ZH. A. BURIBAYEV², ZH. AMIRGALIYEVA¹, S. K. JOLDASBAYEV^{1,2*},
M. S. ZHASSUZAK^{1,2}, A. S. TUREGALI², D. D. DAULETIYA²

¹*Institute of information and computational technologies,*

²*Al-Farabi Kazakh National University,*

Almaty, Kazakhstan

TOMATO MATURITY RECOGNITION USING YOLOv5 MACHINE LEARNING

The implementation of robotic systems and digitalization in agriculture are important tasks today. In this paper, the possibility of using pattern recognition and machine learning methods as a computer model of an agricultural robot for harvesting is considered. The grading of tomato fruits can be classified based on their ripeness according to their life cycles, which can be identified by their color: green in the growing stage, yellow in the pre-ripening stage, and red when ripe. Conventional skill-based methods cannot meet the exact selection criteria for modern production management in the agricultural sector as they are time-consuming and of low accuracy. Automatic feature extraction behavior using machine learning is most effective in image classification and recognition tasks. Thus, the article presents the results of a study on the recognition of ripe tomato fruits by a robotic system, carried out within the framework of the grant project of the Ministry of Education and Science of the Republic of Kazakhstan AP08857573 and implemented classical algorithms based on the HSV color model and color segmentation using the k-means algorithm as comparative algorithms and based on machine learning, a universal intelligent tomato classification system is proposed for practical use using Yolo 5.

This study aims to provide an inexpensive solution with the best performance and accuracy for assessing tomato ripeness. The results are collected in terms of accuracy, loss curves and confusion matrix. The results showed that the proposed model outperforms other machine learning (ML) methods used by researchers for tomato classification problems, providing 99% accuracy.

Keywords: machine learning, tomato recognition, robotic systems, agricultural robot, HSV segmentation, YOLO.

Introduction. Over the past decade, machine learning (ML) has become widely known for the implementation of a wide variety of applied problems. One of the topical directions today is the introduction of intelligent machines in the agro-industrial sector: care and collection of culture. In this work, we use ML to recognize ripe tomato fruits by comparing classical recognition algorithms and ML.

The development of robotic solutions for agriculture requires advanced sensing capabilities that can reliably operate at any stage of the harvest. For example, to automate the process of harvesting tomatoes in greenhouses, the visual perception system must detect a tomato at any stage of its life cycle (from a flower to a ripe tomato). Modern tomato visual detection equipment is mainly focused on ripe tomatoes that have a distinctive color from the background. This document supplements the annotated green and reddish tomato visual dataset. Such a dataset is rare and not available for research purposes. This will enable the further development of advanced artificial intelligence to visually locate tomatoes in situ and in real time needed to develop cleaning robots. With this dataset in

* E-mail корреспондирующего автора: zholdas_87@mail.ru, serykjoldasbaev@mail.ru

mind, five deep learning models were selected, trained and tested to detect green and reddish tomatoes grown in greenhouses. Taking into account the technical characteristics of our robotic platform, the work of experienced specialists was considered, which are briefly described below.

Literature Review. In the course of the study, the work of foreign scientists in this area was considered as [1], a segmentation algorithm was developed to control the robot arm for picking ripe tomatoes and the recognition algorithm adapts to the lighting conditions of the greenhouse, where 110 color images of tomatoes were used in greenhouse lighting, the algorithm works in two steps: (1) by removing the background in RGB color space, followed by extraction of a ripe tomato using a combination of RGB, HSI and YIQ spaces; and (2) localization of a ripe tomato using morphological features of the image.

