



**КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ И АВТОМАТИЗАЦИЯ ПРОЕКТИРОВАНИЯ/COMPUTER
MODELING AND DESIGN AUTOMATION**

DOI: <https://doi.org/10.60797/itech.2026.11.6>

EDN: FALDZV

**АЛГОРИТМЫ ЗАХВАТА И ЛОКАЛИЗАЦИИ ДЛЯ ДВУРУКИХ РОБОТОВ В ИНТЕЛЛЕКТУАЛЬНЫХ
ЛАБОРАТОРИЯХ**

Научная статья

Косимов И.Л.¹, Ма Зенхуай² *^{1, 2} Таджикский национальный университет, Душанбе, Таджикистан

* Корреспондирующий автор (zhenhuai[at]aus.edu)

Предложена: 28.04.2026; Принята: 24.06.2026; Опубликовано: 14.07.2026

Аннотация

Для решения задачи точного определения местоположения прозрачных и полупрозрачных объектов в условиях сложного освещения в интеллектуальных лабораториях в данной статье предлагается алгоритм захвата и локализации, который объединяет улучшенную сверточную нейронную сеть с кластеризацией пиков плотности (DPC-CNN) с улучшенным алгоритмом обнаружения границ Кэнни. Алгоритм DPC оптимизирует процесс извлечения признаков CNN, эффективно устраняя избыточные признаки, вызванные отражениями от стеклянных объектов. Кумулятивное вероятностное преобразование Хафа (PPHT) используется для аппроксимации границ в пределах области интереса (ROI), обеспечивая высокоточную локализацию. Экспериментальные результаты показывают, что этот метод обеспечивает точность распознавания 96% и стабильно высокий процент успешного захвата (выше 98%) в различных условиях освещения в лаборатории, демонстрируя высокую ценность для инженерного применения.

Ключевые слова: интеллектуальная лаборатория, захват роботом, визуальная локализация, кластеризация пиков плотности, глубокое обучение, прозрачное распознавание объектов.

**CAPTURE AND LOCALIZATION ALGORITHMS FOR TWO-ARMED ROBOTS IN INTELLIGENT
LABORATORIES**

Research article

Kosimov I.L.¹, Ma Zhenhuai² *^{1, 2} Tajik National University, Dushanbe, Tajikistan

* Corresponding author (zhenhuai[at]aus.edu)

Suggested: 28.04.2026; Accepted: 24.06.2026; Published: 14.07.2026

Abstract

To solve the problem of accurately determining the location of transparent and translucent objects in difficult lighting conditions in intelligent laboratories, this article proposes a capture and localization algorithm that combines an improved convolutional neural network with density peak clustering (DPC-CNN) with an improved Canny boundary detection algorithm. The DPC algorithm optimizes the CNN feature extraction process, effectively eliminating redundant features caused by reflections from glass objects. The cumulative probabilistic Hough transform (PPHT) is used to approximate boundaries within the region of interest (ROI), providing highly accurate localization. Experimental results show that this method provides 96% recognition accuracy and a consistently high percentage of successful capture (above 98%) in various lighting conditions in the laboratory, demonstrating high value for engineering applications.

Keywords: intelligent laboratory, robot capture, visual localization, density peak clustering, deep learning, transparent object recognition.

Введение

С быстрым развитием биомедицины, материаловедения и других областей высокопроизводительные эксперименты стали ключевым звеном научных исследований и производства. Традиционные лабораторные операции в значительной степени зависят от ручного труда, что не только неэффективно, но и создает угрозу безопасности при работе с высокотоксичными, радиоактивными или хрупкими биологическими образцами. Поэтому разработка интеллектуальных лабораторных роботов, способных заменить ручной труд в точных операциях, стала актуальной темой исследований в области искусственного интеллекта и робототехники [1].

В процессе автоматизации лаборатории система визуального позиционирования является «глазом» робота, и ее основная задача – точно идентифицировать целевые инструменты на сложных экспериментальных столах и обеспечить точные координаты захвата. Однако лабораторная среда имеет свои особенности: преломление и отражение, вызванные прозрачностью стеклянных инструментов, визуальные помехи, вызванные различными уровнями жидкости, и взаимное перекрытие при плотном расположении инструментов – все это создает серьезные проблемы для традиционных алгоритмов визуального позиционирования.

