

РОБОТЫ, МЕХАТРОНИКА И РОБОТОТЕХНИЧЕСКИЕ СИСТЕМЫ



УДК 004.896

DOI:10.30724/1998-9903-2026-28-4-223-233

МЕТОД АВТОМАТИЧЕСКОЙ ИДЕНТИФИКАЦИИ ОБЪЕКТОВ И ОПРЕДЕЛЕНИЯ МЕСТА ЗАХВАТА МАНИПУЛЯТОРОМ РОБОТА С ИСПОЛЬЗОВАНИЕМ СТЗ

Андреев В.П., Карпов М.В.

Московский государственный технологический университет «СТАНКИН»,
Москва, Россия andreevira@yandex.ru

Резюме: *АКТУАЛЬНОСТЬ.* Автоматизация операций манипулирования объектами сложной формы в промышленном производстве является важной задачей современной робототехники. В ряде практических задач объекты располагаются на рабочей плоскости робототехнического комплекса без перекрытий, однако имеют произвольную ориентацию, что существенно усложняет процесс их автоматической идентификации и определения координат места захвата манипулятором. Применение трёхмерных систем технического зрения позволяет решать данную задачу, однако требует использования специальных RGB-D камер и высокопроизводительной вычислительной техники, что существенно увеличивает стоимость комплекса. *ЦЕЛЬ.* Разработка и экспериментальное исследование метода автоматической идентификации объектов сложной формы и определения координат их места захвата роботом-манипулятором с использованием системы технического зрения, снабжённой лишь одной монокулярной камерой. *МЕТОДЫ.* Для распознавания объектов на цифровых изображениях применяется свёрточная искусственная нейронная сеть архитектуры YOLOv8, позволяющая определять положение и ориентацию объектов на рабочей поверхности. Координаты места захвата вычисляются на основе анализа положения ограничивающего прямоугольника и экспериментально определяемого смещения точки захвата относительно геометрического центра объекта. *РЕЗУЛЬТАТЫ.* Проведено экспериментальное исследование метода на примере распознавания шахматных фигур различной формы и цвета. Достоверность обнаружения объектов составила не менее 0,9, а время обработки одного изображения составляет менее 56 мс, что обеспечивает работу системы в режиме реального времени. Показана возможность устойчивого определения координат мест захвата объектов при их произвольной ориентации. *ЗАКЛЮЧЕНИЕ.* Предложенный метод позволяет реализовать автоматическую идентификацию объектов сложной формы и определение координат их захвата с использованием лишь одной монокулярной камеры, что делает его перспективным для автоматизации операций сортировки и иного манипулирования объектами в промышленности.

Ключевые слова: автоматизация; робототехнический комплекс; цифровая обработка изображений; распознавание; система технического зрения; искусственная нейронная сеть; захват объекта.

Для цитирования: Андреев В.П., Карпов М.В. Метод автоматической идентификации объектов и определения места захвата манипулятором робота с использованием СТЗ// Известия высших учебных заведений. ПРОБЛЕМЫ ЭНЕРГЕТИКИ. 2026. Т. 28. № 4. С. 223-233. doi: 10.30724/1998-9903-2026-28-4-223-233.

METHOD FOR AUTOMATIC OBJECT IDENTIFICATION AND ROBOTIC GRASP POINT DETERMINATION USING MACHINE VISION

Andreev V.P., Karpov M.V.

Moscow State University of Technology "STANKIN", Moscow, Russia
andreevipa@yandex.ru

Abstract: RELEVANCE of the study lies in the automation of manipulation operations involving complex-shaped objects in industrial production, which remains an important task in modern robotics. In many practical applications, objects are located on the working surface of a robotic system without overlapping but may have arbitrary orientations, which significantly complicates their automatic identification and the determination of grasp coordinates for a robotic manipulator. Although three-dimensional machine vision systems can solve this problem, they usually require RGB-D cameras and high-performance computing hardware, which significantly increases the cost and complexity of robotic systems. Therefore, the development of efficient methods capable of solving this task using only a monocular camera remains relevant. THE PURPOSE of the study is to develop and experimentally investigate a method for automatic identification of complex-shaped objects and determination of their grasp coordinates by a robotic manipulator using a machine vision system equipped with a single monocular camera. METHODS. Object recognition in digital images is performed using a convolutional neural network based on the YOLOv8 architecture, which allows detection of objects and estimation of their orientation on the working surface. The grasp coordinates are determined by analyzing the bounding box of the detected object and applying an experimentally determined offset of the grasp point relative to the geometric center of the object. RESULTS. Experimental evaluation was carried out using chess pieces of different shapes and colors as test objects. The object detection confidence was not lower than 0.9, while the processing time of a single image was less than 56 ms, which ensures real-time operation. CONCLUSIONS. The proposed method enables automatic identification of complex-shaped objects and stable determination of their grasp coordinates using only a single monocular camera without expensive three-dimensional sensors.

Keywords: automation; robot system; digital image processing; recognition; machine vision; artificial neural network; object capture.

