

Лушников Анатолий, магистрант, Московский Государственный Технический Университет им. Н. Э. Баумана, г. Москва

ОПИСАНИЕ ПРОТОТИПА СИСТЕМЫ ОБНАРУЖЕНИЯ ОБЪЕКТОВ В РЕЖИМЕ РЕАЛЬНОГО ВРЕМЕНИ С ИСПОЛЬЗОВАНИЕМ МОДЕЛИ МАШИННОГО ОБУЧЕНИЯ YOLO (YOU ONLY LOOK ONCE)

Аннотация

В статье представлено описание прототипа системы обнаружения объектов в режиме реального времени на основе модели машинного обучения YOLO (You Only Look Once). Рассмотрены особенности применения методов компьютерного зрения и машинного обучения для решения задач автоматического анализа визуальной информации. Обоснован выбор семейства моделей YOLO, обладающих высокой скоростью обработки данных и обеспечивающих возможность использования в системах реального времени.

Предложенный прототип реализован на языке программирования Python с использованием библиотеки OpenCV для работы с веб-камерой и получения видеопотока. Система осуществляет последовательную обработку кадров, выполняет обнаружение и классификацию объектов, а также отображает пользователю результаты распознавания в виде названия обнаруженного объекта и значения вероятности классификации. Описаны основные сценарии взаимодействия пользователя с системой, включающие этапы сканирования объекта, его обнаружения и просмотра результатов распознавания. Описанный прототип демонстрирует возможность эффективного применения моделей YOLO для создания систем обнаружения объектов, функционирующих в режиме реального времени.

Abstract

The article describes a prototype of a real-time object detection system based on the YOLO (You Only Look Once) machine learning model. The features of the application of computer vision and machine learning methods for solving problems of automatic analysis of visual information are considered. The choice of the YOLO

family of models with high data processing speed and providing the possibility of use in real-time systems is justified.

The proposed prototype is implemented in the Python programming language using the OpenCV library for working with a webcam and receiving a video stream. The system performs sequential frame processing, performs object detection and classification, and displays recognition results to the user in the form of the name of the detected object and the classification probability value. The main scenarios of user interaction with the system are described, including the stages of scanning an object, detecting it, and viewing recognition results. The described prototype demonstrates the possibility of effectively using YOLO models to create real-time object detection systems.

Ключевые слова: компьютерное зрение, машинное обучение, обнаружение объектов, YOLO, OpenCV, Python, обработка видеопотока, распознавание объектов.

Keywords: computer vision, machine learning, object detection, YOLO, OpenCV, Python, video stream processing, object recognition.

С развитием вычислительной техники и технологий искусственного интеллекта всё большую актуальность приобретают задачи автоматического анализа визуальной информации. Одним из наиболее перспективных направлений в данной области является компьютерное зрение, которое позволяет вычислительным системам воспринимать окружающую среду, анализировать изображения и видеопотоки, а также принимать решения на основе полученных данных. В настоящее время технологии компьютерного зрения активно применяются в системах видеонаблюдения, робототехнике,

автономном транспорте, медицине, промышленности и мобильных приложениях.

Одним из ключевых подходов, используемых для решения задач обнаружения объектов, является машинное обучение. В отличие от традиционных алгоритмов, которые основаны на заранее определённых правилах, методы машинного обучения позволяют системе самостоятельно выявлять закономерности на основе обучающих данных и использовать полученные знания для обработки новой информации [6]. Для этого используемые модели проходят этап обучения на наборах данных, благодаря чему появляется возможность распознавания различных категорий объектов с более высокой точностью.

Одной из фундаментальных задач компьютерного зрения является обнаружение объектов, под которым понимается не только определение класса объекта на изображении или видеокadre, но и установление его местоположения [2, 3]. В последние годы данная задача вызывает значительный интерес благодаря применению обработки визуальной информации в режиме реального времени.

Среди большого множества современных алгоритмов обнаружения объектов особое место занимает семейство моделей YOLO (You Only Look Once). Основным преимуществом данного подхода является одноэтапная архитектура, при которой локализация и классификация объектов выполняются за один проход нейронной сети [1, 4]. Благодаря этому модели YOLO обеспечивают высокую скорость обработки данных при сохранении приемлемой точности обнаружения, что делает их эффективным решением для систем реального времени. Несмотря на наличие определённых ограничений, которые связаны со снижением точности локализации в отдельных сценариях, алгоритмы YOLO получили широкое распространение благодаря удачному балансу между скоростью и качеством работы [1].

Для получения информации об объектах окружающей среды будет использована система обнаружения объектов в режиме реального времени.