1.1. Текущее состояние исследований в стране и за рубежом

В области визуальной локализации роботов отечественные и зарубежные ученые добились определенных успехов. За рубежом исследования в основном сосредоточены на обобщенных моделях захвата, основанных на глубоком

обучении с подкреплением, таких как серия Dex-Net, предложенная Малером и др. [2], которая пытается улучшить способность робота обрабатывать неизвестные объекты с помощью большого объема данных моделирования. Для прозрачных объектов Саджан и др. предложили модель ClearGrasp, которая использует технологию восстановления глубины для исправления информации о глубине прозрачных объектов [3]. В Китае исследования по конкретным сценариям (таким как промышленная сортировка и сельскохозяйственная уборка урожая) относительно зрелые.

Исследования визуального позиционирования прозрачных и полупрозрачных объектов в лабораторных условиях все еще находятся на ранней стадии. Существующие методы, такие как улучшенная серия YOLO, хорошо работают при работе с непрозрачными объектами [4], но точность распознавания и позиционирования значительно снижается в условиях частой смены источников света или сильных отражений от стеклянных материалов в лаборатории. В некоторых исследованиях используется технология слияния информации RGB-D и алгоритм кластеризации K-средних для извлечения информации о местоположении, но этапы предварительной обработки являются громоздкими, а эффективность отклика в реальном времени низка.

Для оптимизации точности позиционирования точки захвата лабораторных объектов в данной работе используются преимущества глубокого обучения в синтезе захвата на основе данных и исследованиях сетей предложения регионов [5], [6], и предлагается алгоритм позиционирования, который интегрирует кластеризацию пространственных признаков и геометрическую подгонку краев.

Разработка интеллектуального решения для позиционирования лабораторного робота. Логика определения точек захвата лабораторного оборудования

В процессе работы лабораторного робота основными объектами являются экспериментальные инструменты различных спецификаций. Визуальные особенности экспериментальных инструментов (такие как отверстия пробирок и края контейнеров) значительно зависят от окружающей среды. Например, при сильном холодном свете отражательные особенности края стекла становятся очевидными; в то время как при слабом окружающем освещении или наличии желтого теплого света особенности края легко сливаются с фоном экспериментального стола. Система использует визуальный датчик, установленный на конце, для сбора изображений рабочей зоны, выполняет первоначальную локализацию цели с помощью, улучшенной DPC-CNN, а затем выполняет точное соответствие краев внутри области интереса (ROI) захвата, в итоге направляя двурукого робота для выполнения совместного захвата [7].

2.1. Проверка лабораторного оборудования на основе DPC-CNN

В лабораторных условиях форма объектов (таких как прозрачные пробирки и полупрозрачные коробки с наконечниками для пипеток) нелинейно изменяется под разными углами и при различной интенсивности света, что приводит к неспособности роботизированных манипуляторов точно выравнивать инструменты во время позиционирования. Сверточные нейронные сети (CNN) в глубоком обучении обладают характеристиками совместного использования параметров, что позволяет эффективно использовать информацию о локальной корреляции для извлечения признаков. В данной работе используется механизм совместного использования параметров, позволяющий сети лучше адаптироваться к характеристикам инструментов в различных лабораторных условиях (например, белые столешницы и металлические вытяжные шкафы).

Учитывая, что признаки, извлекаемые с помощью сверточных нейронных сетей, часто носят нелинейный характер, традиционная линейная кластеризация может оказаться неэффективной для их обработки. В данной работе используется улучшенный алгоритм кластеризации пиков плотности (DPC) для оптимизации сверточных слоев, выбора эффективных признаков, извлекаемых сверточными ядрами, и обучения сети с использованием алгоритма обратного распространения ошибки среднего квадрата с обучением под наблюдением. Структура сети состоит из двух сверточных слоев, двух слоев пулинга и полносвязного слоя.

2.2. Механизм оптимизации карты признаков

В качестве входных данных выбирается изображение H , и сверточный слой генерирует последовательность карт признаков $\{h_1, h_2, \dots, h_m\}$. Для фильтрации признаков, чувствительных к прозрачным контурам, определяются локальная плотность w_j и относительное расстояние β_j каждой точки данных признака [8]:

$$w_j = \sum_i \exp\left(-\frac{e_{ji}^2}{e_D^2}\right)$$

$$\beta_j = \min_{i: w_i > w_j} e_{ji}$$

где e_{ji} – евклидово расстояние между векторами признаков, а e_D – расстояние отсечения. Путем вычисления комплексного показателя $\chi_j = w_j \times \beta_j$ выбираются D карт признаков с более высокими показателями в качестве основных признаков, а карты признаков с низким вкладом, вызванные шумом окружающей среды, удаляются.