For citation: Andreev V.P., Karpov M.V. Method for automatic object identification and robotic grasp point determination using machine vision. *Power engineering: research, equipment, technology*. 2026; 28 (4): 223-233. doi: 10.30724/1998-9903-2026-28-4-223-233.

Введение (Introduction)

Современные производственные системы ориентированы на повышение гибкости и снижение доли ручного труда при выполнении операций сортировки и манипулирования различными объектами (например, деталями сборки). При автоматизации подобных операций особую значимость приобретают задачи автоматического распознавания, захвата и перемещения объектов сложной формы, которые могут располагаться в рабочей зоне робототехнического комплекса (РТК) в произвольной ориентации.

В настоящее время одним из ключевых направлений в роботизации промышленного производства является широкое использование в РТК систем технического зрения (СТЗ) с применением методов глубокого обучения для автоматического распознавания и локализации объектов манипулирования. В обзоре [1] рассмотрены основные задачи компьютерного зрения, такие как классификация, локализация и детектирование объектов, отмечена возрастающая роль свёрточных нейронных сетей в обработке изображений для робототехнических приложений. В последние годы применение свёрточных искусственных нейронных сетей (ИНС) в системах технического зрения продемонстрировало высокую эффективность в задачах обнаружения и классификации объектов в реальном времени [2,3]. К примеру, в исследовании [4] подтверждается эффективность интеграции систем компьютерного зрения в контур управления мехатронных станций сортировки для автоматической идентификации изделий. Применение нейросетевых алгоритмов доказало свою эффективность и в смежных задачах промышленной автоматизации, таких как мониторинг устойчивости электротехнических систем, где использование ИНС позволило существенно снизить число необоснованных отключений нагрузки и повысить надёжность непрерывных производств [5]. Аналогично, в задачах технического

диагностирования методы машинного обучения позволяют эффективно обнаруживать аномалии в данных датчиков, подтверждая универсальность данных подходов для обработки сигналов сенсорных систем [6].

Для решения задач обнаружения и классификации объектов в составе СТЗ широко применяются современные архитектуры детекторов. Подходы к детекции объектов можно разделить на двухстадийные и одностадийные. Двухстадийные архитектуры обеспечивают высокую точность за счёт разделения этапов генерации регионов интереса и классификации, однако требуют значительных вычислительных ресурсов [7]. Среди одностадийных детекторов особое место занимают архитектуры семейства YOLO [8]. Развитие архитектур YOLO (*You Only Look Once*) отражено в работах [2, 9, 10, 11], где продемонстрировано последовательное улучшение баланса между скоростью и точностью обнаружения. Современные возможности архитектур семейства YOLOv7 и YOLOv8 описаны в работах [11] и [2], где показана поддержка специализированных задач, включая детекцию с ориентированными ограничивающими прямоугольниками. В российских исследованиях также уделяется внимание анализу производительности YOLO-архитектур: проведено экспериментальное сравнение разновидностей YOLO для задач компьютерного зрения [12], рассмотрено применение свёрточных сетей семейства YOLO в мобильных системах технического зрения [13], а также продемонстрирована эффективность данных архитектур для автоматического обнаружения полезных сигналов на многоканальных дефектограммах [14]. Универсальность нейросетевого подхода подтверждается его успешным применением для решения нелинейных задач обработки данных в различных областях, включая прогнозирование сложных процессов [15].

Важным аспектом построения СТЗ является выбор аппаратного обеспечения. Наиболее точные данные о пространственном расположении объектов обеспечивают RGB-D камеры, формирующие облака точек. В обзоре [16] проанализированы методы обнаружения захвата на основе RGB-D данных, обеспечивающие высокую точность (до 95–98%), но отмечена высокая стоимость и сложность калибровки таких сенсоров. Альтернативой выступают монокулярные камеры, которые дешевле и проще в интеграции [17]. При выборе сенсорной базы важно учитывать не только стоимость, но и устойчивость к помехам. В работах по сенсорному обеспечению мобильных роботов показано, что даже при использовании недорогих датчиков возможно решение задач определения формы и расстояния при условии корректной обработки сигналов и минимизации артефактов в каналах передачи данных [18]. В российских исследованиях также рассматриваются вопросы применения монокулярного зрения: предложен метод позиционирования рабочего инструмента робота-манипулятора с использованием камеры, показавший возможность повышения точности позиционирования при корректной калибровке [19].

Задача определения места захвата является критической для успешного манипулирования. В работе [20] предложены системы, использующие глубокое обучение для планирования надёжных захватов с применением синтетических (сгенерированных) облаков точек.

Вопросы привязки систем координат и определения пространственных параметров в системах технического зрения роботов требуют детальной проработки. В работе [21] рассмотрена методика определения координат в системе технического зрения робота, подчеркнута важность корректной калибровки камер и перехода от пиксельных координат изображения к метрической системе координат манипулятора.

Приведённый анализ публикаций позволяет выделить ряд нерешённых вопросов в пределах общей проблемы, что показывает необходимость разработки метода, позволяющего определять координаты рационального места захвата объектов заданного класса с использованием лишь одной монокулярной камеры, статично закреплённой над рабочей зоной РТК, без применения дорогостоящих RGB-D датчиков.

