

*Стулишенко Михаил Алексеевич
магистрант МТУСИ, Москва, Россия*

*Мохаммад Навар
доцент кафедры ИСУиА МТУСИ, к.т.н., доцент, Москва, Россия*

РАЗРАБОТКА И ЭКСПЕРИМЕНТАЛЬНАЯ АПРОБАЦИЯ СИСТЕМЫ КОНТРОЛЯ СРЕДСТВ ИНДИВИДУАЛЬНОЙ ЗАЩИТЫ И ПРОИЗВОДСТВЕННОЙ БЕЗОПАСНОСТИ НА ОСНОВЕ КОМПЬЮТЕРНОГО ЗЕРНИЯ

Аннотация

В статье представлена разработанная система автоматического контроля соблюдения правил ношения средств индивидуальной защиты и нахождения персонала в опасных зонах на производственных объектах. Система построена на модульной архитектуре и использует нейросетевую модель YOLOv11m для детекции объектов в реальном времени. Для повышения точности распознавания мелких элементов СИЗ применяется тайловая обработка на базе библиотеки SAHI. Реализованы механизмы пространственной ассоциации обнаруженных СИЗ с конкретными работниками, двухрежимное формирование опасных зон (ручное и автоматическое на основе детекции техники), а также темпоральная фильтрация нарушений для снижения ложных срабатываний. Приводятся результаты экспериментального тестирования системы на различных сценариях, включая работу на локальном компьютере и серверном оборудовании. Метрики качества обученной модели ($mAP@0.5 = 76.4\%$, Recall по классам нарушений до 91.2%) подтверждают эффективность предложенного подхода для задач промышленной безопасности.

Ключевые слова: *компьютерное зрение, YOLOv11, средства индивидуальной защиты, детекция объектов, промышленная безопасность, SAHI, темпоральная фильтрация.*

Введение

Рост сложности производственных процессов и увеличение числа контролируемых параметров на промышленных объектах приводят к тому, что традиционный визуальный контроль соблюдения правил безопасности силами оператора становится недостаточно эффективным. По данным Международной организации труда, ежегодно в мире фиксируется около 340 миллионов несчастных случаев на производстве, значительная часть которых связана с человеческим фактором — невнимательностью, усталостью или субъективной оценкой ситуации [1]. Замена ручного мониторинга автоматизированными системами на основе компьютерного зрения позволяет не только снять нагрузку с оператора, но и обеспечить непрерывное наблюдение в реальном времени, исключив пропуски критических нарушений.

В настоящее время на рынке представлен ряд коммерческих систем видеоаналитики, ключевые недостатки систем — монолитная архитектура, не допускающая адаптации под специфику объекта, а также отсутствие автоматического обнаружения опасных зон без ручной настройки. Это определило направления собственной разработки, ориентированной на создание модульного программного комплекса, лишённого отмеченных ограничений.

Целью данной работы является разработка программного комплекса системы контроля средств индивидуальной защиты и производственной безопасности на основе компьютерного зрения, предназначенного для автоматического обнаружения нарушений в реальном времени и адаптации под различные производственные условия. Комплекс реализован как автономная модульная система, что обеспечивает универсальность и возможность развёртывания на различных промышленных объектах без привязки к конкретному производителю оборудования.

Для достижения поставленной цели в работе решаются следующие задачи:

- анализ существующих подходов к видеоаналитике для контроля промышленной безопасности;
- выбор архитектуры и инструментальных средств разработки;
- разработка и обучение нейросетевой модели для детекции СИЗ и опасных зон;
- реализация программного прототипа с модульной структурой;
- экспериментальная апробация системы на тестовых видеопотоках и оценка её эффективности.

Практическая значимость работы обусловлена возможностью снижения производственного травматизма за счёт автоматизации контроля, уменьшения нагрузки на операторов видеонаблюдения, а также гибкостью предлагаемого решения, позволяющего адаптировать систему под различные производственные условия без замены существующей инфраструктуры.

Анализ существующих решений

На сегодняшний день на рынке представлен ряд коммерческих систем видеоаналитики, предназначенных для контроля промышленной безопасности. Среди них можно выделить автоматизированную систему видеонализа концерна «Росэнергоатом», внедрённую на Кольской АЭС на базе платформы VizorLabs [2]. Программное ядро системы построено на каскаде нейросетей: первый детектор отвечает за обнаружение элементов СИЗ на персонале (способен распознавать 26 элементов экипировки на видео разрешением 640×480 при размере лица от 20 пикселей в высоту и угле наклона камеры до 60 градусов), второй — анализирует положение человека относительно границ опасных участков с использованием детектора-построителя скелета. По заявленным характеристикам, система выявляет до 98% инцидентов в режиме реального времени, доля ложных срабатываний не превышает 5%. Пилотный проект показал снижение количества нарушений с 80 до 8 в неделю, а в 2021 году система была включена в реестр отечественного программного обеспечения. Платформа обеспечивает производительность 30 и более камер на одной графической плате и до 200 камер на одном сервере, поддерживая масштабирование до промышленной инсталляции на сеть из 10 000 камер. Однако архитектура является монолитной и не предусматривает адаптации под специфику других производственных объектов, а также не поддерживает автоматическое обнаружение опасных зон.