2.3. Классификация и регрессия

Для снижения частоты ошибок классификации вместо E_{ji} используется расстояние ядра e_{ji} :

$$E_{ji} = \chi_j \left[\epsilon_r(h_i) - \epsilon_r(h_j) \right]^2$$

$$r_{ji} = \frac{1}{1 + E_{ji}}$$

Вероятность принадлежности точки данных к d -му кластеру $q_{j,d}$ рассчитывается на основе сходства r_{ji} . Извлечение признаков на основе улучшенной кластеризации пиков плотности может значительно уменьшить избыточную информацию (например, отражательный шум) и повысить устойчивость к сдвигу и масштабированию.

Слой субдискретизации функционирует на основе алгоритма максимизации (Max Pooling):

$$h_{pool} = \frac{q_{j,d}}{\max(h_{x,y}, h_{x+1,y}, h_{x,y+1}, h_{x+1,y+1})}$$

После уменьшения размерности с помощью максимального пулинга векторы признаков отображаются на полносвязные слои. Для определения типа инструмента используется классификатор softmax, а точность классификации повышается за счет использования методов обучения глубоких остаточных сетей и сверхглубоких сетей [9], [10].

$$Q(\bar{h} = z) = \frac{\exp(\bar{w}_z^T \bar{h})}{\sum_{k=1}^K \exp(\bar{w}_k^T \bar{h})}$$

где $\bar{w}$ – параметр веса, а K – общее количество категорий бытовой техники.

2.4. Точное позиционирование точки захвата экспериментальной установки

Несмотря на то, что усовершенствованные сверточные нейронные сети позволяют получать информацию о положении целевых инструментов (таких как пробирки, наконечники пипеток и т. д.), в реальной работе робота все еще могут возникать ошибки позиционирования. Это в основном связано с высокой отражательной способностью фона, преломлением света в стекле и системным шумом датчиков в лабораторных условиях, что делает невозможным прямое использование выходных координат сверточной нейронной сети для высокоточной настройки роботизированной руки. Алгоритм обнаружения границ Кэнни обладает высокой помехоустойчивостью и подходит для извлечения контуров лабораторных инструментов с высокой степенью прозрачности.

В данной работе отмечаются координаты центра, геометрические вершины и данные описанного прямоугольника инструмента, устанавливается ограничивающий прямоугольник (ROI) под ключевыми частями инструмента, выполняется улучшенное обнаружение границ Кэнни внутри прямоугольника, а затем выполняется обнаружение и подгонка с помощью кумулятивного вероятностного преобразования Хафа на основе обнаруженной информации о границе для извлечения прямых отрезков рукоятки или края инструмента, что обеспечивает точное позиционирование.

2.5. Улучшенное обнаружение границ методом Кэнни

На основе классической теории обнаружения границ Кэнни [11] в данной работе используется адаптивная пороговая сегментация для неравномерного лабораторного освещения. Область ROI разделена на фоновую область $\hat{H}_1$, область-кандидат $\hat{H}_2$ и область границы $\hat{H}_3$. Оптимальная комбинация порогов динамически ищется путем максимизации дисперсии $\vartheta^2(T)$ между тремя областями [12].

$$\vartheta^2(T) = \omega_1 (L_1 - L)^2 + \omega_2 (L_2 - L)^2 + \omega_3 (L_3 - L)^2$$

где ω_k – вероятность появления k -й области, L_k – ожидаемый средний градиент этой области, а L – ожидаемое значение градиента всей области. Эта трехуровневая стратегия пороговой сегментации позволяет эффективно сохранять слабые контуры стеклянной посуды, устраняя ложные края, вызванные световыми пятнами, что обеспечивает надежность извлечения краев в сложных условиях лабораторного освещения.