One of the vulnerabilities on the way to reliable product recognition in agriculture is to reduce the impact of the main obstacles: lighting and overlap. In order to recognize a tomato in the crown of a plant using an inexpensive camera, a robust tomato recognition algorithm based on several characteristic images and image fusion was studied in [2]. Two new characteristic images, an a^* -component image and an I -component image, were extracted from the $L^*a^*b^*$ color space and the luminance, in-phase, quadrature (YIQ) color space. Also in the work, a wavelet transform was adopted to merge two characteristic images at the pixel level, which combined information about the features of the two original images. To separate the target tomato from the background, an adaptive thresholding algorithm was used to obtain the optimal thresholding. The final segmentation result was processed with a morphology operation to reduce a small amount of noise. This indicates that we could use the proposed method for robotic harvesting of tomatoes in an uncontrolled environment. Studies conducted in [3], where data from 600 tomatoes at six stages of maturity were taken and used for the test, absorption and reduced scatter coefficients for tomato fruits were extracted using spatially resolved spectroscopy, partial least squares discriminant analysis (PLSDA) models were built to assess the maturity of tomatoes. The authors have shown that optical absorption and scattering spectra can effectively classify the ripeness stages of tomatoes and have provided new means of non-destructive detection in agricultural products. In [4], the same authors propose a different approach, where partial least squares discriminant analysis PLSDA and support vector machine discriminant analysis SVMDA. Also, in [5], a study of measuring the optical properties of tomato fruit in the wavelength range using the method of spatially resolved diffuse reflection to estimate density, soluble solids content (SSC) and pH. In reviewing this study, we learned that the SRS method, along with the absorption and scattering coefficients, has the potential for non-destructive measurement of quality characteristics, especially the hardness of tomato fruit, which we could then use for the picker arm.

In [6] discusses the detection of ripe and inverted tomato fruits using computer vision and image processing techniques using Raspberry Pi and Pi Camera, Raspbian software using python3. Ripeness is determined using OpenCV and HSV. Researchers are going to connect this equipment to Turtlebot, to a robot that will move around the field to find ripe plants. Since the development ideas are similar for us, the results of this study on the implementation of the development in a mobile robot are important for us.

In [7], the proposed system performs the classification of tomatoes in three stages. Due to the complexity of solving a nonlinear problem, the proposed system is implemented as a cascade of two machine support vector classifiers. The performance of this system is assessed using metrics for accuracy, specificity, sensitivity, and accuracy. Experimental results and comparative analysis of similar methods testify to the effectiveness of the proposed system in comparison with existing systems for sorting and identifying tomatoes. The results obtained at each of the three stages of classification, i.e. tomato/non-tomato, good/defective and the type of defect in the event of a defect, are transmitted to the microcontroller to turn on the corresponding motor, so that the given fruit can be classified and harvested in the corresponding basket.

The work [8] researchers proposed a classification system using a pair of binary SVM classifiers to sort tomato ripeness and defects. The first SVM classifier is unique regardless of the ripeness level for the separation of tomato species. The ripeness level of tomatoes is determined based on the average green component. Thereafter, defective and healthy tomatoes are distinguished using a superpixel segmentation operation. A second SVM classifier is then used to detect black spot defects and ulcers separately in ripe and immature fruits. Although the same feature vectors are used to detect defects regardless of ripeness, the reference image used to create feature afterimages is different for ripe and unripe tomatoes.

Researchers in [9] proposed a fuzzy Mask R-CNN model to automatically determine the ripeness levels of cherry tomatoes. For automatic image annotation, a fuzzy K-means model was used to store the spatial information of various elements of the foreground and background of the image. The Hough transform method was applied to determine the position of the geometric edges of the tomatoes. Each data point in image space has been annotated into a JavaScript object notation file. The annotated images were trained with Mask R-CNN to accurately identify each tomato. To prevent tomato shedding before harvest, the hue-saturation-value color model and fuzzy inference rules were used to predict tomato ripeness. The trigonometric function with Euclidean distance was calculated from the beginning of the calyx and stem to the bottom of the tomato to get the position of the stem head and cut the fruit in time.

In [10], a machine vision system for an apple picking robot was developed based on the YOLOv3 and YOLOv5 algorithms with special pre and post processing. We took this publication as a basis for the development and tested these solutions on tomatoes only using YOLOv5.

In work [11] proposes a method for determining the ripeness of tomatoes by using several streams of a convolutional neural network (ConvNet) and their methodology of stochastic combination of solutions (SDF). The authors named the general pipeline SDF-ConvNets, which can correctly determine the ripeness of tomatoes by performing the sequential steps: (1) determining the initial ripeness of tomatoes for multi-view images based on a deep learning model, and (2) stochastically combining these initial results to obtain the final classification result. No follow-up study was undertaken in this paper to develop an integrated system that can determine the appropriate harvest time and monitor the growing state of the crop by recognizing and assessing the ripening stage in real time by observing the tomatoes before harvest.