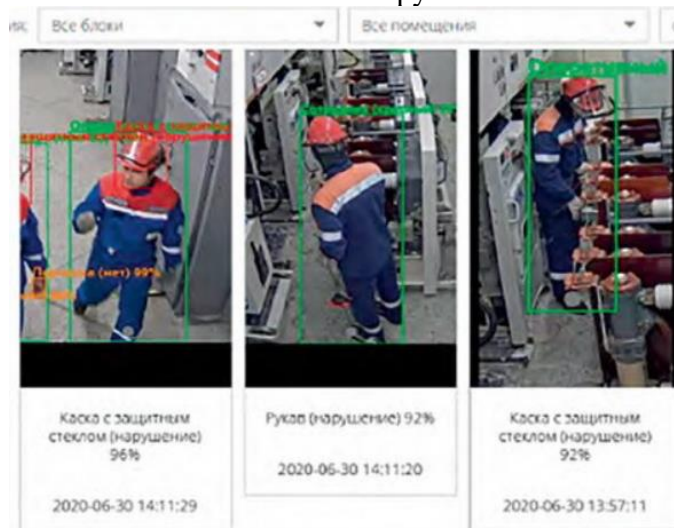


Рис.1 Детекция СИЗ системой

Платформа EYECONT.VMS от компании «Маллен Системс» представляет собой модульное клиент-серверное решение, спроектированное для эксплуатации в сложных производственных

условиях — при наличии пыли, грязи, тусклого освещения, высокой влажности и вибраций от оборудования [3]. Ключевыми функциональными модулями являются «Контроль опасных зон» и «Контроль СИЗ». Модуль контроля СИЗ способен обнаруживать наличие, комплектность и корректность ношения более 15 типов средств индивидуальной защиты, включая каски, спецодежду, респираторы, противошумные наушники, подбородочные ремни и сигнальные жилеты. Модуль контроля опасных зон обеспечивает детекцию присутствия персонала в статических и динамических запретных зонах в реальном времени. По заявленным характеристикам, точность обнаружения инцидентов достигает 99% при минимальном количестве ложных активаций, что подтверждено в ходе пилотных внедрений на промышленных объектах. Система обеспечивает интеграцию с цеховым оборудованием через API, включая АСУ ТП, что позволяет реализовывать сложные логические сценарии — например, активировать контроль нарушений только в момент работы оборудования. К достоинствам можно отнести адаптивность и возможность дообучения под требования производства, однако стоимость системы существенно возрастает при добавлении дополнительных модулей, а автоматическое обнаружение зон также не реализовано.

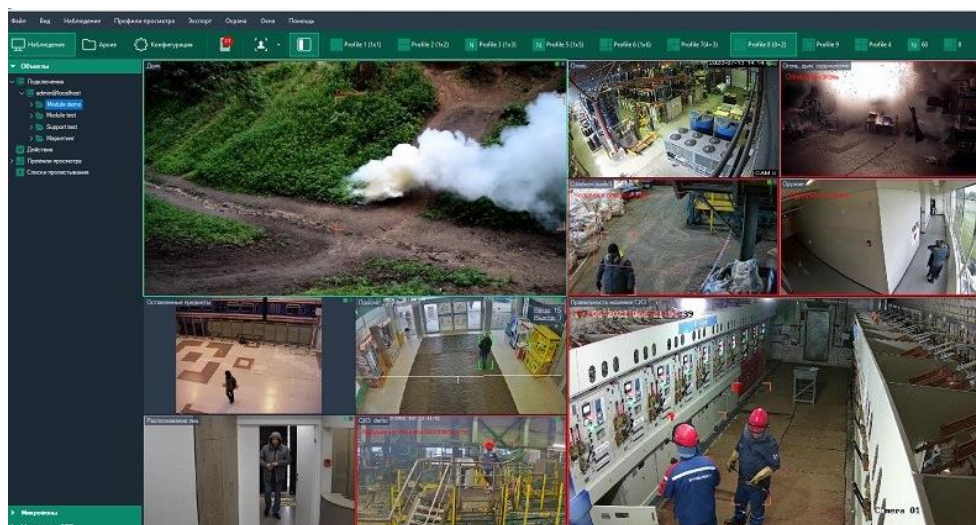


Рис.2 Мониторинг системы EYECONT.VMS

Решение Axis Object Analytics от AXIS Communications реализует встроенную видеоаналитику непосредственно на борту сетевых камер [4]. Приложение классифицирует движущиеся объекты на категории «человек» и «транспортное средство», поддерживает одновременное выполнение нескольких сценариев анализа. Ключевая особенность - периферийная архитектура: анализ видеопотока выполняется непосредственно на камере (на базе чипсетов ARTPEC-7 или ARTPEC-8 с аппаратным ускорением нейросетей), что исключает необходимость в выделенных серверах и снижает нагрузку на сетевую инфраструктуру.

Для уверенного распознавания человека в производственных условиях рекомендуется освещённость не менее 50 люкс, минимальный размер объекта - не менее 4% высоты кадра, а время движения - не менее 2 секунд для исключения ложных срабатываний. Система интегрируется с платформой AXIS Camera Station и основными VMS-системами сторонних производителей, что позволяет встраивать её в комплексные решения по промышленной безопасности.

Касательно управления и обнаружения разного рода зон и объектов рядом с ними, Axis отлично справляется и имеет большое количество сценариев работы с территориальными зонами.

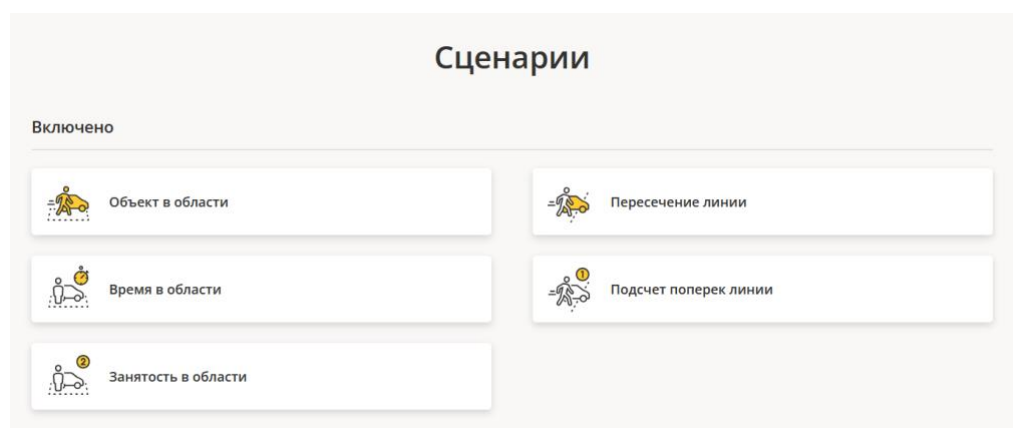


Рис.3 Пример сценариев работы с областями

Настройка сценариев позволяет гибко настроить взаимодействие триггер-зоны и объектов.