Система реализована на языке программирования Python с использованием библиотеки OpenCV для взаимодействия с веб-камерой пользователя, получения видеопотока и отображения кадров пользователю [5]. После запуска приложения на экране будет отображаться изображение с камеры, а пользователь сможет направить её на интересующий объект для его дальнейшего анализа.

В процессе работы системы отдельные кадры видеопотока будут последовательно извлекаться и далее передаваться на обработку. После получения очередного кадра запускается этап обнаружения объектов, реализованный на основе модели машинного обучения YOLO.

По результатам обработки система определяет наличие объектов на изображении и принадлежность к определённым классам. И в случае успешного обнаружения пользователю отображается информация о распознанном объекте, включая его название и значение вероятности классификации. Если объекты на текущем кадре не обнаружены, то система не прекращает работу, а продолжает анализ последующих кадров видеопотока до момента успешного распознавания объекта.

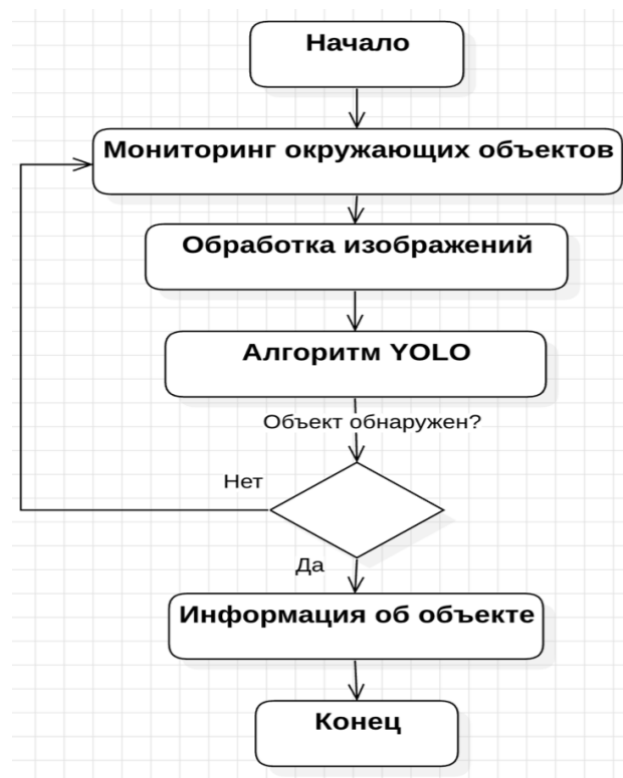


Рисунок 1. Алгоритм работы системы обнаружения объектов с использованием модели машинного обучения YOLO

Основные сценарии взаимодействия пользователя с системой обнаружения объектов в режиме реального времени включают этапы сканирования, обнаружения и просмотра результатов обнаружения объектов.

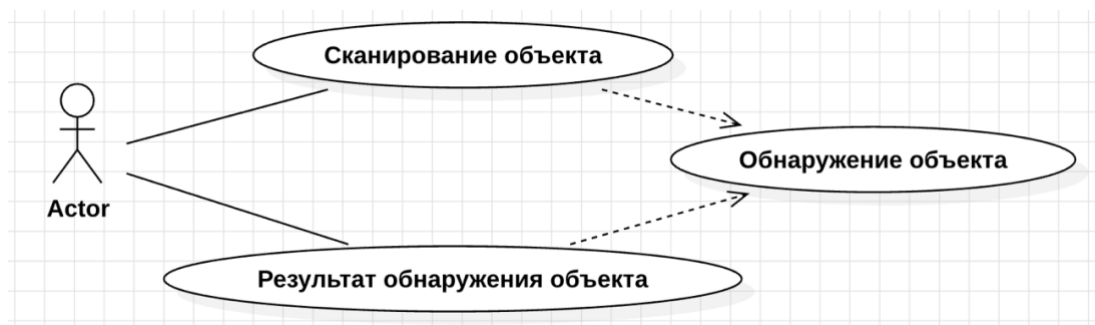


Рисунок 2. Диаграмма взаимодействия пользователя с системой

На этапе сканирования объекта пользователь направляет веб-камеру на интересующий его объект. Система получает видеопоток с камеры и извлекает кадры, которые содержат изображение окружающей среды пользователя. Данный вариант использования является обязательным условием для выполнения последующих этапов обработки.