Цель исследования заключается в разработке и экспериментальном исследовании метода автоматической идентификации объектов сложной формы и определения координат их рационального места захвата роботом-манипулятором с использованием системы технического зрения на основе одной монокулярной камеры. Научная значимость исследования состоит в предложении алгоритма определения 3-мерных координат рационального места захвата на основе анализа двумерного изображения, инвариантного к ориентации объекта, что позволяет обойтись без необходимости получения трёхмерных данных о сцене. Практическая значимость заключается в возможности внедрения предложенного метода при использовании РТК для автоматизации операций сортировки и манипулирования объектами сложной формы при условии их расположения на рабочей плоскости без взаимного перекрытия, что существенно снижает стоимость системы за счёт отказа от дополнительных

дорогостоящих датчиков глубины и использования высокопроизводительной вычислительной техники.

Распознавание объектов на цифровом изображении

В задаче автоматизации манипулирования объектами, расположенными на рабочей плоскости без взаимного перекрытия, первоначальным этапом является достоверная идентификация объекта и определение его пространственной ориентации по двумерному цифровому изображению. Поскольку ориентация объекта может быть произвольной, недостаточно определить только координаты его центра; необходимо также получить информацию об угле поворота в системе координат изображения, которая жёстко «привязывается» к системе координат РТК в процессе калибровки.

Для решения задачи распознавания предлагается использовать архитектуру ИНС семейства YOLO (*You Only Look Once*), относящуюся к одноэтапным детекторам объектов [2]. В работе применена модификация YOLOv8 с поддержкой ориентированных ограничивающих прямоугольников (*Oriented Bounding Boxes, OBB*). В отличие от классической версии YOLOv8, формирующей модель объектов в виде ограничивающего прямоугольника, оси которого параллельны осям изображения, OBB-модификация позволяет определять угол поворота ограничивающего прямоугольника. Таким образом, результатом работы такой нейронной сети является не только идентификация объекта из заданного класса, но и положение объекта (координаты центра и размеры) на изображении, а также его ориентация. Определение параметров ориентированных ограничивающих прямоугольников является принципиально важным для решаемой задачи, поскольку:

- объекты на рабочей плоскости располагаются без перекрытий и с произвольной ориентацией;
- корректное определение угла поворота позволяет определить положение главной оси объекта;
- информация об ориентации используется при вычислении рационального места захвата.

Таким образом, применение YOLOv8-OBB позволяет объединить этапы детекции и оценки ориентации объекта в рамках одной модели, что упрощает алгоритмическую структуру системы и снижает вычислительные затраты за счёт исключения отдельной процедуры определения угла ориентации объекта на основе анализа его контура.

«Привязка» координат цифрового изображения к системе координат РТК

Искусственная нейронная сеть требует предварительного обучения, по завершении которого ИНС готова принимать на вход цифровые изображения сцен на рабочей поверхности РТК с целью распознавания объектов. ИНС определяет наличие того или иного объекта на изображении и размечает их ограничивающими прямоугольниками. На выходе YOLOv8-OBB формируется набор параметров в системе координат цифрового изображения:

$$(x_c, y_c, w, h, \theta, p),$$

где x_c, y_c – координаты центра ограничивающего прямоугольника (*pixels*),

w, h – его размеры (*pixels*),

θ – угол поворота прямоугольника (объекта),

p – вероятность принадлежности к классу.

Необходимо выполнить перерасчёт приведённых координат в метрическую систему при помощи следующих выражений («привязка» к системе координат рабочей зоны РТК):

$$\begin{aligned} x_c^* &= \frac{x_c}{w_{изобр}} \cdot w_{pz}, \\ y_c^* &= \frac{y_c}{h_{изобр}} \cdot h_{pz}, \end{aligned}$$

где x_c^*, y_c^* – приведённые координаты (*мм*);

$w_{изобр}, h_{изобр}$ – ширина и высота цифрового изображения рабочей зоны РТК (*pixels*);

w_{pz}, h_{pz} – ширина и высота рабочей зоны РТК (*мм*).

Используя данные формулы, можно перевести значения координат на изображении из пикселей в миллиметры. Необходимо также выполнить «привязку» пространственных координат рабочего стола к системе координат манипулятора РТК. На рисунке 1 представлен процесс «привязки» разных систем координат.

Поскольку положение зоны поля зрения камеры фиксировано, а робот-манипулятор также жёстко закреплён и неподвижен, то их системы координат также статичны. Согласно схеме (рис.1), для «привязки» системы координат рабочей зоны к манипулятору необходимо сделать смещение по оси X на расстояние dx , а по оси Y на расстояние dy . Рабочая зона может находиться чуть выше основания робота-манипулятора по оси Z на величину dz . Расстояние dy соответствует половине ширины рабочей зоны, а dx является расстоянием от основания манипулятора до рабочей зоны. Данные значения необходимо

учесть при определении пространственных координат идентифицированного объекта в системе координат РТК.