Рис.4 Присутствие человека в потенциально опасной или закрытой к посещению зоне

Таким образом, выявленные ограничения существующих решений обосновывают необходимость разработки собственного модульного программного комплекса, лишённого отмеченных недостатков.

Архитектура системы

Разработанная система построена по модульному принципу, что обеспечивает масштабируемость, гибкость настройки и возможность поэтапного внедрения на производственных объектах. Архитектура системы включает аппаратную и программную составляющие, тесно интегрированные между собой для обеспечения непрерывного цикла обработки видеоданных. Аппаратная часть в тестовой конфигурации представлена IP-камерой Cisco CIVS-IPC-6020 с разрешением 1920×1080 пикселей и частотой 30 кадров в секунду, управляемым коммутатором Cisco Catalyst 2960G для передачи видеопотоков, локальным вычислительным узлом на базе сервера Dell PowerEdge R610 с двумя процессорами Intel Xeon и дисковым массивом RAID 1 для отказоустойчивого хранения данных, а также операторским терминалом.



Рис.5 Демонстрация тестового стенда

Выбор компонентов обусловлен необходимостью обеспечения стабильной работы системы в условиях промышленных помещений, а также достаточной производительности для параллельной обработки нескольких видеопотоков с применением нейросетевых алгоритмов.

Важнейшим элементом аппаратной части является камера видеонаблюдения, качество работы которой напрямую зависит от разрешения, типа сенсора и условий освещения. Для эффективного контроля наличия СИЗ в зоне шириной 2–3 метра достаточно камеры с разрешением 1–2 Мп (Full HD), что подтверждается расчётом рекомендуемого разрешения исходя из минимального размера объекта (каска диаметром около 200 мм) и требуемого числа пикселей на объект (100×100) [5]. Так, для зоны шириной 2 метра и объекта 200 мм расчёт даёт 1000 пикселей по горизонтали, что соответствует разрешению 1 Мп. Рекомендуется использовать камеры с глобальным затвором (global shutter) для исключения искажений при съёмке движущихся объектов, монохромные сенсоры (при отсутствии необходимости различать цвета СИЗ) для повышения светочувствительности, а также объективы с ИК-коррекцией и защитным корпусом класса не ниже IP65 для работы в запылённых и влажных производственных средах.

Программная часть системы реализована на языке Python и состоит из семи взаимосвязанных модулей, образующих единый конвейер обработки видеокадров.

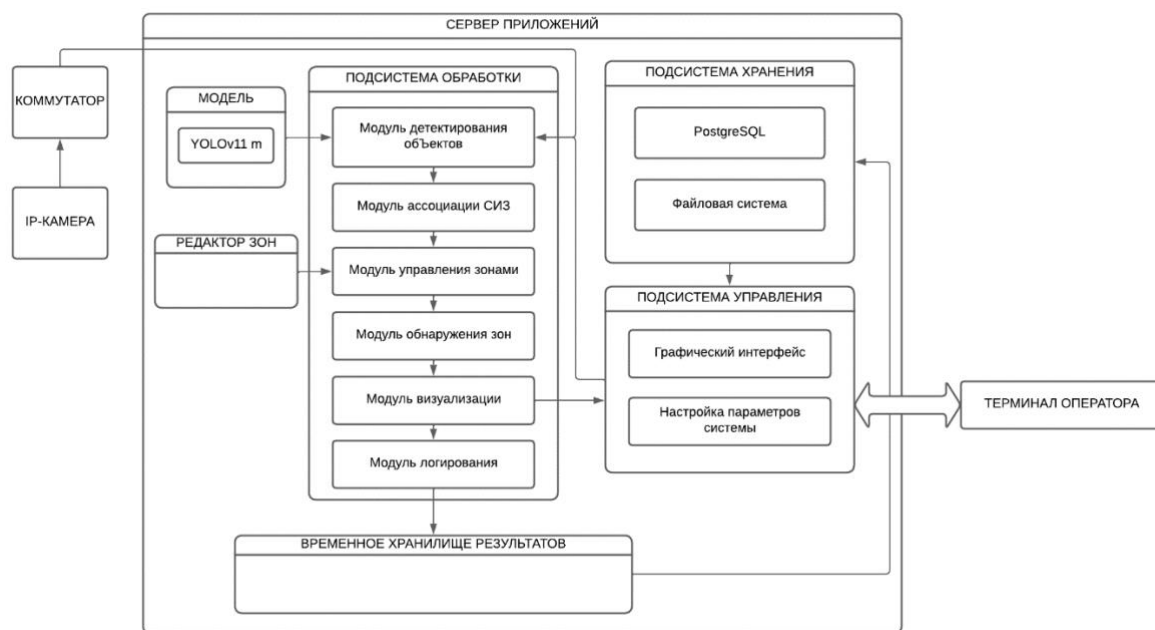


Рис.6 Структура программного комплекса системы контроля СИЗ и производственной безопасности

Модуль детектирования объектов принимает кадр с камеры, приводит его к стандартному разрешению, подавляет шумы с помощью фильтра Гаусса и улучшает локальный контраст в цветовом пространстве LAB с использованием алгоритма CLAHE [6], что обеспечивает устойчивость работы при неравномерном освещении. Далее выполняется тайловое разбиение кадра на перекрывающиеся фрагменты, каждый из которых обрабатывается нейросетевой моделью YOLOv11m, а результаты объединяются алгоритмом подавления немаксимумов. Список обнаруженных объектов, включающий класс, уверенность и координаты ограничивающего прямоугольника, передаётся в модуль ассоциации СИЗ с персоналом, который определяет, какие средства защиты надеты на конкретного человека, на основе пространственного перекрытия областей (порог перекрытия не менее 25 % площади бокса человека). При этом СИЗ, не пересекающиеся ни с одним человеком в кадре (например, каска, лежащая на полу или висящая на стойке), игнорируются, что исключает ложные срабатывания.

