem, trips were organized to the fields of the village of Koram of the Enbekshikazakh district, Saty of the Raiym-

bek district and Kyrgauyldy of the Karasay district. Weeds such as Bindweed, Amaranthus, Bromus and Ambrosia were selected. The reason for choosing these plants is their widespread distribution in fields, gardens and vegetable gardens, complicating the processes of tillage, crop care and harvesting. More than 1,000 photos were taken for each weed species, thus, the database with images of the aforementioned weeds was formed. The collected dataset can be used by other researchers to conduct further research related to the detection of pests of agricultural lands.

2. In the process of segmentation, the image is cleared of noise to provide images of good quality. According to the results of the assessment, the accuracy of weed detection by the K-Nearest Neighbors, Random Forest and Decision Tree classifiers was 83.3 %, 87.5 %, and 80 %. The average cross-validation index, which is a method of evaluating machine learning models, was 0.68. Due to the fact that the images of weeds of each species differ in resolution and light level, the results of the neural network have corresponding indicators in the intervals of 0.82–0.92 for each class. In general, the YOLOv5 architecture showed a good result.

3. A weed detection system based on the YOLOv5 neural network architecture with the addition of the attention module to improve performance has been developed. Quantitative results obtained on real data demonstrate that the proposed approach can provide good results in classifying low-resolution images of weeds.

Acknowledgments

This work was supported by a grant from the Ministry of Education and Science of the Republic of Kazakhstan within the framework of project No. AP08857573 “Development of intelligent information technologies based on machine vision and pattern recognition with the construction of a mobile robot for farmland maintenance”.

References

1. Fawakherji, M., Youssef, A., Bloisi, D., Pretto, A., Nardi, D. (2019). Crop and Weeds Classification for Precision Agriculture Using Context-Independent Pixel-Wise Segmentation. 2019 Third IEEE International Conference on Robotic Computing (IRC). doi: <https://doi.org/10.1109/irc.2019.00029>
2. Zhang, S., Huang, W., Wang, Z. (2021). Combing modified Grabcut, K-means clustering and sparse representation classification for weed recognition in wheat field. *Neurocomputing*, 452, 665–674. doi: <https://doi.org/10.1016/j.neucom.2020.06.140>
3. Potena, C., Nardi, D., Pretto, A. (2017). Fast and Accurate Crop and Weed Identification with Summarized Train Sets for Precision Agriculture. *Advances in Intelligent Systems and Computing*, 105–121. doi: https://doi.org/10.1007/978-3-319-48036-7_9
4. Di Cicco, M., Potena, C., Grisetti, G., Pretto, A. (2017). Automatic model based dataset generation for fast and accurate crop and weeds detection. 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). doi: <https://doi.org/10.1109/iros.2017.8206408>
5. Lavania, S., Matey, P. S. (2015). Novel Method for Weed Classification in Maize Field Using Otsu and PCA Implementation. 2015 IEEE International Conference on Computational Intelligence & Communication Technology. doi: <https://doi.org/10.1109/cict.2015.71>
6. Wu, Z., Chen, Y., Zhao, B., Kang, X., Ding, Y. (2021). Review of Weed Detection Methods Based on Computer Vision. *Sensors*, 21 (11), 3647. doi: <https://doi.org/10.3390/s21113647>
7. Selvi, C. T., Sankara Subramanian, R. S., Ramachandran, R. (2021). Weed Detection in Agricultural fields using Deep Learning Process. 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS). doi: <https://doi.org/10.1109/icaccs51430.2021.9441683>
8. Boyina, L., Sandhya, G., Vasavi, S., Koneru, L., Koushik, V. (2021). Weed Detection in Broad Leaves using Invariant U-Net Model. 2021 International Conference on Communication, Control and Information Sciences (ICCISc). doi: <https://doi.org/10.1109/iccisc52257.2021.9485001>