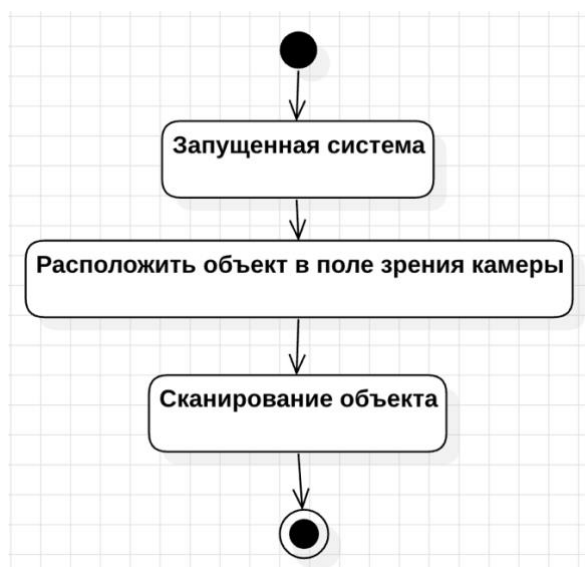


Рисунок 3. Алгоритм работы этапа сканирования объекта

Этап обнаружения объекта является основным вариантом использования системы. После получения кадра видеопотока запускается процесс его анализа с применением модели машинного обучения YOLO. В ходе обработки система

определяет наличие объектов на изображении, устанавливает принадлежность объектов к определённым классам и вычисляет уровень уверенности предсказания. Данный процесс выполняется автоматически и не требует дополнительных действий со стороны пользователя.



Рисунок 4. Алгоритм работы этапа обнаружения

Просмотр результатов обнаружения объектов представляет собой завершающий этап взаимодействия пользователя с системой. После успешного распознавания на экране отображается информация об обнаруженном объекте, включая его наименование и процент точности классификации. Как было описано ранее, если на текущем кадре объект не был обнаружен, то система продолжает анализировать последующие кадры видеопотока до получения результата.



Рисунок 5. Алгоритм работы этапа просмотра результатов обнаружения объектов

Литература

1. Васильев М. Е., Шалимов А. С., Савина О. А. Обзор версий YOLO: одноэтапная модель сверточной нейронной сети / М. Е. Васильев, А. С. Шалимов, О. А. Савина // Universum: технические науки. – 2025. – № 6 (135).
2. Зуева Е. Ю. Компьютерное зрение в ИПМ им. В.М. Келдыша РАН – история развития [Электронный ресурс] // Келдыш.ру. –

URL: http://www.keldysh.ru/papers/2009/art04/Zueva_09.htm (дата обращения: 24.05.2026).

3. Компьютерное (машинное) зрение (computer vision) [Электронный ресурс] // ES-prof.com. – URL: http://es-prof.com/m_mvision.php (дата обращения: 29.05.2026).

4. Маралов А. А., Серимбетов Б. А. Современные методы распознавания образов: подходы, алгоритмы и реализация в информационных системах / А. А. Маралов, Б. А. Серимбетов // Вестник науки. – 2025. – № 2 (83).

5. Синевич Е. В. Программное обеспечение для распознавания образов [Электронный ресурс] // БГУИР. – URL: https://libeldoc.bsuir.by/bitstream/123456789/58143/1/Sinevich_Programmnoe.pdf (дата обращения: 27.05.2026).

6. Флах П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных / П. Флах. – М. : ДМК Пресс, 2023. – 401 с.

Literature

1. Vasilyev M. E., Shalimov A. S., Savina O. A. A Review of YOLO Versions: A One-Stage Convolutional Neural Network Model / M. E. Vasilyev, A. S. Shalimov, O. A. Savina // Universum: Technical Sciences. – 2025. – No. 6 (135).

2. Zueva E. Yu. Computer Vision at the Keldysh Institute of Applied Mathematics of the Russian Academy of Sciences – History of Development [Electronic resource] // Keldysh.ru. –

URL: http://www.keldysh.ru/papers/2009/art04/Zueva_09.htm (accessed:24.05.2026)

. 3. Computer (Machine) Vision [Electronic resource] // ES-prof.com. – URL: http://es-prof.com/m_mvvision.php (accessed: 29.05.2026).

4. Maralov A. A., Serimbetov B. A. Modern Methods of Pattern Recognition: Approaches, Algorithms and Implementation in Information Systems / A. A. Maralov, B. A. Serimbetov // Bulletin of Science. – 2025. – No. 2 (83).

5. Sinevich E. V. Software for Pattern Recognition [Electronic resource] // BSUIR–

URL: https://libeldoc.bsuir.by/bitstream/123456789/58143/1/Sinevich_Programmnoe.pdf (accessed: 27.05.2026).

6. Flach P. Machine Learning: The Art and Science of Algorithms that Make Sense of Data / P. Flach. – Moscow: DMK Press, 2023. – 401 p.