Модуль управления зонами загружает заранее определённые оператором конфигурации запретных зон из файла и проверяет каждого человека на попадание в эти зоны с учётом буферного отступа, обеспечивающего заблаговременное срабатывание. Для гибкости настройки предусмотрен визуальный редактор зон, в котором оператор может рисовать прямоугольные запретные области непосредственно на кадре с камеры. Параллельно работает модуль автоматического обнаружения зон, использующий нейронную сеть, предобученную на датасете СОСО, для детекции в кадре транспортных средств и объектов техники (экскаваторов, кранов, погрузчиков и т.п.), после чего вокруг них автоматически формируются запретные зоны с настраиваемым буферным отступом.

Поддерживаются три режима работы: «Вручную» (только зоны, заданные оператором), «Авто» (только автоматически обнаруженные зоны) и «Совместно» (объединение обоих источников), что позволяет адаптировать систему под различные производственные условия и требования безопасности.

Модуль визуализации формирует итоговое изображение с наложенными зонами, цветными рамками вокруг людей и метками, указывающими тип нарушения и уровень уверенности. При подтверждении нарушения (после прохождения темпоральной фильтрации) в верхней части кадра отображается сигнальный баннер с числом событий и текущим временем. Модуль логирования сохраняет аннотированные кадры на диск и фиксирует событие в базе данных PostgreSQL [7], включая временную метку, идентификатор камеры, номер человека, тип нарушения, уровень уверенности и координаты бокса. Использование PostgreSQL обеспечивает высокую надёжность

хранения благодаря поддержке транзакций по стандарту ACID, а также широкий набор типов данных и возможность интеграции с Python через библиотеку psycopg2. Модуль графического интерфейса, реализованный на Tkinter, предоставляет оператору средства для управления системой, просмотра видеопотока в реальном времени, настройки контролируемых типов СИЗ и режимов работы зон, а также просмотра журнала нарушений с аннотированными изображениями и детальной информацией по каждому событию.

Алгоритмы и методы

В качестве основного детектора объектов в системе выбрана нейросетевая модель YOLOv11m [8], которая обеспечивает высокую скорость обработки и хорошую точность за счёт однопроходной архитектуры. Архитектура YOLOv11 организована в три последовательных функциональных блока: экстрактор признаков (backbone), модуль агрегации признаков (neck) и выходной модуль детекции (head).

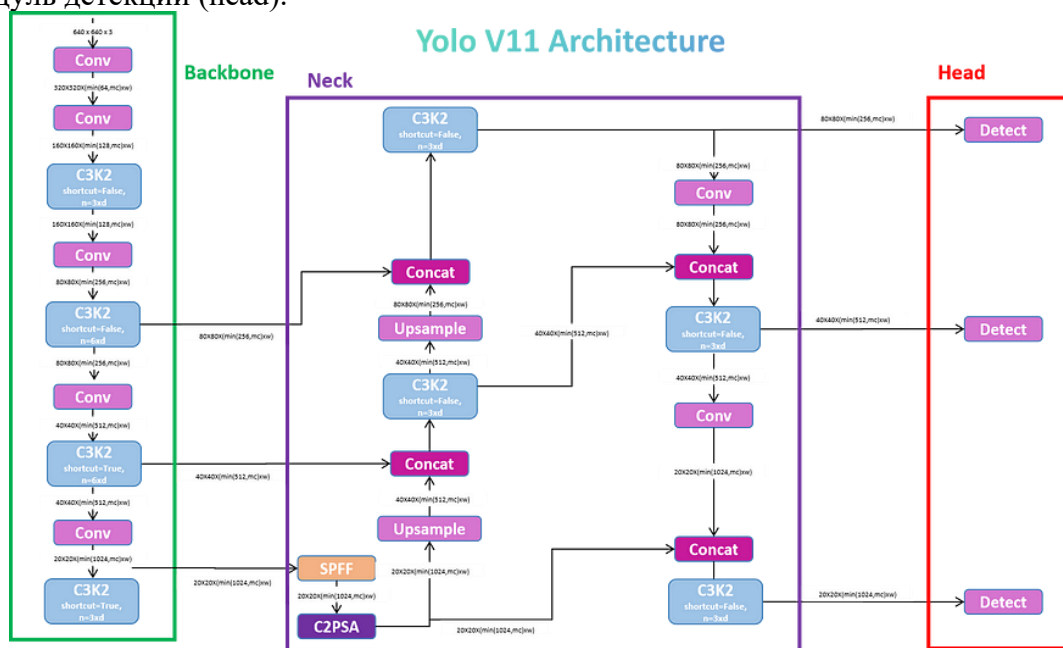


Рис.7 Архитектура YOLOv11

Входное изображение размером 640×640×3 пикселей последовательно понижается по разрешению через свёрточные слои и блоки C3K2, которые, в отличие от блоков C2F из YOLOv8, не разделяют входные данные, а пропускают их через последовательность bottleneck-слоёв с конкатенацией, что позволяет достичь баланса между точностью и производительностью при меньшем числе параметров. На выходе backbone формируются три карты признаков с шагами дискретизации 8, 16 и 32 пикселя, кодирующие объекты разных масштабов. В модуле neck применяется пространственная пирамидальная агрегация (SPPF) для учёта широкого контекста, а также блок C2PSA, реализующий механизм частичного пространственного внимания, который усиливает наиболее информативные области признаков и повышает точность обнаружения мелких или частично перекрытых объектов. Выходной модуль формирует предсказания для 11 классов, включая «Человек», «Каска», «Жилет», «Перчатки», «Очки», «Ботинки» и соответствующие классы нарушений («Нет каски», «Нет очков», «Нет перчаток», «Нет жилета», «Нет ботинок»). Итоговый список детекций формируется после фильтрации по порогу уверенности и подавления дублирующихся боксов алгоритмом Non-Maximum Suppression.

