

## **СИСТЕМА ВИДЕОКОНТРОЛЯ СРЕДСТВ ИНДИВИДУАЛЬНОЙ ЗАЩИТЫ НА ПРОМЫШЛЕННОМ ПРЕДПРИЯТИИ**

**Аннотация.** В статье рассматриваются методы автоматизации контроля средств защиты на производстве с применением видеонаблюдения. Представлены подходы к распознаванию и мониторингу использования необходимых элементов, оценена эффективность системы. Предложена комплексная архитектура системы видеоконтроля средств защиты для промышленных предприятий.

**Ключевые слова:** видеоконтроль, средства индивидуальной защиты, компьютерное зрение, системы автоматизации.

**Введение.** Обеспечение безопасности на производственных предприятиях остается одним из важнейших направлений современной промышленной безопасности. Согласно статистике Международной организации труда, ежегодно в мире происходит около 340 миллионов несчастных случаев на производстве, причем значительная часть из них связана с несоблюдением правил использования средств индивидуальной защиты (СИЗ) [1]. Анализ данных производственного травматизма за 2020-2023 годы показывает, что около 27% несчастных случаев связаны с неиспользованием или неправильным использованием СИЗ [2]. При этом тяжесть последствий таких случаев в среднем на 43% выше, чем у происшествий, не связанных с нарушением требований к СИЗ [3].

Контроль за использованием СИЗ традиционно осуществляется посредством ручных проверок, что неизбежно приводит к низкой оперативности, субъективности оценок и высоким трудовым затратам. Эффективность таких проверок существенно ограничена человеческим фактором, особенно на крупных промышленных объектах с множеством производственных участков.

Практика показывает, что внедрение систем видеоконтроля способствует снижению количества нарушений использования СИЗ на 62% в течение первых трех месяцев после внедрения за счет постоянного мониторинга и своевременного реагирования на выявленные нарушения [4, 5]. Более того, время реагирования на нарушения сокращается до 30-60 секунд при использовании автоматизированных систем, что значительно снижает риск возникновения несчастных случаев.

Целью данного исследования является изучение и сравнительный анализ алгоритмов детекции объектов и модулей оптического распознавания текста (OCR) для создания единой системы видеоконтроля СИЗ, обеспечивающей высокую оперативность и точность мониторинга в условиях реального производства. Для достижений поставленной цели решаются следующие задачи:

1. Проведение обзора существующих алгоритмов компьютерного зрения для детекции СИЗ и оценка их применимости в условиях промышленного производства.
2. Изучение эффективности OCR систем для идентификации сотрудников по информации на одежде.
3. Проектирование целостной архитектуры системы видеоконтроля с учетом особенностей промышленного предприятия.

4. Модификация и оптимизация алгоритмов компьютерного зрения под специфику детекции средств защиты.

5. Оценка эффективности спроектированной системы в условиях схожих с условиями реального предприятия.

**Проблема исследования.** Существующие автоматизированные системы контроля СИЗ сталкиваются с рядом технических ограничений. Большинство коммерческих систем оценки рисков демонстрируют падение точности детекции на 25-40% при условиях, затрудняющих работу систем видеоаналитики: вибрации, экстремальные температуры, пыль, переменное освещение [6]. Все эти факторы требуют разработки специализированных решений, адаптированных к условиям промышленного производства.

Значительной проблемой также является масштабируемость систем видеонаблюдения. По мере увеличения количества камер возрастает нагрузка на вычислительные ресурсы и каналы передачи данных, что может привести к снижению производительности системы в целом. Для полного охвата требуется в среднем одна камера на 100 м<sup>2</sup> площади, что для среднего предприятия составляет от 30 до 80 камер. При использовании централизованной архитектуры обработки данных это составляет значительную нагрузку на каналы связи и серверное оборудование.