Experimental

Segmentation of tomato according to the HSV model. Two main filters were used to segment the tomato - the first filter determines the shape of the object and the second filter to determine the color. In this study, we selected the HSV color model for the color segmentation of tomatoes (Figure 1).

A comparative analysis is performed using a color model (HSV), which changes the hue, saturation and intensity of the pixel of the image in question. After determining the dominant property based on these three parameters to segment the extracted image. This color space model is very close to human perception when compared to other models such as RGB or CMYK. The analysis shows that with the help of this segmentation it is possible to obtain a more accurate identification of objects in comparison with other algorithms. The HSV model can be visualized as a cylinder, where the intensity will be located on the central vertical axis, and the values of the shadow parameters will be determined by the tilt angle in the range of 0-360°. For red, green and blue, the values correspond to 0, 120°, 240°, and the saturation parameter is responsible for determining the depth or purity of a color, which is measured as the radial distance with a value from center to center. outer surface, calculated from 0 to 1. If written as a percentage, 0 will be black and 100% will be white. By decreasing the saturation, each color could be converted to a shade of gray. For low saturation values, the color may correspond to gray.

Figure 1 – HSV color model for color segmentation of tomatoes.

The advantages of the HSV model over RGB are in Figure 2 that the hue of a color (hue) is specified by only one coordinate - hue, which makes it easier to filter out only the necessary colors from the picture. In addition, using the saturation and value coordinates, you can minimize or completely eliminate the effect of low light, darkening or shadows in the working area of the camera. This significantly increases the efficiency of determining the ripeness of a tomato by its color.

Figure 2 – Image conversions from BGR to HSV color space

Then we need to define the lower and upper range of values for a particular color, which in turn consists of three values (hue, saturation, intensity). After that we create a mask for the image by setting the threshold and adding the required range to the value.

Figure 3 – Histogram of an image in RGB and HSV format

In Figure 3, you can clearly see the difference in the intensity distribution. In the histogram, the pixel values are on the abscissa, and the ordinate corresponds to the number of pixels. We can see that in the first picture there were quite a lot of intensity values in the area of pixels from 0 to 100, and after converting the intensity values from 0 to 50, very little. The maximum value can be seen in section 50. Also, if in the first image the value from 50 has decreased, then in the second, starting from 100 to 250, it leads the same way.

Next, you will need to make morphological transformations using the OpenCV library. We need this operation to remove white noise, separate and establish the boundary of two related objects. Morphological operations are performed on binary images. To do this, we need to give our original image and a threefold image at the input, which is called a structural element. A structural element or core does the operation we need.

There are several morphological operators, and in order to choose the most suitable one it is necessary to consider all of them.

Figure 4 – Morphological transformation of the original image

The first image is the result of the Dilation operation, which enlarges the white area in the original image. Also, this operation increases the size of the foreground object. It is also useful for connecting separated parts of an object. In our case, we see that indeed the size of the foreground objects has increased significantly. However, this is unnecessary for us, since our main task is to divide the boundaries of objects in such a way as to segment them separately. The second image is Erosion. This operation blurs the boundaries of the foreground object, which is the exact opposite of the first. The main advantage of which is the separation of the two areas, which is what we need. However, it is hard not to notice that this operation, although it separates two objects, it also reduces their size. This is not what we want at all. The third image is Opening, the principle of which is the same as for erosion, but after that it increases the white area in the image in question. A feature of this operation is the removal of noise. The fourth is Closing, which is the reverse of closing. The main advantage is the removal of small holes inside foreground objects and unnecessary pixels. Fifth Image The Morphological Gradient is the result of the difference between Dilation and Erosion. In the original image, you can see that only the border of our object has been received. The last Tophat image is the result of the work between the original image and the Closing operation. In the figure, we see that this operation does not quite give a good result.

Having received all the results, it was decided to choose the Closing operation for our segmentation. After we can segment objects by color, a function was written for each color we needed (red and green). After segmentation, we can define their outline and use the boundingRect function to constrain them to a rectangle.