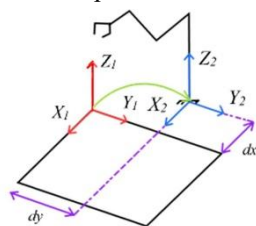


Рис.1. «Привязка» систем координат (красным – система координат рабочей зоны камеры, синим – система координат рабочей зоны манипулятора, чёрным – поле зрения камеры

Fig. 1. Coordinate system alignment (the camera working zone coordinate system is marked in red, the manipulator working zone coordinate system is marked in blue, the camera field of view is marked in black)

*Источник: Составлено авторами Source: compiled by the author.

Следует отметить, что перед выполнением пересчёта координат изображения в метрическую систему была проведена предварительная калибровка камеры смартфона по шаблону шахматной доски. Калибровка позволила определить внутренние параметры камеры и уменьшить влияние геометрических искажений объектива на результат измерения координат объектов на изображении. После калибровки выполнялось сопоставление координат рабочей области на изображении с реальными размерами рабочей зоны РТК в миллиметрах. В данной работе, с учётом фиксированного положения камеры относительно рабочей поверхности, использовалась плоская модель пересчёта координат, достаточная для рассматриваемых условий эксперимента. В настоящей работе калибровка телекамеры и оценка её качества осуществлялись по методике, описанной в [21] и [22]. Полученная точность оказалась достаточной при экспериментальной проверке решения задачи определения точек захвата шахматных фигур, поскольку для захвата манипулятором свободно лежащих объектов не требуется микронная точность.

Технология определения координат места захвата объектов

В проведённых экспериментах в качестве объектов сложной формы выбраны шахматные фигуры, в произвольном порядке лежащие на относительно равномерно освещённом поле рабочей плоскости стола (Рис.2) – это типичная ситуация при решении задач сортировки в промышленности. Выбор в качестве объектов манипулирования шахматных фигур обусловлен сложностью и значительным разнообразием их форм, а также различием по цвету – 6 объектов сложных форм и 2-х цветов.



Рис. 2. Объекты сортировки – шахматные фигуры на относительно равномерно освещённом рабочем столе

Fig. 2. Sorting objects – chess pieces on a relatively uniformly illuminated work table

*Источник: Составлено авторами Source: compiled by the author.

Для захвата объекта необходимо знать пространственные координаты места захвата. Поэтому, необходимо найти практические способы вычислить его точное положение в рабочем пространстве РТК, а именно значения координат (x, y, z) .

Казалось бы, что для удобства захвата и дальнейшего манипулирования объектами роботу необходимо схватить объект в области центра его масс (см. рисунок 3). Однако, на производстве у объектов манипулирования далеко не всегда удобное место захвата располагается в центре масс.

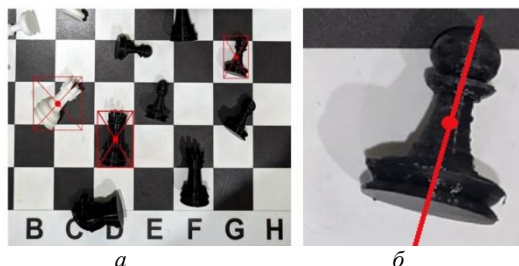


Рис. 3. Нахождение геометрического центра фигуры – а; ось симметрии – б

Fig. 3. Finding the shape geometric center – а; axis of symmetry – б

*Источник: Составлено авторами Source: compiled by the author.

В данном случае высока вероятность того, что объект может соскользнуть с губок схвата, поскольку, центр масс у шахматных фигур располагается ближе к основанию.

Отсюда вытекает очень важный вывод: способ нахождения места захвата зависит от класса объектов, которыми требуется манипулировать.

Но для нахождения оси симметрии фигуры на её цифровом изображении необходимо определить ориентацию фигур в координатах рабочей поверхности.

В отличие от универсальных методов, предполагающих определение точки захвата непосредственно из геометрического центра объекта, в данной работе место рационального захвата определяется заранее экспериментально для каждого класса объектов. На подготовительном этапе оператор выполняет серию пробных захватов (собственноручно или же при помощи ручного управления манипулятором) и экспериментально определяет положение места захвата относительно геометрического центра фигуры. В ряде случаев рациональная точка может быть смещена относительно центра прямоугольника (например, выше или ниже него), что связано с неоднородностью формы объекта или особенностями его геометрии и распределения массы.

По результатам экспериментальной настройки определяется смещение $(\Delta x, \Delta y)$, характерное для конкретного класса объектов, и вычисляются координаты x, y места захвата:

$$\begin{aligned}x &= x_c^* + \Delta x \\ y &= y_c^* + \Delta y\end{aligned}$$

Полученные параметры заносятся в базу данных (БД) системы и в дальнейшем автоматически используются при расчёте координат места захвата.

Стоит подчеркнуть, что в настоящей работе основной акцент сделан на объектах, для которых рациональная точка их захвата может быть задана через экспериментально (заранее) определяемое смещение относительно геометрического центра для всех объектов из конечного набора. Для асимметричных объектов (шахматный конь) указанное смещение может зависеть от его ориентации на рабочей поверхности, что требует отдельной предварительной настройки и рассматривается как направление дальнейшего развития метода.