**Материалы и методы.** Для обучения и оценки моделей был сформирован специализированный набор данных, включающий 15 тыс. изображений размером 624 x 624 пикселя с более чем 20 тыс. объектами. Изображения были получены с использованием различных источников данных, включая свободно распространяемые наборы данных, такие как: “SH17” [7], “Hard Hat Dataset” [8] и другие, менее известные наборы данных. Для их поиска использовались специализированные платформы машинного обучения Roboflow, GitHub и Kaggle. Такой комплексный подход позволил собрать разнообразный материал, охватывающий различные производственные операции.

Собранный набор данных включает изображения, полученные с производственных видеокамер, охватывающие различные производственные помещения: цеха, складские помещения, лаборатории, открытые площадки. Многообразие сцен позволяет повысить устойчивость модели компьютерного зрения.

Аннотирование изображений выполнено с использованием предварительно размеченных наборов данных и инструментов разметки. Аннотации включают ограничивающие рамки для сотрудников и различных типов СИЗ: каски, защитные очки, респираторы, перчатки, защитная обувь, защитная одежда, наушники. Пример изображения из обучающего набора данных с разметкой представлен на рис. 1.

Особое внимание было уделено созданию негативных классов (no\_helmet, no\_glasses), которые представляют собой явное отсутствие необходимого СИЗ. Включение таких классов в набор данных дает следующие преимущества:

1. Непосредственное обнаружение нарушений — вместо логического вывода об отсутствии СИЗ, модель обучается явно распознавать случаи отсутствия требуемого средства защиты.

2. Снижение количества ложных срабатываний — модель учится различать ситуации, когда СИЗ действительно отсутствует, от ситуаций, когда оно просто не видно из-за ракурса или перекрытия.

3. Повышение точности в сложных сценариях — при наличии в обучающей выборке негативных примеров модель лучше обучается распознавать пограничные случаи.



Рис. 1. Пример изображения из собранного набора данных с ограничивающими рамками:

- 1 — рамки для сотрудников (оранжевые прямоугольники); 2 — рамки для каски (желтые элементы);  
3 — рамки для защитной обуви (синие элементы); 4 — рамки для жилета (фиолетовые элементы);  
5 — рамки для перчаток (зеленые элементы)

Набор данных был разделен на обучающую (70%) и тестовую (30%) выборки. Разделение осуществлялось с учетом баланса классов и условий съемки, чтобы обеспечить репрезентативность обучающей выборки и эффективность модели в различных сценариях.

Разработанная система представляет собой многоуровневое решение, интегрирующее аппаратные и программные компоненты для обеспечения непрерывного мониторинга СИЗ. Общая архитектура системы представлена на рис. 2.