Поскольку стандартный размер входного изображения 640×640 пикселей не всегда позволяет хорошо распознавать очень мелкие объекты, такие как перчатки или защитные очки, особенно при большом расстоянии от камеры до работника, в системе применяется тайловая обработка с использованием библиотеки SAHI (Slicing Aided Hyper Inference) [9]. Кадр делится на перекрывающиеся фрагменты размером 640×640 пикселей с перекрытием 20% по горизонтали и

вертикали, каждый фрагмент обрабатывается нейронной сетью независимо, после чего все результаты объединяются алгоритмом GREEDYNMM с порогом IoU 0.25, что позволяет эффективно обнаруживать мелкие объекты на границах фрагментов и избегать дублирования. Для классов СИЗ и персонала установлены пороги уверенности 0,35–0,5, а для классов-нарушений порог намеренно снижен до 0,2, поскольку пропуск реального нарушения техники безопасности несёт значительно большие риски, чем ложное срабатывание. Такой подход обеспечивает высокую чувствительность к потенциальным угрозам при сохранении приемлемого уровня ложных срабатываний.

Для пространственной привязки обнаруженных СИЗ к конкретным работникам разработан модуль ассоциации, который для каждой камеры использует активный фильтр множества элементов СИЗ, обязательных к контролю. Принадлежность элемента СИЗ класса «С» к сотруднику w_k определяется условием, что центр ограничивающего прямоугольника этого элемента находится внутри прямоугольника сотрудника (формула 1.1)

$$p_k(c) = 1 \Leftrightarrow (x_c^*, y_c^*) \in b_k, \quad c \in R_i, \quad (1.1)$$

где $c = \{\text{«Человек»}, \text{«Каска»}, \text{«Отсутствие каски»}, \text{«Жилет»}, \text{«Отсутствие жилета»}\}$ и т.п., (x_c^*, y_c^*) — центр ограничивающего прямоугольника элемента СИЗ класса c ; b_k — прямоугольник сотрудника w_k . Проверка выполняется только для элементов активного фильтра R_i данной камеры.

Нарушение по СИЗ фиксируется при отсутствии хотя бы одного элемента из активного фильтра (формула 1.2)

$$\text{SIZ_violation}(w_k, R_i) = 1 \Leftrightarrow \exists c \in R_i : p_k(c) = 0 \quad (1.2)$$

При этом если на человеке одновременно обнаружено средство защиты и сигнал о его отсутствии, приоритет отдаётся подтверждённому наличию, что реализовано через логику взаимного исключения на основе карты соответствия классов нарушения и подтверждения.

Для дополнительного снижения ложных срабатываний, вызванных кратковременными помехами, временным перекрытием объектов или ошибками детекции, в системе реализован механизм темпоральной фильтрации. Нарушение считается подтверждённым только в том случае, если оно последовательно фиксируется на протяжении N видеокадров подряд (формула 1.3)

$$\text{confirmed}(w_k) = 1 \Leftrightarrow \sum_{t=0}^{N-1} \text{violation}(w_k, f_t) = N, \quad (1.3)$$

где N — настраиваемый параметр (в системе используется $N=5$, что соответствует задержке примерно 0,17 с при частоте 30 кадров/с). До момента подтверждения система продолжает работу в режиме мониторинга, отображая рамки вокруг людей и контуры зон, но не формируя активные оповещения. Итоговое условие генерации события нарушения объединяет все этапы:

$$E \neq \emptyset \Leftrightarrow W(f_t) \neq \emptyset \wedge \text{violation}(w_k, f_t) = 1 \wedge \text{confirmed}(w_k) = 1. \quad (1.4)$$

это гарантирует регистрацию только тех инцидентов, которые устойчиво наблюдаются в течение заданного интервала времени.

Обучение модели и результаты

Для обучения модели YOLOv11m был сформирован объединённый набор данных из трёх источников: датасет Person Detection Dataset [10], содержащий комплексные аннотации людей на разном расстоянии и в разных положениях; датасет SH17 [11] с разнообразными производственными сценами; и датасет Hard Hat Workers [12], включающий множество изображений работников в различной спецодежде и с разными средствами защиты. Общий объём датасета составил 3685 изображений с 14 279 объектами, при этом количество примеров для

каждого класса варьировалось: наибольшее число объектов приходилось на классы «Каска» (4502 объекта) и «Жилет» (2587 объектов), а наименьшее — на классы «Перчатки» (293 объекта) и «Ботинки» (335 объектов).

Процесс обучения включает несколько итераций с различными конфигурациями модели и наборов данных. Первоначальная версия v1 на базе YOLOv11n с использованием базового датасета construction-prp (содержит 1000 изображений) показала нестабильные результаты. Расширение датасета за счёт SH17 (v2) позволило увеличить mAP@0.5 до 0,523, при этом класс goggles достиг высокой точности (0,886), однако класс helmet остался на низком уровне (0,392) из-за недостатка примеров касок крупным планом. Добавление датасета Hard Hat Workers (v3) привело к резкому росту метрик по классу helmet до 0,964, но при этом была обнаружена критическая ошибка маппинга: классы head и helmet были перепутаны, что привело к инвертированию результатов. Исправление маппинга и добавление одноклассового датасета Person Detection (v4) позволило поднять точность класса Person до 83,3%, однако точность распознавания классов по_goggle и по_gloves снизились до 0,209 и 0,292 соответственно из-за дисбаланса в валидационной выборке.