Определение аппликаты z места захвата выполняется исходя из следующих условий постановки эксперимента. Телекамера жёстко закреплена в РТК на известной высоте H (Рис.4) и располагается по нормали по отношению к плоскости, где располагаются объекты.

Место захвата по оси z будет располагаться выше опорной поверхности на половину толщины лежащего объекта – b . При предварительном экспериментальном определении места захвата для каждого объекта также можно заранее определить этот параметр в миллиметрах и занести в БД.

Таким образом результирующие формулы для подсчёта координат места захвата будут выглядеть следующим образом:

$$\begin{aligned}x &= x_c^* + \Delta x + dx = \frac{x_c}{w_{изобр}} \cdot w_{pz} + \Delta x + dx \\ y &= y_c^* + \Delta y + dy = \frac{y_c}{h_{изобр}} \cdot h_{pz} + \Delta y + dy \\ z &= z_{изм} + \frac{b}{2} + dz,\end{aligned}$$

где dx – смещение по оси X СК для «привязки» системы координат рабочей зоны к манипулятору;

dy – смещение по оси Y ;

dz – высота опорной поверхности рабочего стола относительно базовой системы координат манипулятора.



Рис.4. Определение аппликаты места захвата
Fig. 4. Determination of the Z-coordinate of the grasp point

*Источник: Составлено авторами Source: compiled by the author.

Постановка эксперимента

Для проведения эксперимента применялось следующее оборудование:

– Стандартный набор из 32-х шахматных фигур.

– Смартфон, телекамера смартфона (разрешающая способность 64Мп, значение диафрагмы f/1.89, изображение цветное).

– Компьютер со следующими характеристиками: процессор: AMD Ryzen 5 3550H with Radeon Vega Mobile Gfx 2.10GHz; видеокарта: NVIDIA GeForce GTX 1650; ОЗУ: 15,8 ГБ.

– USB-кабель.

– Штатив для закрепления смартфона.

– 4х-звенный робот-манипулятор, оснащённый 2-пальцевым схватом.

На рисунке5 приведено фото робота-манипулятора с пятью степенями подвижности и фото экспериментальной установки.



Рис.5. Фото экспериментальной установки



Fig. 5. Photo of the experimental setup

*Источник: Составлено авторами Source: compiled by the author.

Возле манипулятора установлен штатив с закреплённым смартфоном, который выполняет роль телекамеры и формирует цифровое изображение сцены. Для создания однотонного фона шахматная доска, нарисованная на рабочем столе робота, закрывается зелёной бумагой, и организовано относительно равномерное бестеневое освещение сцены. Это значительно повышает достоверность распознавания объектов на цифровом изображении. Предварительно выполнена калибровка камеры смартфона по рисунку шахматной доски.

Подготовка данных для обучения ИНС

Для обучения искусственной нейронной сети архитектуры YOLOv8-OBV создана обучающая выборка. Для этого при помощи телекамеры смартфона создан набор изображений реальных сцен с шахматными фигурами (объектами), лежащими без перекрытия на рабочей поверхности РТК. Шахматные фигуры из стандартного набора разбросаны в случайном порядке. В результате было создано 212 цифровых изображений размером 4624x3472 пикселей. На каждом снимке расположение фигур, а также их количество различны, в результате чего были получены максимально разнообразные по содержанию изображения. Это требуется для лучшего обучения ИНС.

Разметка объектов на изображении выполнялась при помощи сервиса RoboFlow. После загрузки на сервис всех изображений были объявлены классы объектов. Названия классов указываются на английском языке во избежание ошибок при работе на иностранном сервисе. Используемые классы: WhitePawn (Белая пешка), BlackPawn (Чёрная пешка), WhiteHorse (Белый конь), BlackHorse (Чёрный конь), WhiteBishop (Белый слон), BlackBishop (Чёрный слон), WhiteQueen (Белый ферзь), BlackQueen (Чёрный ферзь), WhiteKing (Белый король), BlackKing (Чёрный король), WhiteRook (Белая ладья), BlackRook (Чёрная ладья). Затем выполнялась ручная разметка: открывается каждое изображение, и вручную обводятся индивидуальные рамки вокруг фигур, одновременно назначается для каждой фигуры её принадлежности к классу. Для удобства задаётся свой цвет рамки для каждого класса. Результат разметки одного из изображений приведён на рис.6.

Для повышения разнообразия исходных изображений был создан дополнительный набор данных с помощью сервиса RoboFlow, который позволяет выполнить соответствующие преобразования исходных изображений. При этом эффективность обучения значительно возрастает. В нашем случае применён эффект перевода цветных изображений в чёрно-белый формат, поскольку работа производится с объектами двух цветов – чёрные и белые (или иное сочетание только 2-х цветов). Если бы набор объектов был бы разноцветным, то применение данного преобразования, скорее всего, снизило бы достоверность распознавания объектов. Помимо этого, были применены: повышение яркости, изменение оттенков, увеличение и уменьшение экспозиции, насыщенности, а

также повороты изображений на 90 и 180 градусов. Также, размер изображений был уменьшен до 416x416 пикселей – стандартный размер, с которыми работает ИНС YOLOv8. В результате количество исходных изображений увеличилось до 604.