2.6. Визуальная локализация точки захвата

После обнаружения границ, учитывая, что лабораторное оборудование (например, ручки пробирок и края наконечников пипеток) обычно обладает очевидными геометрическими прямолинейными особенностями, а отражающий шум проявляется в виде нерегулярных кривых, вероятностное кумулятивное преобразование Хафа (РРНТ) может эффективно идентифицировать прямолинейные цели путем подсчета количества голосов в пикселях. Поэтому в данной работе используется метод РРНТ для обнаружения прямолинейных линий на обнаруженных сегментах границ [13]. В качестве разумных параметров для эффективной фильтрации ложных границ, вызванных преломлением света в стекле, установлены пороговое значение накопленной плоскости и минимальная длина отрезка линии. Оставшиеся прямолинейные сегменты представляют собой геометрические оси оборудования, а уравнение аппроксимирующей прямой линии имеет вид:

$$\lambda(\varphi) = \frac{B^2}{2} - \tau \left\{ - \left[\frac{\varphi \vartheta^2(n, m)}{B} \right]^2 \right\}$$

В формуле $\lambda(\varphi)$ – геометрическая линия, аппроксимирующаяся роботом; B – константа аппроксимации (значение 1,4203); и τ – ошибка аппроксимации. Координаты точки пересечения аппроксимирующей линии $\lambda(\varphi)$ и горизонтальной оси симметрии интересующего прямоугольника задаются в качестве положения точки захвата. На основе этого робот корректирует положение концевой эффектора роботизированной руки с помощью алгоритма ПИД-регулирования, основанного на информации обратной связи, полученной от датчиков (таких как камеры и датчики силы), так что его захват точно перемещается в целевые координаты и выполняет захватное действие. Этот метод, интегрирующий извлечение признаков с помощью глубокого обучения и классическую геометрическую аппроксимацию, значительно повышает точность визуального позиционирования лабораторных роботов в условиях нестабильного освещения.

2.7. Настройка и оценка эксперимента

Экспериментальная платформа состоит из двух независимых легких шести осевых (шестиступенных) роботизированных манипуляторов, образующих единую двурукую робототехническую систему, камеры на концевом захвате и промышленного вычислительного блока, как показано на рисунке 1.

Экспериментальный сценарий имитирует типичный интеллектуальный лабораторный стол с оборудованием, включающим стеклянные пробирки, колбы Эрленмейера, центрифужные пробирки и коробки с наконечниками для пипеток различных размеров. Набор данных содержит 9677 изображений типичной лабораторной сцены. 70% изображений используются в качестве обучающего набора, 20% – в качестве проверочного набора и 10% – в качестве тестового набора.

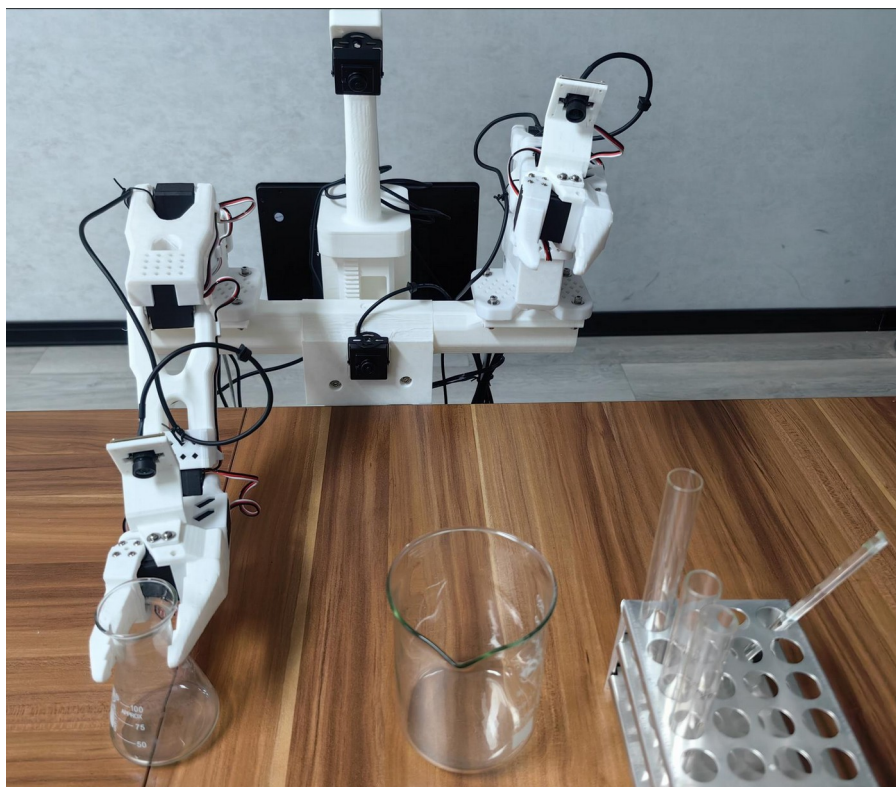


Рисунок 1 - Лабораторный робот с двумя манипуляторами
DOI: <https://doi.org/10.60797/itech.2026.11.6.1>

В эксперименте первый слой сверточной нейронной сети имел размер ядра 3×3 , количество ядер 16 и шаг 1; второй слой имел размер ядра 3×3 , количество ядер 32 и шаг 1. На основе этого была обучена сеть, и изменение точности распознавания устройств в зависимости от количества итераций показано на рисунке 2. Как видно из рисунка, когда количество итераций достигает 220, точность распознавания достигает 96%, что доказывает, что предложенный метод может точно распознавать экспериментальные устройства.