Финальная версия была получена путём перехода на модель YOLOv11m (среднего масштаба) и объединения всех трёх датасетов с фиксированием уже обученных метрик классов и отсутствием переобучения классов с высокими показателями, после чего модель обучалась 150 эпох. Итоговые метрики показали значительное улучшение: mAP@0.5 достигла 76,4%, а mAP@0.5:0.95 составила 48,7%. Наибольшую точность демонстрирует класс «Жилет» (mAP@0.5 = 95,7%), что объясняется хорошей контрастностью и формой жилетов в обучающих данных. Класс «Каска» показал результат 81,7% при полноте (Recall) 93,9%, что является отличным показателем для системы безопасности. Что касается классов нарушений, ключевым является «Нет каски», для которого Recall достиг 91,2% при точности 77,6% и mAP@0.5 = 84,9%, что означает, что система пропускает лишь около 9% реальных нарушений, что критически важно для предотвращения несчастных случаев. Наихудшие результаты наблюдаются для классов «Нет ботинок» (mAP@0.5 = 52,5%) и «Нет перчаток» (68,8%), что связано с малым количеством обучающих примеров и сложностью распознавания мелких объектов. График F1-Confidence Curve показал, что оптимальный порог уверенности для баланса точности и полноты составляет 0,47, при котором F1-мера достигает 0,75, что и было использовано для настройки системы.

Экспериментальная апробация

Тестирование разработанной системы проводилось в трёх экспериментах, каждый из которых был направлен на проверку работоспособности и эффективности в различных условиях.

В первом эксперименте система запускалась на локальном компьютере без аппаратного ускорителя для проверки функционирования работоспособности системы. Была настроена ручная опасная зона (обозначена как «Ненадёжная конструкция») и активирован контроль трёх типов СИЗ: каска, перчатки и ботинки. На тестовом видеопотоке с изображением строительной площадки система верно определила всех работников, корректно идентифицировала наличие и отсутствие указанных СИЗ, а также зафиксировала нарушение — нахождение одного из сотрудников в опасной зоне. Инцидент был задокументирован в журнале с аннотированным кадром, временной меткой и уровнем уверенности, что подтвердило работоспособность логики ассоциации и контроля зон даже без использования мощного GPU.

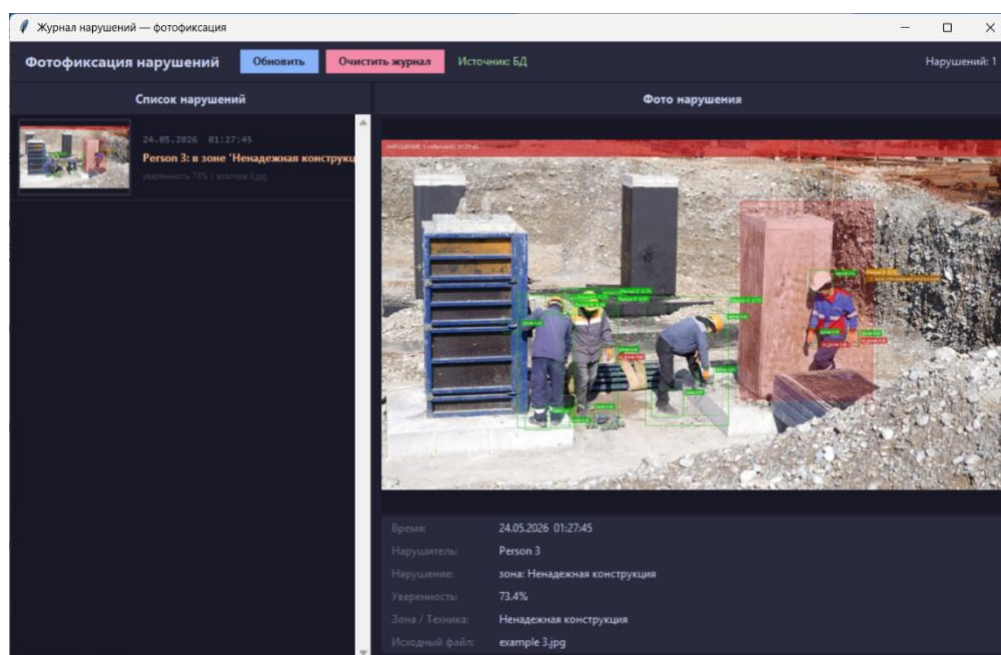


Рис.8 Эксперимент 1. Обнаружение нарушения и фиксация нарушения в журнале и базе данных

Во втором эксперименте проверялась работа системы в автоматическом режиме обнаружения опасных зон и при одновременном контроле всех пяти типов СИЗ. На видеопотоке, содержащем работников на производственной площадке, система выявила два нарушения — отсутствие защитных очков и перчаток у одного из сотрудников — и одновременно автоматически обнаружила две потенциально опасные зоны на основе детекции строительной техники (экскаватора и крана). Поскольку сотрудник не находился в пределах этих зон, нарушение по пересечению зон зафиксировано не было, что подтвердило корректность анализа потенциально опасных зон окружения.

