

АНАЛИТИЧЕСКИЙ ОБЗОР АРХИТЕКТУРНЫХ РЕШЕНИЙ РЕЗУЛЬТАТИВНОСТИ РОБОТИЗИРОВАННЫХ МАНИПУЛЯТОРОВ С ИНТЕГРАЦИЕЙ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА И ВИЗУАЛЬНОГО РАСПОЗНОВАНИЯ ОБЪЕКТОВ

Плеханов Тимофей Михайлович

**Студент 2 курса факультета государственного управления,
магистратура (дневное отделение) Московского государственного
университета имени М. В. Ломоносова**

В настоящей работе представлен аналитический обзор структурных и алгоритмических основ функционирования автоматизированных манипуляционных систем, оснащённых модулями искусственного интеллекта (далее — ИИ) и средствами компьютерного зрения. Детально описан замкнутый цикл управления, включающий этапы восприятия сцены, семантического анализа изображения глубокими свёрточными нейронными сетями и генерации управляющих воздействий на исполнительные механизмы. Рассмотрена трехуровневая архитектура «сенсор — нейросетевое вычислительное ядро — исполнительное устройство». Приведены количественные данные о состоянии мирового рынка промышленной робототехники за 2024— 2025 гг., а также верифицированные технические показатели эффективности систем визуального распознавания. На основе анализа научных публикаций и отраслевых отчётов сделан вывод о ключевой роли интеграции ИИ и компьютерного зрения в реализации концепции «Индустрия 4.0» (Industry 4.0 — четвёртая промышленная революция).

Ключевые слова: нейросетевое зрение, глубокие свёрточные сети, промышленный манипулятор, адаптивная роботизация, технико-экономическая эффективность, умное производство.

Современный этап цифровой трансформации промышленного производства характеризуется активным внедрением киберфизических систем, объединяющих физические исполнительные механизмы с вычислительными платформами на основе алгоритмов искусственного интеллекта [1]. Количественным свидетельством этого процесса служат данные Международной федерации робототехники (International Federation of Robotics, IFR): в 2024 году глобальный объём установок промышленных роботов достиг 542 000 единиц, а общий эксплуатационный парк составил 4 664 000 единиц при годовом росте в 9 % [4, 5]. Одним из наиболее динамично развивающихся направлений является создание роботизированных манипуляторов, способных автономно воспринимать производственную среду посредством оптических сенсоров, идентифицировать объекты по визуальным признакам и адаптировать пространственные траектории в реальном масштабе времени [2].

1. Принцип работы и техническая реализация системы «зрение — управление»

Функционирование современного манипулятора с интегрированным модулем AI— зрения (от англ. Artificial Intelligence — искусственный интеллект) может быть представлено в виде замкнутого кибернетического цикла, состоящего из трёх последовательных фаз: восприятие сцены, семантический анализ и генерация управляющего воздействия.

1.1. Восприятие сцены (сенсорный уровень). Входным информационным потоком служат данные, поступающие с оптических

сенсоров. В отличие от классических систем технического зрения с жёстко фиксированной над конвейером камерой, в передовых разработках применяется технология On-Arm Vision (от англ. — «зрение на руке») — установка 3D-камеры непосредственно на подвижном звене манипулятора [6]. Такой подход обеспечивает непрерывное обновление информации о рабочей сцене в процессе движения и позволяет компенсировать механические вибрации. Для построения трёхмерной карты глубины используются методы структурированного подсвета либо стереозрения, формирующие облако точек с плотностью, достаточной для вычисления пространственной ориентации объекта.

1.2. Семантический анализ (нейросетевое вычислительное ядро). Полученное изображение или облако точек передаётся на вычислительный модуль, реализующий алгоритмы глубокого обучения. В задачах промышленной робототехники, критичных ко времени отклика, предпочтение отдаётся одноэтапным детекторам семейства YOLO (от англ. You Only Look Once — «ты смотришь лишь однажды»). В частности, в работе [3] использована облегчённая архитектура YOLOv5s (You Only Look Once version 5 small — пятая версия алгоритма YOLO, облегчённый вариант), выполняющая прямой проход (англ. inference) за временной интервал порядка нескольких миллисекунд при работе на промышленных контроллерах. Нейросетевая модель не только осуществляет классификацию объекта, но и вычисляет координаты ограничивающей рамки (англ. bounding box), а в совокупности с данными о глубине — определяет углы Эйлера для точного позиционирования схвата. При необходимости достижения субмиллиметровой точности (например, в системе CapSen Robotics PIC 2.0 — программно-аппаратной платформе 3D-зрения компании CapSen Robotics, США) точность позиционирования составляет $\pm 0,10$ мм, и применяется дополнительная обработка облака

точек с совмещением цветовой сегментации и данных о кривизне поверхности [7].