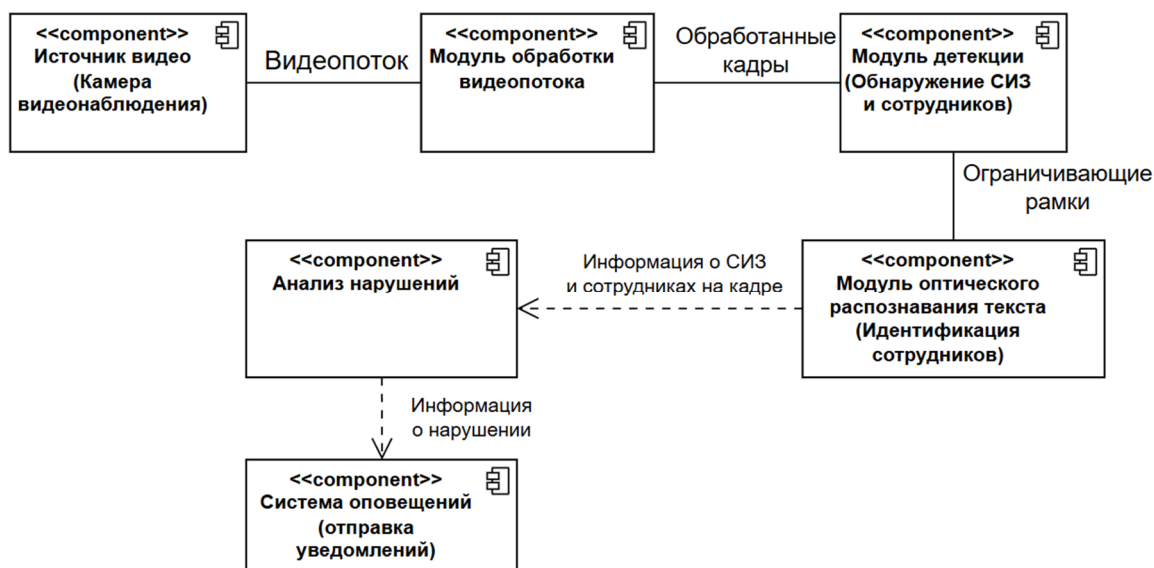


Рис. 2. Архитектура системы видеоконтроля СИЗ

Система включает следующие основные компоненты:

1. Уровень сбора данных — сеть камер видеонаблюдения, расположенных в важных точках предприятия, видеопоток с которых передается для анализа системой.
2. Модуль обработки видеопотока — получает видеопоток с камер видеонаблюдения, извлекает кадры и нормализует полученные изображения. На выходе модуля идут обработанные кадры, используемые для последующего анализа.
3. Модуль детекции — принимает кадры с предыдущего модуля и применяет к ним нейросетевые алгоритмы для обнаружения сотрудников и СИЗ. После обработки передает следующему модулю координаты ограничивающих рамок обнаруженных объектов, их классы и вероятность обнаружения.
4. Модуль идентификации — использует рамки обнаруженных сотрудников для распознавания идентификационных номеров сотрудников с использованием OCR-системы.
5. Уровень анализа нарушений — на основе переданных ему данных о сотрудниках и средствах защиты находящихся на них делает итоговый вывод о наличии или отсутствии нарушений.

Важным аспектом системы является адаптивность к условиям конкретного производства и цеха. Система позволяет настраивать правила контроля СИЗ для различных зон с учетом специфического набора средств защиты, которые обязательны именно для данного цеха.

Наиболее важным модулем разрабатываемой системы является модуль детекции, для того чтобы выбрать оптимальный алгоритм необходимо изучить и сравнить какие варианты существуют. В задаче обнаружения СИЗ можно выделить два основных подхода: одностадийные и двухстадийные детекторы. Одностадийные детекторы выполняют локализацию и классификацию объектов за один проход, что обеспечивает высокую скорость обработки при достаточной точности. Двухстадийные детекторы сначала генерируют множество потенциальных регионов интереса, а затем классифицируют и уточняют параметры ограничивающих рамок для каждого региона. Двухстадийные детекторы, как правило, обеспечивают более высокую точность, но требуют больше вычислительных ресурсов.

В контексте системы видеоконтроля СИЗ требуется обработка в режиме реального времени при сохранении достаточной точности выявления нарушений. Поэтому в данном исследовании рассматриваются как одностадийные, так и двухстадийные детекторы с целью определения оптимального баланса между скоростью и точностью. Среди одностадийных детекторов были выбраны YOLO (You Only Look Once) и SSD (Single Shot MultiBox Detection), а среди двухстадийных детекторов выбор остановился на Faster R-CNN.

Обучение проводилось на графических ускорителях NVIDIA 2060 с использованием фреймворка PyTorch. Для оптимизации применялся алгоритм SGD с моментом 0.9, кадры подавались группами по 16 изображений, что обусловлено ограничениями доступной видеопамяти. Обучение длилось в течении 50 эпох. В качестве функции потерь использовалась комбинация Focal Loss для классификации и GIoU Loss для регрессии координат рамок.

Focal Loss — это функция потерь, разработанная для улучшения точности классификации в условиях сильного дисбаланса классов. Преимущество Focal Loss в том, что она придает больший вес объектам, которые модель классифицирует с низкой уверенностью, что позволяет эффективно бороться с проблемой дисбаланса классов, возникающей в среде обнаружения СИЗ, где нарушения встречаются реже, чем случаи соблюдения правил.