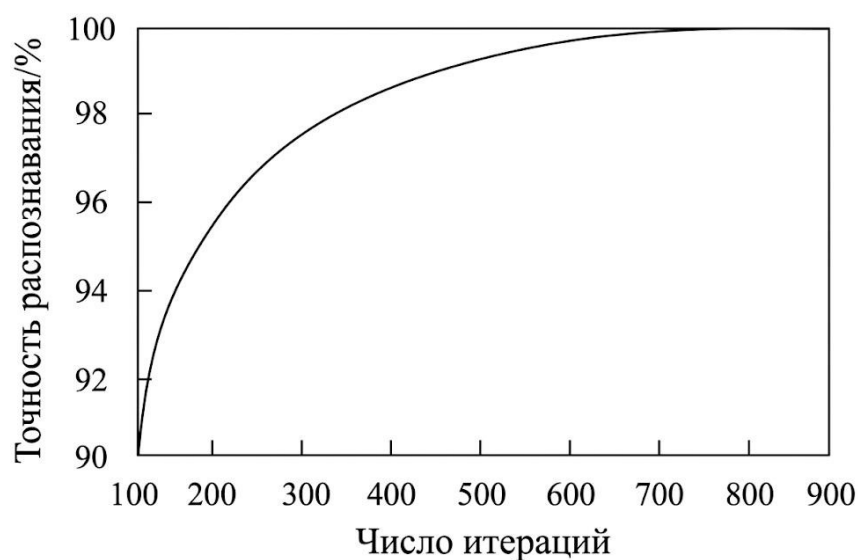


Рисунок 2 - Точность распознавания
DOI: <https://doi.org/10.60797/itech.2026.11.6.2>

Помимо указанных выше параметров, параметры, используемые при реализации описанного в данной статье метода, перечислены в таблице 1.

Таблица 1 - Настройки параметров алгоритма

DOI: <https://doi.org/10.60797/itech.2026.11.6.3>

Название метода	Параметры	Значения настроек
Алгоритм DPC	Радиус отсечения e_D	0,15
Структура CNN	Размер ядра и Шаг	3×3 / 1
Обнаружение Кэнни	Размер ядра Гаусса	5×5
Метод PRHT	Минимальная длина сегмента и максимальная косвенность	50 px / 10 px

Обсуждение результатов эксперимента

1. Анализ в реальном времени: после внедрения оптимизации DPC время обработки одного кадра изображения сокращается с 0,08 с до 0,045 с. В сочетании со стратегией захвата с координацией рук и глаз, основанной на больших объемах данных [13], это полностью отвечает требованиям онлайн-работы робота.

2. Ошибка позиционирования: ошибка позиционирования пикселя Ф определяется как евклидово расстояние между прогнозируемой точкой захвата и истинным значением. В условиях сильного отражения средняя ошибка составляет всего 0,17d (d – диаметр прибора), что значительно лучше, чем 0,45d у традиционной комбинации YOLO-Canny.

Таблица 2 - Статистические данные по показателю успешности захвата роботом в различных условиях работы

DOI: <https://doi.org/10.60797/itech.2026.11.6.4>

Условия освещения	Тип цели	Ошибка позиционирования	Процент успешности захвата
Стандартное комнатное освещение	Стеклоанная колба Эрленмейера	0,11d	99,2%
Высокая отражательная способность (полдень)	Стеклоанная колба Эрленмейера	0,21d	98,1%
Низкая освещенность (тьма)	Стеклоанная колба Эрленмейера	0,19d	98,5%

Заключение

В данной статье предлагается алгоритм визуальной локализации, основанный на DPC-CNN и слиянии геометрических признаков, для сценариев интеллектуальной лаборатории. Исследование показывает, что оптимизация пространства признаков с помощью алгоритмов кластеризации значительно повышает устойчивость робота к помехам от прозрачных объектов. Этот алгоритм не только обеспечивает высокоточный захват и локализацию (с вероятностью успеха 98%), но и отвечает требованиям реального времени промышленных приложений. Дальнейшая работа будет сосредоточена на предотвращении сложных перекрытий в многоруких средах совместной работы и восприятии направления с помощью слияния данных от нескольких датчиков, чтобы еще больше повысить операционную гибкость робота в сложных и неструктурированных средах.