9. Tang, J., Wang, D., Zhang, Z., He, L., Xin, J., Xu, Y. (2017). Weed identification based on K-means feature learning combined with convolutional neural network. *Computers and Electronics in Agriculture*, 135, 63–70. doi: <https://doi.org/10.1016/j.compag.2017.01.001>
10. Luo, T., Zhao, J., Gu, Y., Zhang, S., Qiao, X., Tian, W., Han, Y. (2021). Classification of weed seeds based on visual images and deep learning. *Information Processing in Agriculture*. doi: <https://doi.org/10.1016/j.inpa.2021.10.002>
11. Jiang, H., Zhang, C., Zhang, Z., Mao, W., Wang, D., Wang, D. (2020). Detection Method of Corn Weed Based on Mask R-CNN. *Nongye Jixie Xuebao/Transactions of the Chinese Society for Agricultural Machinery*, 51 (6), 220–228. doi: <https://doi.org/10.6041/j.issn.1000-1298.2020.06.023>
12. Dos Santos Ferreira, A., Matte Freitas, D., Gon alves da Silva, G., Pistori, H., Theophilo Folhes, M. (2017). Weed detection in soybean crops using ConvNets. *Computers and Electronics in Agriculture*, 143, 314–324. doi: <https://doi.org/10.1016/j.compag.2017.10.027>
13. Wang, A., Zhang, W., Wei, X. (2019). A review on weed detection using ground-based machine vision and image processing techniques. *Computers and Electronics in Agriculture*, 158, 226–240. doi: <https://doi.org/10.1016/j.compag.2019.02.005>
14. Ma, X., Deng, X., Qi, L., Jiang, Y., Li, H., Wang, Y., Xing, X. (2019). Fully convolutional network for rice seedling and weed image segmentation at the seedling stage in paddy fields. *PLOS ONE*, 14 (4), e0215676. doi: <https://doi.org/10.1371/journal.pone.0215676>
15. Asad, M. H., Bais, A. (2020). Weed detection in canola fields using maximum likelihood classification and deep convolutional neural network. *Information Processing in Agriculture*, 7 (4), 535–545. doi: <https://doi.org/10.1016/j.inpa.2019.12.002>
16. Bosilj, P., Duckett, T., Cielniak, G. (2018). Analysis of Morphology-Based Features for Classification of Crop and Weeds in Precision Agriculture. *IEEE Robotics and Automation Letters*, 3 (4), 2950–2956. doi: <https://doi.org/10.1109/lra.2018.2848305>
17. Sarvini, T., Sneha, T., Sukanya Gowthami, G. S., Sushmitha, S., Kumaraswamy, R. (2019). Performance Comparison of Weed Detection Algorithms. 2019 International Conference on Communication and Signal Processing (ICCSP). doi: <https://doi.org/10.1109/iccsp.2019.8698094>
18. Xiao, L., Ouyang, H., Fan, C., Umer, T., Poonia, R. C., Wan, S. (2020). Gesture image segmentation with Otsu's method based on noise adaptive angle threshold. *Multimedia Tools and Applications*, 79 (47-48), 35619–35640. doi: <https://doi.org/10.1007/s11042-019-08544-7>
19. Belgiu, M., Drăguț, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 24–31. doi: <https://doi.org/10.1016/j.isprsjprs.2016.01.011>
20. Starovoitov, V. V., Golub, Y. I. (2020). Comparative study of quality estimation of binary classification. *Informatics*, 17 (1), 87–101. doi: <https://doi.org/10.37661/1816-0301-2020-17-1-87-101>
21. Thuan, D. (2021). Evolution of YOLO algorithm and YOLOv5: The state-of-the-art object detection algorithm. *Information Technology Oulu University of Applied Sciences*, 61. Available at: https://www.theseus.fi/bitstream/handle/10024/452552/Do_Thuan.pdf?isAllowed=y&sequence=2
22. Xu, R., Lin, H., Lu, K., Cao, L., Liu, Y. (2021). A Forest Fire Detection System Based on Ensemble Learning. *Forests*, 12 (2), 217. doi: <https://doi.org/10.3390/f12020217>
23. Wang, Q., Wu, B., Zhu, P., Li, P., Zuo, W., Hu, Q. (2020). ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). doi: <https://doi.org/10.1109/cvpr42600.2020.01155>
24. Osorio, K., Puerto, A., Pedraza, C., Jamaica, D., Rodriguez, L. (2020). A Deep Learning Approach for Weed Detection in Lettuce Crops Using Multispectral Images. *AgriEngineering*, 2 (3), 471–488. doi: <https://doi.org/10.3390/agriengineering2030032>
25. Ahmad, T., Ma, Y., Yahya, M., Ahmad, B., Nazir, S., Haq, A. ul. (2020). Object Detection through Modified YOLO Neural Network. *Scientific Programming*, 2020, 1–10. doi: <https://doi.org/10.1155/2020/8403262>
26. Yeshmukhametov, A. N., Koganezawa, K., Buribayev, Z., Amirgaliyev, Y., Yamamoto, Y. (2020). Study on multi-section continuum robot wire-tension feedback control and load manipulability. *Industrial Robot: The International Journal of Robotics Research and Application*, 47 (6), 837–845. doi: <https://doi.org/10.1108/ir-03-2020-0054>
27. Yeshmukhametov, A., Khaleel, L. A., Koganezawa, K., Yamamoto, Y., Amirgaliyev, Y., Buribayev, Z. (2020). Designing of CNC Based Agricultural Robot with a Novel Tomato Harvesting Continuum Manipulator Tool. *International Journal of Mechanical Engineering and Robotics Research*, 9 (6), 876–881. doi: <https://doi.org/10.18178/ijmerr.9.6.876-881>