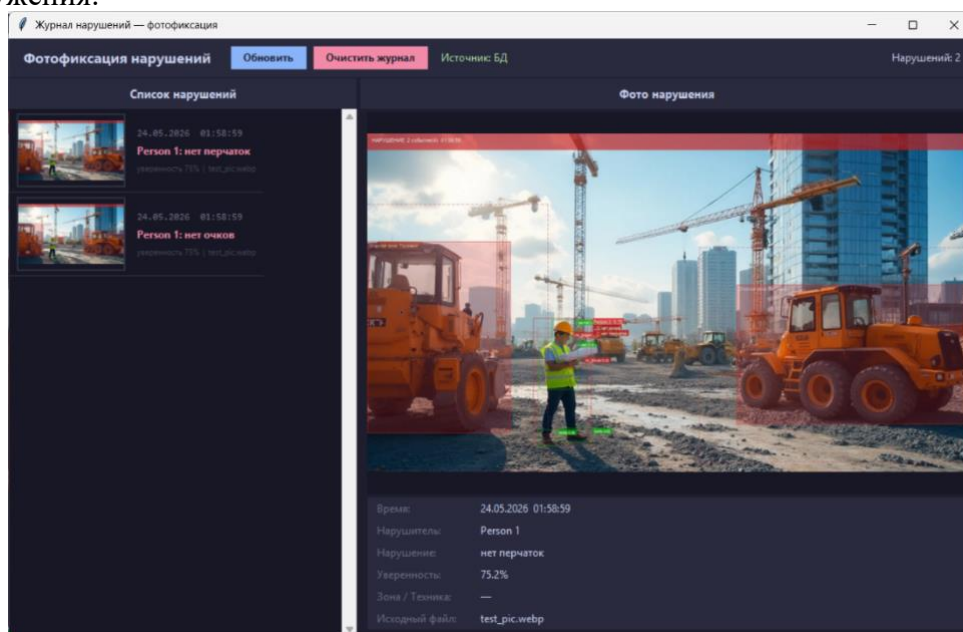


Рис.9 Эксперимент 2. Обнаружение нарушения и фиксация нарушения в журнале и базе данных

В третьем эксперименте система была развёрнута на серверном оборудовании Dell PowerEdge R610 с подключённой IP-камерой, что обеспечило производительность, достаточную для обработки видеопотока в реальном времени с частотой не менее 15 кадров в секунду. Были протестированы сценарии: отсутствие одного из элементов (каска), а также нахождение в опасной зоне, заданной вручную. Важно отметить, что система игнорировала СИЗ, находящиеся в кадре вне

зоны человека (например, каски на вешалке), что защищает алгоритм от ложных срабатываний и является прямым следствием реализованной пространственной ассоциации.

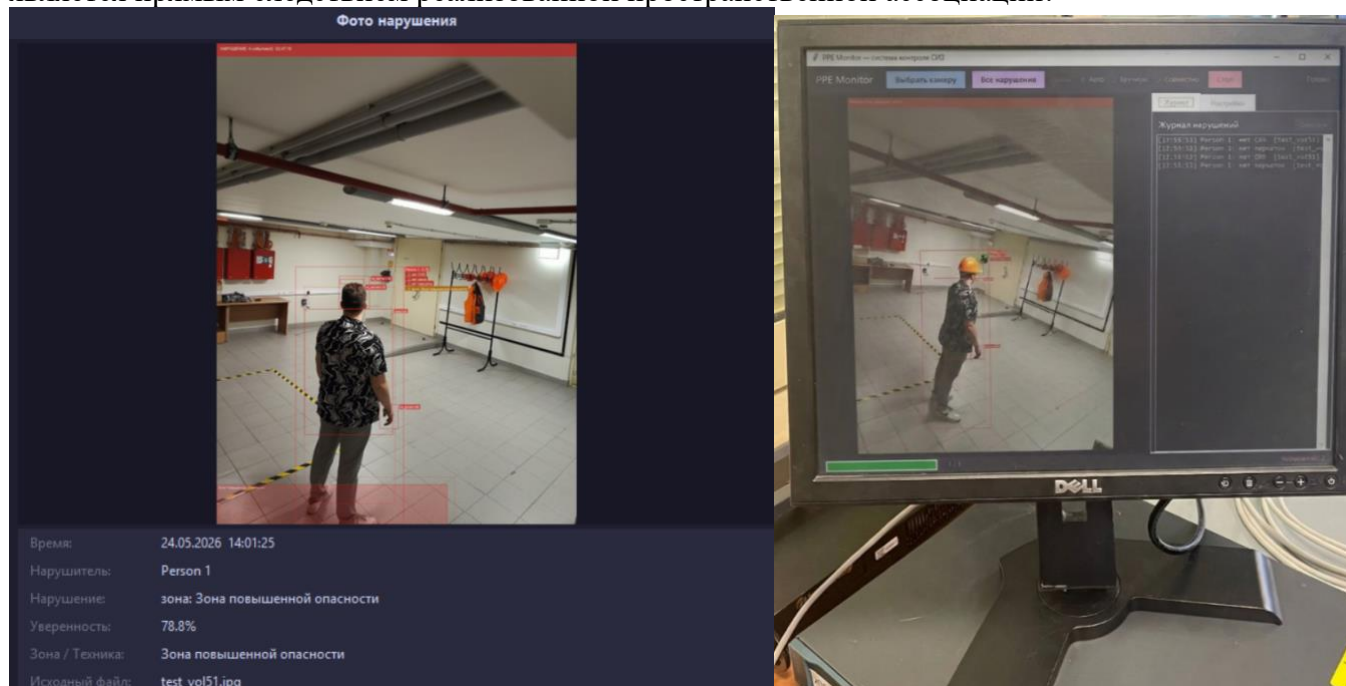


Рис.10 Эксперимент 3. Обнаружение нарушения и фиксация нарушения в журнале и базе данных

Во всех случаях система корректно определяла нарушения, фиксировала их в журнале с аннотированными кадрами и выводила уведомления на интерфейс оператора. Важным результатом стала демонстрация того, что система способна работать в реальных производственных условиях с минимальными задержками и без потери точности, что подтверждает её готовность к промышленному внедрению. Дополнительно было проверено, что использование тайловой обработки SAPH позволяет распознавать мелкие объекты (очки, перчатки) даже при значительном удалении от камеры, что невозможно при стандартном одноэтапном инференсе.