GIoU Loss — это функция потерь для регрессии ограничивающих рамок, которая улучшает стандартный метод оценки пересечений над объединением путем учета не только площади пересечения предсказанной и истинной ограничивающих рамок, но и их общей охватывающей области, что позволяет лучше учитывать относительное положение рамок, улучшая точность работы.

Первым исследуемым алгоритмом обнаружения стал одностадийный YOLO. YOLO — это детектор, обеспечивающий высокую скорость обработки при хорошей точности [9]. Архитектура этого детектора включает три основных компонента. Основную сеть, построенную на базе CSPNet и основанную на архитектуре DarkNet. Это фундаментальная нейросеть, специально разработанная для семейства моделей YOLO, способная разделять входной слой на две части и обрабатывать их по отдельности, что позволяет снизить вычислительную сложность. Второй компонент — промежуточная модель Path Aggregate Network, которая эффективно агрегирует признаки с разных уровней сети. Третий компонент — классификатор, состоящий из нескольких параллельных сверточных слоев для предсказания координат рамок, а также вероятностей классов.

YOLOv11 представлен в нескольких вариантах с различной архитектурной сложностью. Результаты предварительных тестов на обучающих данных представлены в табл. 1.

Таблица 1

**Сравнительный анализ различных версий моделей YOLO**

| <i>Модель</i> | <i>mAP0.5</i> | <i>Precision</i> | <i>Recall</i> | <i>Размер модели</i> |
|---------------|---------------|------------------|---------------|----------------------|
| YOLOv11s      | 0.732         | 0.864            | 0.812         | 7.2                  |
| YOLOv11m      | 0.776         | 0.891            | 0.867         | 21.1                 |
| YOLOv11l      | 0.788         | 0.898            | 0.883         | 46.5                 |
| YOLOv11x      | 0.793         | 0.907            | 0.891         | 86.7                 |

В данном исследовании акцентировано внимание на вариантах YOLOv11s и YOLOv11m, которые обеспечивают наилучшее соотношение производительности и вычислительных затрат для систем реального времени.

Для адаптации модели к обнаружению средств индивидуальной защиты была модифицирована стандартная архитектура YOLOv11. Было увеличено количество слоев для лучшего обнаружения мелких объектов. Добавлены шаблоны для поиска объектов разной формы, соответствующие различным типам защитного оборудования. Также использована техника объединения информации для улучшения распознавания в условиях плохого освещения. На рис. 3 представлены графики изменения функций потерь для модели YOLOv11m в процессе обучения.

Левый график показывает изменение ошибки локализации, а правый — классификационную ошибку. Как видно из графиков, модель демонстрирует устойчивое снижение ошибки на протяжении всего процесса обучения, с выходом на плато примерно после 35-й эпохи, что свидетельствует о хорошей сходимости процесса обучения и отсутствии переобучения.

Следующий рассматриваемый алгоритм — SSD, который является еще одним представителем одностадийных детекторов [10]. Основной сетью SSD является VGG16 — глубокая

нейронная сеть, предварительно обученная на наборе данных ImageNet, с удалением полносвязных слоев. Эта база дополнена сверточными слоями для извлечения признаков разного масштаба. Завершает архитектуру классификатор использующий сверточные слои с ядром 3 x 3.