Рис.6. Изображение с размеченными фигурами и панель инструментов
Fig. 6. Image with annotated figures and toolbar

*Источник: Составлено авторами Source: compiled by the author.

Следующим шагом является разделение всех полученных данных на тренировочный (train), валидационный (valid) и тестовый (test) наборы. На рисунке 7 представлено процентное соотношение между наборами данных.



Рис.7. Процентное соотношение между наборами данных
Fig.7. Percentage distribution among datasets

*Источник: Составлено авторами Source: compiled by the author.

Результаты (Discussions)

Процесс выполнения эксперимента состоял из следующих этапов:

1. Подготовка рабочего места: установка штатива, создание однотонного фона на рабочей поверхности.
2. Процесс калибровки: привязка системы координат цифрового изображения рабочей поверхности к системе координат манипулятора.
3. Подготовка данных и обучение ИНС: 604 изображений размером 4624x3472 пикселей, которые вследствие были уменьшены до размера 416x416 пикселей. Обучение проходило в течение 100 эпох и завершилось за 51 минуту. Достоверность обнаружения каждого объекта на цифровом изображении составила не менее 0,9 при скорости обработки одного изображения 56,3 миллисекунды.
4. Создание цифрового изображения и его подача на обученную ИНС.
5. Определение места захвата (x, y, z).
6. Координаты передаются в программу, решающую обратную задачу кинематики (ОЗК) манипулятора.
7. Значения обобщённых координат вносятся в инженерную панель манипулятора, и манипулятор со схватом начинает перемещение.
8. Управление схватом и, в последствии, – захват объекта.
9. Объект захвачен и готов к манипулированию.

Результат идентификации объектов представлен на рис.8. Вывод настроен таким образом, чтобы на цифровом изображении выводились предсказания с достоверностью не менее 0,9.

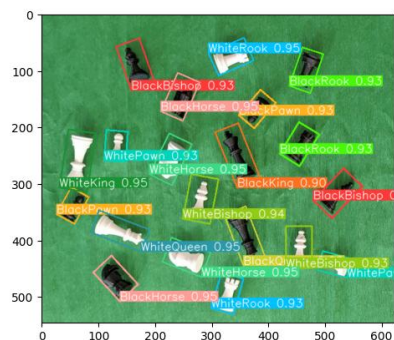


Рис.8. Результат идентификации
Fig.8. Identification result

*Источник: Составлено авторами Source: compiled by the author.

По результатам работы ИНС можно отметить два важных момента. Во-первых, скорость обработки одного изображения составила 56,3 миллисекунды, что обеспечивает работу системы в режиме реального времени. Во-вторых, в ходе идентификации

достоверность обнаружения каждой из шахматных фигур составила не менее 0,9, что видно из таблицы.

Таблица
Table

Достоверность распознавания каждой фигуры
The accuracy of recognizing each figure

Фигура	Достоверность
Белая пешка	0.93
Чёрная пешка	0.93
Белый конь	0.95
Чёрный конь	0.95
Белая ладья	0.93
Чёрная ладья	0.93
Белый слон	0.94
Чёрный слон	0.93
Белый король	0.95
Чёрный король	0.90
Белая королева	0.95
Чёрная королева	0.92

*Источник: Составлено авторами Source: compiled by the author.

Заключение (Conclusions)

Разработан и исследован метод автоматической идентификации объектов сложной формы и определения мест их захвата манипулятором РТК при условии их произвольного расположения на рабочем столе без взаимного перекрытия. Предложен алгоритм определения координат рационального места захвата на основе анализа двумерного изображения, полученного с использованием монокулярной камеры. Применение архитектуры YOLOv8-ОВВ позволило учитывать ориентацию объекта при определении его положения, что является необходимым условием для корректного расчёта координат захвата. Особенностью метода является учёт экспериментально определяемого смещения точки захвата относительно геометрического центра объекта, зависящего от его класса. Экспериментальная проверка метода подтвердила возможность устойчивого распознавания и определения координат захвата объектов сложной формы в реальном времени без использования дорогостоящих RGB-D датчиков.

Разработанный метод применим для объектов из заранее известного конечного набора. Объекты из набора обладают осевой симметрией (шахматные фигуры, болты и т.п.) и располагаются на рабочей плоскости в произвольной ориентации без взаимного перекрытия. Монокулярная камера жёстко закреплена относительно рабочей зоны, что соответствует условиям применения промышленных робототехнических комплексов.

При наличии перекрытий объектов или сложной трёхмерной структуры сцены точность определения рационального места захвата снижается. Поэтому расширение данного метода для новых условий требует разработки дополнительных методов анализа сцены. Предполагается проведение дальнейших исследований, которые будут направлены на решение задачи разбора навала и обработки частично перекрывающихся объектов.