It is clear that it is impossible to simultaneously determine two colors on the input image, since during the detection of one color everything else becomes a black background and goes to the background.

Therefore, to solve this problem, two images were created, one which consists entirely of black pixels Fig5a, and the other is a copy of the input and both of them are identical in size with the input Fig5b. We don't need the first image, but it doesn't hurt for comparison and visualization. After that, while detecting each color, we write down the coordinates of the bounding rectangle and also draw them on the two images we just created (Figure 5).

Figure 4 – Identifying objects by color and creating a bounding box

Object detection with YOLO. Object detection using CNN convolutional neural networks is one of the urgent tasks. When reviewing the literature on this problem, many researchers used the COCO dataset to compare different convolutional neural network models. Researchers and developers of the YOLO neural network also used it to compare with other models of convolutional non-ironic networks [2].

To train the YOLO neural network, we collected a dataset of 800 photographs that were taken in the Almaty Tomato Greenhouse, Kazakhstan. And also the photos were taken under different lighting conditions. When marking the data, the tomatoes were divided into 3 cash registers: 0-red, 1-green, 2-yellow. The training itself was carried out in 70 eras and the following results were obtained:

This image shows the PR curve for each class separately and one for all classes. The areas under these curves were calculated for each class in order to belong to a specific class and to evaluate the algorithm. The lowest value is shown by the second class, that is, the class for the yellow tomato, and the highest value for the red tomato. This is due to the fact that the set contains a relatively small amount of data for this class. Also, the dark blue curve shows the arithmetic mean of 0.83. For all classes of measuring the accuracy of the object detector in a particular dataset, the value is 0.5. This value is standard for YOLO5 and the AP of all images of each category is calculated, and then all categories are averaged, namely in mAP.

Figure 5 – Yolo learning outcomes

Results and Discussion.

Figure 6 – The results of the recognition of the algorithm, where different images with different illumination, from different angles and with different levels of background noise are selected.

Training and testing of the model is implemented on a computer with the following characteristics: Intel® Core™ i5-9400F CPU @ 2.90GHz, NVIDIA Corporation GP104 [GeForce GTX 1070 Ti], 24 GiB.

Conclusion. The color representation model has significant significant flaws, and very often these deficiencies are serious problems in the field of computer vision. For example, the strong influence of shadows and low light can significantly reduce the color recognition performance of tomatoes. That is, with the slightest change in illumination, some pixels of the image will not be recognized as part of the contour, or those pixels that are not part of it will also be associated with the contour. It is difficult to establish the desired color tone, as the tone determines the degree of color difference. Based on these two disadvantages, it was

decided to use the Yolo neural network. Based on the results of the research, the proposed solution recognizes ripe fruits in 99% of cases under different lighting conditions.

The results obtained are planned to be implemented in the future on the developed robotic manipulator in greenhouse conditions.

Acknowledgments. This work was carried out and funded under the grant project of the Ministry of Education and Science of the Republic of Kazakhstan AP08857573.