Конфликт интересов

Не указан.

Рецензия

Гибадуллин Р.Ф., Казанский национальный исследовательский технический университет им. А.Н. Туполева – КАИ, Казань Российская Федерация
DOI: <https://doi.org/10.60797/itech.2026.11.6.5>

Conflict of Interest

None declared.

Review

Gibadullin R.F., Kazan National Research Technical University named after A.N. Tupolev – KAI, Kazan Russian Federation
DOI: <https://doi.org/10.60797/itech.2026.11.6.5>

Список литературы на английском языке / References in English

1. Burger B. A mobile robotic chemist / B. Burger, P.M. Maffettone, V.V. Gusev [et al.] // Nature. — 2020. — Vol. 583. — № 7815. — P. 237–241. — DOI: 10.1038/s41586-020-2442-2.
2. Mahler J. Dex-Net 2.0: Deep learning to plan robust grasps with synthetic point clouds and analytic grasp metrics / J. Mahler, J. Liang, S. Niyaz [et al.] // Robotics: Science and Systems (RSS). — 2017. — URL: <https://www.roboticsproceedings.org/rss13/p58.pdf> (accessed: 28.04.2026).
3. Sajjan S. ClearGrasp: 3D shape estimation of transparent objects for manipulation / S. Sajjan, M. Moore, M. Pan [et al.] // Proceedings of the IEEE International Conference on Robotics and Automation (ICRA). — 2020. — P. 2381–2387. — DOI: 10.1109/ICRA40945.2020.9197518.



4. Bochkovskiy A. YOLOv4: Optimal speed and accuracy of object detection / A. Bochkovskiy, C.Y. Wang, H.Y.M. Liao // arXiv. — 2020. — URL: <https://arxiv.org/abs/2004.10934> (accessed: 28.04.2026).
5. Bohg J. Data-driven grasp synthesis: A survey / J. Bohg, A. Morales, T. Asfour [et al.] // IEEE Transactions on Robotics. — 2014. — Vol. 30. — № 2. — P. 289–309. — DOI: 10.1109/TRO.2013.2289018.
6. Ren S. Faster R-CNN: Towards real-time object detection with region proposal networks / S. Ren, K. He, R. Girshick [et al.] // Advances in Neural Information Processing Systems. — 2015. — Vol. 28. — P. 91–99.
7. Kragic D. Vision for robotics / D. Kragic, M. Vincze // Foundations and Trends in Robotics. — 2009. — Vol. 1. — № 1. — P. 1–78. — DOI: 10.1561/23000000001.
8. Rodriguez A. Clustering by fast search and find of density peaks / A. Rodriguez, A. Laio // Science. — 2014. — Vol. 344. — № 6191. — P. 1492–1496. — DOI: 10.1126/science.1242072.
9. He K. Deep residual learning for image recognition / K. He, X. Zhang, S. Ren [et al.] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). — 2016. — P. 770–778.
10. Simonyan K. Very deep convolutional networks for large-scale image recognition / K. Simonyan, A. Zisserman // arXiv. — 2014. — URL: <https://arxiv.org/abs/1409.1556> (accessed: 28.04.2026).
11. Canny J. A computational approach to edge detection / J. Canny // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 1986. — № 6. — P. 679–698. — DOI: 10.1109/TPAMI.1986.4767851.
12. Otsu N. A threshold selection method from gray-level histograms / N. Otsu // IEEE Transactions on Systems, Man, and Cybernetics. — 1979. — Vol. 9. — № 1. — P. 62–66. — DOI: 10.1109/TSMC.1979.4310076.
13. Matas J. Robust line detection using the progressive probabilistic Hough transform / J. Matas, C. Galambos, J. Kittler // Computer Vision and Image Understanding. — 2000. — Vol. 78. — № 1. — P. 119–137. — DOI: 10.1006/cviu.1999.0831.
14. Levine S. Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection / S. Levine, P. Pastor, A. Krizhevsky [et al.] // The International Journal of Robotics Research. — 2018. — Vol. 37. — № 4–5. — P. 421–436. — DOI: 10.1177/0278364917710318.