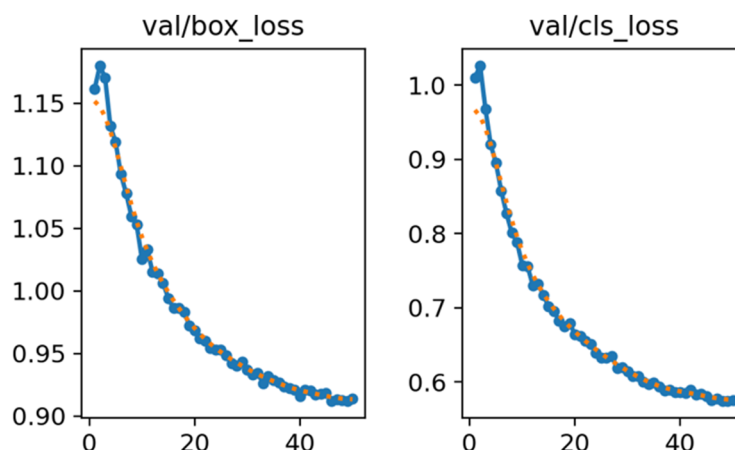


Рис. 3. График изменения функций ошибок со временем

SSD использует множество предопределенных ограничивающих рамок с различными соотношениями сторон и размерами для повышения точности детекции объектов различной формы. Особенностью SSD является использование многослойной пирамидальной структуры для детекции объектов разного размера.

Для адаптации SSD к обнаружению средств защиты были оптимизированы размеры и пропорции шаблонов поиска для лучшего соответствия различным типам защитного снаряжения. Добавлены дополнительные слои для улучшения обнаружения мелких объектов. Применена техника отбора сложных примеров с преобладанием негативных случаев в соотношении 3:1. Данная техника заключается в целенаправленном отборе таких областей изображения, которые модель ошибочно классифицировала как содержащие объекты.

В качестве представителя двухстадийных детекторов был выбран Faster R-CNN. Этот детектор включает несколько ключевых компонентов. Основная сеть построена на архитектуре глубокой нейросети с 50 слоями — ResNet-50 с пирамидальной структурой. Также модель включает модуль Region Proposal Network, который выделяет области-кандидаты, с высокой вероятностью содержащие объекты.

Особенностью Faster R-CNN является интеграция процесса генерации предложенных регионов в архитектуру нейронной сети. Это обеспечивает более высокую точность по сравнению с одностадийными детекторами, но требует больше времени для обработки одного изображения.

Для идентификации сотрудников была использована OCR-система EasyOCR. Архитектура EasyOCR включает в себя сверточную нейронную сеть для извлечения признаков, двуправленную для моделирования последовательностей и классификации для дальнейшего декодирования. Эта комбинация позволяет эффективно распознавать текст на изображениях даже в сложных условиях, с небольшими потерями точности.

В настоящем исследовании проблема масштабируемости системы решается посредством оптимизации алгоритмов для работы на граничных устройствах, что позволяет распределить

вычислительную нагрузку и уменьшить объем передаваемых данных. Предварительные эксперименты показали, что применение 8-битной квантизации к модели YOLOv11s позволяет снизить требования к вычислительным ресурсам на 72% при снижении точности детекции всего на 3.7%, что является приемлемым компромиссом для систем реального времени.

Для объективного сравнения эффективности различных моделей были использованы следующие метрики:

1. Mean Average Precision (mAP) — стандартная метрика для оценки точности детекторов объектов, рассчитывается при заданном пороге вероятности 0.5 и в диапазоне вероятности от 0.5 до 0.95.

2. Precision (точность) — доля правильно обнаруженных объектов среди всех обнаружений.

3. Recall (полнота) — доля правильно обнаруженных объектов среди всех объектов, присутствующих на изображении.

Результаты оценки производительности различных моделей детекции объектов на тестовой выборке представлены в табл. 2.

Таблица 2

Сравнительный анализ моделей детекции СИЗ

| Модель       | <i>mAP</i> 0.5 | <i>mAP</i> [.5:0.95] | <i>Precision</i> | <i>Recall</i> | <i>FPS</i> | Размер модели |
|--------------|----------------|----------------------|------------------|---------------|------------|---------------|
| YOLOv11s     | 0.732          | 0.578                | 0.864            | 0.812         | 62.3       | 28.7          |
| YOLOv11m     | 0.776          | 0.623                | 0.891            | 0.867         | 45.8       | 84.5          |
| SSD          | 0.721          | 0.543                | 0.835            | 0.798         | 32.1       | 96.3          |
| Faster R-CNN | 0.802          | 0.692                | 0.914            | 0.895         | 18.3       | 167.2         |

Как видно из табл. 2, двухстадийный детектор Faster R-CNN демонстрирует наиболее высокую точность, однако значительно уступает одностадийным детекторам в скорости обработки. В условиях промышленного видеоконтроля, когда требуется обработка видеопотока в реальном времени, скорость имеет критическое значение.