Литература

1. Счетчиков Ю.А., Кайнова Т.Д., Чупаков М.В., Демидова К.А. Распознавание целей на базе технологий компьютерного зрения для беспилотных систем // Известия Тульского государственного университета. Технические науки. – 2023. – Вып. 7. – С 628–629. – DOI: 10.24412/2071-6168-2023-7-628-629.
2. Jocher G., Chaurasia A., Qiu J. YOLO by Ultralytics. Version 8.0.0. Ultralytics, 2023. URL: <https://docs.ultralytics.com>
3. Redmon J., Divvala S., Girshick R., Farhadi A. You Only Look Once: Unified, Real-Time Object Detection. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 779–788.
4. Казначеева А.А., Власенко О.М., Эпов А.А. Алгоритм управления мехатронной станцией сортировки изделий с применением системы компьютерного зрения // Инженерный вестник Дона. – 2025. – № 5.
5. Ершов М.С., Журавлев С.М. Применение искусственных нейронных сетей для мониторинга параметров устойчивости промышленной электротехнической системы. *Известия высших учебных заведений. ПРОБЛЕМЫ ЭНЕРГЕТИКИ*. 2025;27(5):41-52.
6. Политова Т.О., Загреддинов А.Р., Сидоров М.В., Зиганшин Ш.Г., Ваньков Ю.В. Методы машинного обучения в разработке методики обнаружения утечек. *Известия высших учебных заведений. ПРОБЛЕМЫ ЭНЕРГЕТИКИ*. 2025;27(2):177-186.
7. Liu W., Anguelov D., Erhan D. et al. SSD: Single Shot MultiBox Detector // *ECCV*. – 2016. – pp. 21–37.

8. Redmon J., Farhadi A. YOLOv3: An Incremental Improvement. arXiv preprint arXiv:1804.02767, 2018.
9. Bochkovskiy A., Wang C.-Y., Liao H.-Y.M. YOLOv4: Optimal Speed and Accuracy of Object Detection. arXiv:2004.10934.
10. Redmon J., Farhadi A. YOLO9000: Better, Faster, Stronger // CVPR. – 2017.
11. Wang C.-Y., Bochkovskiy A., Liao H.-Y.M. YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors // CVPR. – 2023. – pp. 7464–7475.
12. Боков П.А., Кравченя П.Д. Экспериментальный анализ точности и производительности разновидностей архитектур YOLO для задач компьютерного зрения // Программные продукты и системы. – 2020. – №4. – С. 635–640.
13. Небаба С.Г., Марков Н.Г. Сверточные нейронные сети семейства YOLO для мобильных систем компьютерного зрения // Компьютерные исследования и моделирование. – 2024. – №3. – С. 615–631. Zhang H., Lan X., Zhou X. RGB-D Based Robotic Grasp Detection: A Survey. *Sensors*, Vol. 22, No. 15, 2022, p. 5634.
14. Гладков А.Н., Быстров Л.Ю., Кузьмин Е.В. Применение нейронных сетей семейства YOLO для обнаружения полезных сигналов на вихретоковых дефектограммах рельсов // Труды Института системного программирования РАН. – 2025. – Т.37. – №6. – С. 131–150. DOI: 10.15514/ISPRAS-2025-37(6)-25.
15. Шилин А.Н., Богале М.А., Коновалова Л.А. Модель искусственной нейросети для прогнозирования объема воды в водохранилище // Известия высших учебных заведений. Проблемы энергетики. – 2024. – Т.26. – №5. – С. 104–117.
16. Zhang H., Lan X., Zhou X. RGB-D Based Robotic Grasp Detection: A Survey // *Sensors*. – 2022. – Vol.22. – No.15. – p.5634.
17. Ma Y., Zhu W., Zhou Y. Automatic grasping control of mobile robot based on monocular vision // *Int. Journal of Advanced Manufacturing Technology*. – 2022. – Vol.121. – pp.1785–1798.
18. Андреев В.П. Определение формы препятствия сенсорной системой мобильного робота с помощью ИК-датчиков расстояния типа GP2Y0A (Sharp) // Известия высших учебных заведений. Проблемы энергетики. – 2024. – Т.26. – №1. – С.195–207.
19. Рублева Е.А., Мудрич А.Б. Определение положения рабочего инструмента робота-манипулятора с помощью камеры и машинного зрения // Международный научно-исследовательский журнал. – 2024.
20. Mahler J. et al. Dex-Net 2.0: Deep Learning to Plan Robust Grasps with Synthetic Point Clouds and Analytic Grasp Metrics // *Robotics: Science and Systems*. – 2017.
21. Поливанов А.Ю., Дабул Ю. Методика определения координат сварочных стыков в системе технического зрения робота для дуговой сварки // Вестник МГТУ «Станкин». – 2025. – №1(72). – С.16–25.
22. Поливанов А.Ю., Иванов Ю.В. Калибровка видеосенсора системы технического зрения промышленного робота для лазерной сварки // Вестник МГТУ «Станкин». – 2019. – №49, – с. 119-126.

Авторы публикации

Андреев Виктор Павлович – д-р техн. наук, профессор кафедры «Сенсорные и управляющие системы» Московского государственного технологического университета «СТАНКИН».

Карнов Максим Вячеславович – студент (магистрант) Московского государственного технологического университета «СТАНКИН».