1.3. Планирование траектории и управление (исполнительный уровень). На основе вычисленных нейросетью пространственных координат модуль управления (в референсной реализации [3] — программируемый логический контроллер Siemens S7-1200) решает обратную задачу кинематики, формируя массив углов поворота сервоприводов. В сценариях bin picking (от англ. — захват деталей из хаотично насыпанной тары) дополнительно активируются алгоритмы планирования бесколлизонной траектории, например RRT (от англ. Rapidly-exploring Random Tree — быстроисследующее случайное дерево), и реализуется принцип «видения в руке»: после выполнения захвата производится повторное сканирование сцены для верификации правильности позиционирования и внесения корректирующих поправок в реальном времени [6]. Синхронизация между вычислителем и исполнительными устройствами осуществляется по промышленным протоколам реального времени, в частности Profinet (от англ. Process Field Network), гарантирующим детерминированную задержку передачи данных [3].

2. Архитектурная организация интеллектуальных роботизированных систем.

С точки зрения структурной организации описанный выше функциональный цикл реализуется в рамках классической трехуровневой архитектуры, включающей сенсорную подсистему, вычислительное ядро и исполнительные механизмы [1]. Такая декомпозиция обеспечивает возможность независимой модернизации каждого из уровней и масштабирования системы. Современные тенденции предполагают

дополнение архитектуры уровнем цифрового двойника (англ. digital twin), что позволяет осуществлять предиктивное обслуживание и адаптивное управление в симулированной среде без остановки физического оборудования [8, 9].

3. Режимы функционирования и типовые производственные сценарии

В зависимости от степени априорной неопределённости рабочей среды выделяют два основных режима работы интеллектуальных манипуляторов [2].

3.1. Режим статической детекции. Применяется в условиях, когда объекты перемещаются по конвейеру с известной скоростью, однако их точное положение и ориентация варьируются в заданных пределах. Система однократно захватывает кадр, вычисляет пространственное смещение и передаёт скорректированные координаты на исполнительный уровень [3].

3.2. Режим динамического сопровождения (bin picking). Характерен для задач выборки деталей из хаотично насыпанной тары (bin picking). В данном сценарии необходима непрерывная коррекция траектории в условиях изменяющейся геометрии сцены. Согласно данным компании Inbolt, реализация технологии On-Arm Vision («зрение на руке») в задачах bin picking позволяет достичь 95 % успешных захватов при времени цикла менее 1 секунды в реальных производственных условиях [6].

4. Обзор верифицированных научных исследований и промышленных внедрений

4.1. Полностью автоматизированная система роботизированной укладки. В исследовании, опубликованном в журнале *Sensors* (MDPI, 2025), представлена система на основе глубокого обучения, интегрирующая алгоритм YOLOv5s, промышленный контроллер Siemens S7-1200 и манипулятор ABB IRB6640 [3]. Система обучена классификации 13 типов промышленных мешков. Валидационные испытания показали точность классификации 99,08 % на начальном этапе и 100 % после оптимизации условий освещения и угла съёмки камеры при эксплуатации на выборке из 9600 единиц продукции в течение пяти дней. Среднее время цикла составило 10–11 секунд, что соответствует производительности до шести единиц в минуту.

4.2. Технология bin picking компании Inbolt. В январе 2026 года компания Inbolt анонсировала решение для неструктурированного bin picking на основе бортового ИИ-зрения, функционирующее на аппаратной платформе NVIDIA [6]. Система использует собственные модели наведения, обеспечивающие оценку положения объекта в реальном времени и непрерывную коррекцию траектории. Ключевые заявленные характеристики включают время обработки менее 1 секунды на захват и надёжность до 95 % успешных попыток.

4.3. Высокоточное позиционирование. Платформа CapSen Robotics PIC 2.0 демонстрирует достижимые в настоящее время показатели точности 3D-зрения: распознавание объектов с погрешностью не более $\pm 0,10$ мм даже в условиях нагромождения деталей [7]. Обработка данных и планирование движений выполняются менее чем за 200 мс, что потенциально позволяет достигать скорости до 60 объектов в минуту.