Среди одностадийных детекторов наиболее оптимальное соотношение точности и скорости демонстрирует модель YOLOv11m, обеспечивая *mAP*0.5 равное 0.776 при скорости обработки 46 кадров в секунду.

Таким образом, на основе проведенного исследования было принято решение использовать модифицированную YOLOv11m в качестве модуля детекции, а также EasyOCR для идентификации работников на предприятии. Для выявления нарушений был написан код для серверной части системы, с использованием библиотеки Flask, который анализирует данные полученные от других модулей и делает вывод о наличии или отсутствии СИЗ и оповещает инспекторов, ответственных за технику безопасности.

**Заключение.** В рамках исследования была разработана система для выявления нарушений, проведен сравнительный анализ алгоритмов детекции использования СИЗ. Экспериментальные результаты показывают, что применение современных моделей, таких как YOLO, позволяет существенно повысить точность обнаружения СИЗ в условиях переменного освещения и динамичных объектов, а OCR решения демонстрируют умеренную эффективность в

распознавании идентификационных знаков. Полученные данные подтверждают, что выбранные методы способны обеспечить надежную автоматизацию контроля, что существенно повышает уровень безопасности на производственных объектах.

Дальнейшее развитие разработанной системы возможно в ее интеграции с существующими системами управления безопасностью предприятия для создания единой экосистемы контроля. Также для повышения точности возможно дообучение на основе поступающих данных с производства, что позволит адаптировать систему к специфическим условиям конкретного предприятия.

## СПИСОК ЛИТЕРАТУРЫ

1. International Labour Organization. World Statistics on Occupational Safety and Health. — 2023. — 156 p.
2. Морозов В.Н., Сидоренко А.П. Анализ причин производственного травматизма в России и за рубежом // Охрана труда и техника безопасности. — 2023. — № 2. — С. 32-41.
3. Huang J., Wei L., Chen Y. Deep learning-based PPE detection systems: implementation and performance evaluation // IEEE Transactions on Industrial Informatics. — 2021. — Vol. 17, № 8. — P. 5532-5541.
4. Аналитический отчет по результатам аудитов промышленной безопасности в России 2022-2023 гг. — Москва : ВНИИ Охраны труда, 2023. — 78 с.
5. Li X., Zhang T., Wang R. Real-time monitoring of personal protective equipment using computer vision // Safety Science. — 2022. — Vol. 148. — P. 105622.
6. Смирнов А.В., Иванов К.П. Автоматизированные системы контроля безопасности труда на производстве // Вестник промышленной безопасности. — 2022. — № 3. — С. 45-52.
7. SH17 / GitHub. — URL: <https://github.com/ahmadmughees/SH17dataset> (дата обращения: 01.04.2025).
8. HARD HAT WORKERS / Public Roboflow. — URL: <https://public.roboflow.com/object-detection/hard-hat-workers> (дата обращения: 01.04.2025).
9. YOLOv11 / Ultralytics Docs. — URL: <https://docs.ultralytics.com/models/yolo11> (дата обращения: 01.04.2025).
10. Liu W., Anguelov D., Erhan D., Szegedy C., Reed S., Fu C.-Y., Berg A.C. SSD: Single Shot MultiBox Detector // Computer Vision — ECCV 2016: 14th European Conference on Computer Vision. — Amsterdam, The Netherlands, 2016. — С. 21-37. — DOI: 10.1007/978-3-319-46448-0\_2.