References

1. Schetcikov Yu.A., Kainova T.D., Chupakov M.V., Demidova K.A. Target recognition based on computer vision technologies for unmanned systems. *Proceedings of Tula State University. Technical Sciences*. 2023, Issue 7, pp. 628–629. DOI: 10.24412/2071-6168-2023-7-628-629.
2. Jocher G., Chaurasia A., Qiu J. YOLO by Ultralytics. Version 8.0.0. Ultralytics, 2023. Available at: <https://docs.ultralytics.com>
3. Redmon J., Divvala S., Girshick R., Farhadi A. You Only Look Once: Unified, Real-Time Object Detection. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 779–788.
4. Kaznacheeva A.A., Vlasenko O.M., Epov A.A. Control algorithm for a mechatronic product sorting station using a computer vision system. *Engineering Journal of Don*. 2025, No. 5.
5. Ershov M.S., Zhuravlev S.M. Application of artificial neural networks for monitoring

stability parameters of industrial electrical systems. News of Higher Educational Institutions. Energy Sector Problems. 2025, Vol. 27, No. 5, pp. 41–52.

6. Politova T.O., Zagretidinov A.R., Sidorov M.V., Ziganshin Sh.G., Vankov Yu.V. Machine learning methods in the development of leak detection techniques. News of Higher Educational Institutions. Energy Sector Problems. 2025, Vol. 27, No. 2, pp. 177–186.

7. Liu W., Anguelov D., Erhan D., Szegedy C., Reed S., Fu C.-Y., Berg A.C. SSD: Single Shot MultiBox Detector. European Conference on Computer Vision (ECCV), 2016, pp. 21–37.

8. Redmon J., Farhadi A. YOLOv3: An Incremental Improvement. arXiv preprint arXiv:1804.02767.

9. Bochkovskiy A., Wang C.-Y., Liao H.-Y.M. YOLOv4: Optimal Speed and Accuracy of Object Detection. arXiv preprint arXiv:2004.10934.

10. Redmon J., Farhadi A. YOLO9000: Better, Faster, Stronger. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

11. Wang C.-Y., Bochkovskiy A., Liao H.-Y.M. YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023, pp. 7464–7475.

12. Bokov P.A., Kravchenya P.D. Experimental analysis of accuracy and performance of YOLO architectures for computer vision tasks. Software & Systems. 2020, No. 4, pp. 635–640.

13. Nebaba S.G., Markov N.G. Convolutional neural networks of the YOLO family for mobile computer vision systems. Computer Research and Modeling. 2024, No. 3, pp. 615–631.

14. Gladkov A.N., Bystrov L.Yu., Kuzmin E.V. Application of YOLO neural networks for detection of useful signals in eddy-current rail defectograms. Proceedings of the Institute for System Programming of the Russian Academy of Sciences. 2025, Vol. 37, No. 6, pp. 131–150. DOI: 10.15514/ISPRAS-2025-37(6)-25.

15. Shilin A.N., Bogale M.A., Konovalova L.A. Artificial neural network model for forecasting water volume in a reservoir. News of Higher Educational Institutions. Energy Sector Problems. 2024, Vol. 26, No. 5, pp. 104–117.

16. Zhang H., Lan X., Zhou X. RGB-D Based Robotic Grasp Detection: A Survey. Sensors. 2022, Vol. 22, No. 15, p. 5634.

17. Ma Y., Zhu W., Zhou Y. Automatic grasping control of a mobile robot based on monocular vision. The International Journal of Advanced Manufacturing Technology. 2022, Vol. 121, pp. 1785–1798. DOI: 10.1007/s00170-022-09438-z.

18. Andreev V.P. Determining the shape of an obstacle using a mobile robot sensor system based on GP2Y0A infrared distance sensors (Sharp). News of Higher Educational Institutions. Energy Sector Problems. 2024, Vol. 26, No. 1, pp. 195–207.

19. Rubleva E.A., Mudrich A.B. Determining the position of a robot manipulator tool using a camera and machine vision. International Research Journal. 2024.

20. Mahler J., Liang X., Niyaz S., Peters D., Rich A., Lee J., Gonzalez A., Goldberg K. Dex-Net 2.0: Deep Learning to Plan Robust Grasps with Synthetic Point Clouds and Analytic Grasp Metrics. Robotics: Science and Systems (RSS), 2017.

21. Polivanov A.Yu., Dabul Yu. Method for determining coordinates of welding joints in a robot vision system for arc welding. STANKIN University Bulletin. 2025, No.1 (72), pp. 16–25.

22. Polivanov A.Yu., Ivanov Y.V. Calibration of the video sensor of the technical vision system of an industrial robot for laser welding. Messenger of MSUT “STANKIN”. 2019. No.49, pp.119-126.

Authors of the publication

Victor P. Andreev – Moscow State University of Technology "STANKIN", Moscow, Russia.

Maksim V. Karpov – Moscow State University of Technology "STANKIN", Moscow, Russia.

Шифр научной специальности: 2.5.4. Роботы, мехатроника и робототехнические системы

Получено 18.04.2026 г.

Отредактировано 21.04.2026 г.

Принято 29.04.2026 г.