REFERENCES

- 1 A. Arefi, A.M. Motlagh, K. Mollazade, R.F. Teimourlou, Recognition and localization of ripen tomato based on machine vision // Australian Journal of Crop Science. - Vol. 5, Issue 10, - 2011, p. 1144-1149 ISSN: 18352693.
- 2 Y. Zhao, L. Gong, Y. Huang and C. Liu Robust tomato recognition for robotic harvesting using feature images fusion // Sensors (Switzerland). - Vol. 16, Issue 2, - 29 January 2016, №173 ISSN: 14248220, DOI: 10.3390/s16020173.
- 3 Y.P. Huang, D.Z. Wang, H.Y. Zhou, Y.T Yang, K.J. Chen Ripeness assessment of tomato fruit by optical absorption and scattering coefficient spectra // Spectroscopy and Spectral Analysis Volume 40, Issue 11, 1 November 2020, P. 3556-3561 DOI: 10.3964/j.issn.1000-0593(2020)11-3556-06.
- 4 Y.P. Huang, R. Lu, C. Qi, K.J. Chen Tomato maturity classification based on spatially resolved spectra // Spectroscopy and Spectral Analysis Vol. 38, Issue 7, 1 July 2018, P. 2183-2188 DOI: 10.3964/j.issn.1000-0593(2018)07-2183-06.
- 5 Y. Huang, R. Lu, D. Hu, K. Chen Quality assessment of tomato fruit by optical absorption and scattering properties // Postharvest Biology and Technology Volume 143, September 2018, P. 78-85, DOI: 10.1016/j.postharvbio.2018.04.016.
- 6 Z. Al-Mashhadani, B. Chandrasekaran Autonomous Ripeness Detection Using Image Processing for an Agricultural Robotic System // 2020 11th IEEE Annual Ubiquitous Computing, Electronics and Mobile Communication Conference, UEMCON 2020 №9298168, p. 0743-0748 DOI: 10.1109/UEMCON51285.2020.9298168.
- 7 K. S. Dhakshina, S. Esakkirajan, S. Bama, B. Keerthiveena A microcontroller based machine vision approach for tomato grading and sorting using SVM classifier // Microprocessors and Microsystems Volume 76, July 2020, ISSN: 01419331 DOI: 10.1016/j.micpro.2020.103090.
- 8 K. S. Dhakshina, S. Esakkirajan, S. Bama A non destructive tomato grading and sorting mechanism using binary SVM pair based classifier for sustainable agricultural practices // Journal of Green Engineering Volume 10, Issue 7, July 2020, P. 3637-3654 ISSN: 19044720.
- 9 Y.P. Huang, T.H. Wang, H. Basanta, Using Fuzzy Mask R-CNN Model to Automatically Identify Tomato Ripeness // IEEE DOI: 10.1109/ACCESS.2020.3038184.
- 10 Kuznetsova, A., Maleva, T., Soloviev, V. Detecting Apples in Orchards Using YOLOv3 and YOLOv5 in General and Close-Up Images Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) Volume 12557 LNCS, 2020, Pages 233-243 17th International Symposium on Neural Networks, ISNN 2020; Cairo; Egypt; 4 December 2020 до 6 December 2020, ISBN: 978-303064220-4 doi: 10.1007/978-3-030-64221-1_20.
- 11 K. Ko, I. Jang, J.H. Choi, J.H. Lim and D.U. Lee Stochastic decision fusion of convolutional neural networks for tomato ripeness detection in agricultural sorting systems // Sensors (Switzerland) Vol. 21, Issue 3, 1 February 2021, №917, Pages 1-14, ISSN: 14248220 DOI: 10.3390/s21030917.
- 12 Sandro Augusto Magalhães, Luís Castro, Germano Moreira, Filipe Neves dos Santos, Mário Cunha, Jorge Dias, António Paulo Moreira Evaluating the single-shot multibox detector and yolo

deep learning models for the detection of tomatoes in a greenhouse // Sensors 2021, 21(10), 3569; <https://doi.org/10.3390/s21103569>.