5. Анализ текущего состояния рынка промышленной робототехники

Согласно отчёту World Robotics 2025, опубликованному IFR (Международной федерацией робототехники), в 2024 году глобальные установки промышленных роботов достигли 542 000 единиц, что является вторым по величине показателем за всю историю наблюдений [4]. Азия сохраняет доминирующее положение: на Китай пришлось 295 000 установок (54 % мирового объёма), Япония установила 44 500 роботов, Южная Корея — 30 600. В Европе зафиксировано 85 000 новых установок, в странах Америки — 50 077. Ведущим сектором автоматизации выступила электронная промышленность (128 899 установок), за ней следуют автомобилестроение и металлообработка.

Ключевыми трендами 2025—2026 гг., по данным IFR, являются: переход от узкоспециализированных роботов к многоцелевым системам, оснащённым ИИ; расширение применения коллаборативных роботов (коботов); интеграция роботов в «умные» подключённые фабрики с возможностью управления в реальном времени [5].

6. Перспективные направления развития

Основной вектор эволюции — переход от аналитического ИИ к физическому ИИ (Physical AI — направление, объединяющее ИИ с робототехникой и симуляцией физических процессов), при котором роботы обучаются в виртуальных средах на цифровых двойниках и переносят приобретённые навыки в физический мир [8, 9]. Технология цифровых двойников обеспечивает безопасную среду тестирования алгоритмов, минимизирует риски повреждения оборудования и сокращает время вывода новых решений на рынок. Другим значимым направлением является использование генеративного ИИ для синтеза обучающих

данных, что позволяет обучать нейросетевые модели без ручной разметки тысяч реальных изображений. Наконец, наблюдается тенденция к унификации AI-модулей, что должно снизить порог входа для предприятий малого и среднего бизнеса [5].

Проведённый анализ свидетельствует о том, что интеллектуальные роботизированные манипуляторы с AI-зрением перешли из категории экспериментальных разработок в разряд зрелых промышленных технологий. Верифицированные результаты научных исследований и промышленных внедрений подтверждают возможность достижения 100 % точности классификации объектов и производительности до шести единиц в минуту в реальных производственных условиях [3]. Технические характеристики решений компаний Inbolt [6] и CapSen Robotics [7] дополнительно иллюстрируют высокий уровень технологической готовности. Интеграция методов искусственного интеллекта и компьютерного зрения в робототехнику является ключевым драйвером реализации концепции «Индустрия 4.0» (Industry 4.0) и формирования умных производственных систем будущего [1, 5].

СПИСОК ЛИТЕРАТУРЫ

1. Kumar A., et al. Advances in intelligent industrial manipulators for smart manufacturing and standardized automation technologies // Discover Robotics. Springer, 2025. Vol. 1, Article No. 12. DOI: 10.1007/s44430-025-00012-2.
2. Industrial Robotic Setups: Tools and Technologies for Tracking and Analysis in Industrial Processes // Applied Sciences. MDPI, 2025. Vol. 15, No. 18, 10249. DOI: 10.3390/app151810249.
3. Özer A.S., et al. Real-Time and Fully Automated Robotic Stacking System with Deep Learning-Based Visual Perception // Sensors. MDPI, 2025. Vol. 25, No. 22, 6960. DOI: 10.3390/s25226960.
4. Global Robot Demand in Factories Has Doubled Over the Past 10 Years // IEN Europe, 2025. URL: <https://www.iен.eu/article/global-robot-demand-in-factories-has-doubled-over-the-past-10-years/>.
5. Industrial robotics in 2025: trends, figures and global outlook // Roboticstomorrow, 2025. URL: <https://www.roboticstomorrow.com/news/2025/12/18/industrial-robotics-in-2025-trends-figures-and-global-outlook/25933>.
6. Inbolt Unveils Human-Like Bin Picking Solution Powered by On-Arm AI Vision // Metrology and Quality News, 2026. URL: <https://metrology.news/inbolt-unveils-human-like-bin-picking-solution-powered-by-on-arm-ai-vision/>.
7. CapSen Robotics review: The pros and cons // Standard Bots, 2025. URL: <https://standardbots.com/blog/capsen-robotics>.
8. Digital twin-enabled real-time control for robot arm-based manufacturing via reinforcement learning // NSF PAR, 2025. URL: <https://par.nsf.gov/biblio/10647078/media/xml>.

9. DTMA: visual object inspection and mechanism for digital twin with robotic arm // Signal, Image and Video Processing. Springer, 2025. Vol. 19, Article No. 852. DOI: 10.1007/s11760-025-04134-3.