**Ж.А. БУРИБАЕВ², Ж.АМИРГАЛИЕВА¹, С.К. ДЖОЛДАСБАЕВ^{1,2},
М.С. ЖАСҰЗАҚ^{1,2}, А.С. ТУРЕГАЛИ², Д. ДАУЛЕТІЯ²**

¹Ақпараттық және есептеу технологиялары институты,

²Әл-Фараби атындағы Қазақ Ұлттық Университеті,

Алматы, Қазақстан

YOLO 5 МАШИНАЛЫҚ ОҚЫТУДЫ ҚОЛДАНЫП ҚЫЗАНАҚ ТҮЙІНДЕРІНІҢ ЖЕТІЛУ ДЕҢГЕЙІН ТАҢУ

Ауыл шаруашылығында роботтандырылған жүйелерді енгізу маңызды міндеттерің бірі. Бұл жұмыста егін жинау үшін ауылшаруашылық роботының компьютерлік үлгісі ретінде үлгіні таңу мен машиналық оқыту әдістерін қолдану мүмкіндігі қарастырылған. Қызанақ жемістерін бағалау олардың өмірлік циклдеріне сәйкес жектеудің тиімді әдісі – оны түсі бойынша анықтау: өсіп келе жатқан кезеңде жасыл, орта кезеңінде сары және піскен кезде қызыл түске боялады. Дағдылардың әдеттегі әдістері ауылшаруашылық секторындағы заманауи өндірісті басқарудың нақты таңдау критерийлеріне сәйкес келе алмайды, өйткені олар көп уақытты және төмен дәлдікті береді. Машинаны оқытуды қолдана отырып, функцияларды автоматты түрде алу әрекеті кескіндерді жектеуде және таңу тапсырмаларында тиімді. Осылайша, мақалада Қазақстан Республикасы Білім және ғылым министрлігінің AP08857573 гранттық жобасы аясында жүргізілген және классикалық алгоритмдерге негізделген роботты жүйе ретінде піскен қызанақ жемістерін таңу бойынша зерттеу нәтижелері келтірілген. HSV түсті моделі және k-алгоритмін қолдана отырып түсті сегментация-салыстырмалы Алгоритмдер ретінде және машиналық оқыту негізінде Yolo қолдану кезінде практикалық қолдану үшін қызанақ бойынша әмбебап зияткерлік жектеу жүйесі ұсынылған.

Бұл зерттеу қызанақтың пісіп жетілуін бағалау үшін ең жақсы сипаттамалары мен дәлдігі бар арзан шешімді ұсынуға бағытталған. Нәтижелер дәлдік, жоғалту қисықтары және матрицаның шатасуы тұрғысынан жиналады. Нәтижелер ұсынылған модель 99% дәлдікті қамтамасыз ететін қызанақты жектеу міндеттері үшін зерттеушілер қолданатын машинаны оқытудың басқа әдістерінен (ML) асып түсетінін көрсетті.

Түйін сөздер: машиналық оқыту, қызанақты таңу, роботты жүйелер, ауылшаруашылық робот, HSV сегментациясы, Yolo.

**Ж.А. БУРИБАЕВ², Ж. АМИРГАЛИЕВА¹, С.К. ДЖОЛДАСБАЕВ^{1,2},
М.С. ЖАСҰЗАҚ^{1,2}, А.С. ТУРЕГАЛИ², Д. ДАУЛЕТІЯ²**

¹Институт информационных и вычислительных технологий,

²Казахский национальный университет имени аль-Фараби,

Алматы, Казахстан

РАСПОЗНАВАНИЕ ЗРЕЛОСТИ ПЛОДОВ ТОМАТА С ИСПОЛЬЗОВАНИЕМ МАШИННОГО ОБУЧЕНИЯ YOLO 5

Внедрение робототехнических систем и цифровизация в сельском хозяйстве на сегодняшний день является важной задачей. В данной работе рассмотрена возможность применения методов

распознавания образов и машинного обучения как компьютерная модель сельскохозяйственного робота, предназначенной для сбора урожая. Одной из актуальных задач является классификация томата по зрелости плодов в соответствии с их жизненными циклами, которые можно идентифицировать по их цвету: зеленый – на этапе растущего, желтый – в стадии предварительной созревания, а красный – созревший. Обычные методы навыков не могут соответствовать точным критериям выбора для современного управления производством в сельскохозяйственном секторе, поскольку они потребляют много времени и низкой точности. Автоматическое поведение экстракции функций с использованием машинного обучения наиболее эффективно в классификации изображений и задачах распознавания. Таким образом, в статье представлены результаты исследования по признанию спелых томатных фруктов роботизированной системой, проведенной в рамках грантового проекта Министерства образования и науки Республики Казахстан AP08857573 и реализованы классические алгоритмы, основанные на основе Цветовая модель HSV и сегментация цвета с использованием алгоритма k -ближайших в качестве сравнительных алгоритмов и на основе машинного обучения. Для практического использования при использовании Yolo предложена универсальная интеллектуальная система классификации томата.

Это исследование направлено на предоставление недорогого решения с наилучшими характеристиками и точностью для оценки зрелости томата. Результаты собираются с точки зрения точности, кривых потерь и путаницы матрицы. Результаты показали, что предлагаемая модель превосходит другие методы обучения машины (ML), используемые исследователями для задач классификации томатов, обеспечивая точность 99%.

Ключевые слова: машинное обучение, распознавание помидоров, роботизированные системы, сельскохозяйственный робот, сегментация HSV, Yolo.