Выявление проблем и доработка системы

В ходе разработки и тестирования системы был выявлен ряд проблем, которые потребовали доработки для достижения заявленного уровня надёжности. Одной из первых проблем стало отображение СИЗ, не находящихся на человеке: любые СИЗ в зоне видимости камеры детектировались и отрисовывались на кадре, создавая визуальный шум и потенциально мешая восприятию оператора. Для устранения этой проблемы был добавлен фильтр, проверяющий, пересекается ли ограничивающий прямоугольник объекта СИЗ с боксом хотя бы одного обнаруженного человека; объекты, не пересекающиеся ни с одним человеком, исключаются из отображения и анализа. Другой проблемой стало некорректное определение людей в кадре, особенно в нестандартных позах (положение сидя, положение нааклона, работающие с инструментом). Для решения этой проблемы модель была дообучена на нескольких датасетах, содержащих большое количество изображений людей в разнообразных позах и на разном фокусном расстоянии, что позволило повысить полноту обнаружения класса Person с 76,7% в финальной версии до значений, обеспечивающих устойчивую работу в реальных сценариях.

Кроме того, динамика изменения эффективности алгоритмов в процессе обучения показала, что критическое значение для практического применения имеет не только средняя точность, но и сбалансированность метрик по классам нарушений. Переход от версии v5 к v6, включающий замену модели с YOLOv11n на YOLOv11m и объединение всех датасетов, позволил поднять общую mAP@0.5 с 0,746 до 0,764, а для класса no_helmet — до 0,849, что является лучшим результатом за всю историю обучения. Тем не менее, классы no_boots и no_gloves по-прежнему демонстрируют относительно низкие показатели (mAP@0.5 = 0,525 и 0,688 соответственно), что указывает на

необходимость дальнейшего расширения обучающей выборки для этих категорий, а также, возможно, применения специализированных методов аугментации для повышения устойчивости к вариативности формы и цвета обуви и перчаток. Потенциал масштабируемости системы позволяет в процессе эксплуатации собирать дополнительные данные и проводить дообучение модели без риска ухудшения распознавания других классов, что является важным конкурентным преимуществом по сравнению с монолитными коммерческими решениями.

Заключение

В ходе выполнения работы разработана и экспериментально апробирована система контроля средств индивидуальной защиты и производственной безопасности с использованием компьютерного зрения. Система построена на модульной архитектуре, обеспечивающей гибкость настройки и масштабируемость, и использует современную нейросетевую модель YOLOv11m в сочетании с тайловой обработкой SAHI для повышения точности распознавания мелких объектов. Реализованные механизмы пространственной ассоциации СИЗ с конкретными работниками и темпоральной фильтрации нарушений позволяют существенно снизить количество ложных срабатываний, что критически важно для систем безопасности, работающих в реальных производственных условиях. В отличие от рассмотренных коммерческих аналогов (Росэнергоатом, EYECONT.VMS, Axis Object Analytics), разработанная система поддерживает автоматическое обнаружение опасных зон на основе детекции техники, что исключает необходимость ручной настройки для каждого производственного участка, а также обеспечивает возможность комбинирования ручных и автоматических зон для максимальной гибкости.

Результаты экспериментального тестирования подтвердили работоспособность системы на различных сценариях — от локального компьютера без GPU до серверной конфигурации с IP-камерами. Метрики качества обученной модели ($mAP@0.5 = 76,4\%$, Recall по классам нарушений до 91,2%) соответствуют критериям эффективности, определённым для систем промышленной безопасности. Выявленные недостатки, такие как относительно низкая точность распознавания ботинок и перчаток, могут быть устранены путём дальнейшего расширения обучающей выборки и применения специализированных методов аугментации. Предложенные направления модернизации включают интеграцию с мультимодальными языковыми моделями (CLIP) для семантического анализа сцены и обнаружения нарушений, не заданных заранее фиксированными классами, а также подключение модулей активного обучения для автоматической адаптации к изменяющимся условиям производства. Таким образом, разработанная система представляет собой перспективное решение для автоматизации контроля промышленной безопасности, которое может быть адаптировано под широкий спектр производственных объектов и требований.

Литература

1. Министерство промышленности и торговли Республики Беларусь. Узел № 6244 [Электронный ресурс]. – URL: <https://www.mpt.gov.by/be/node/6244> (дата обращения: 13.11.2024).
2. Росэнергоатом. CCTV: Система видеоанализа соблюдения требований техники безопасности [Электронный ресурс]. – URL: <https://ai-russia.ru/library/rosenergo-cctv> (дата обращения: 17.11.2024).
3. Mallenom. EyeCont VMS [Электронный ресурс]. – URL: <https://mallenom.com/products/eyecont-monitoring-and-controlling-of-people-and-events/eyecontvms/> (дата обращения: 08.10.2024).
4. Axis Communications. Axis Object Analytics [Электронный ресурс]. – URL: <https://www.axis.com/products/axis-object-analytics> (дата обращения: 20.12.2024).
5. Complete Guide to Industrial Camera Selection for Machine Vision [Электронный ресурс] // Elementary ML. – URL: <https://www.elementaryml.com/blog/complete-guide-to-industrial-camera-selection-for-machine-vision> (дата обращения: 01.09.2025).

6. Черников Р.С. Улучшение качества потоковых изображений в реальном времени методом CLAHE // Программные продукты и системы. – 2023. – № 4. – С. 628-635.
7. What is PostgreSQL? [Электронный ресурс] // PostgreSQL. – URL: <https://www.postgresql.org/about/> (дата обращения: 20.03.2026).
8. Khanam R. YOLOv11: An Overview of the Key Architectural Enhancements / R. Khanam, M. Hussain // arXiv.org. – 2024. – 23 Oct. – DOI: 10.48550/arXiv.2410.17725.
9. Using YOLO with SAHI for Sliced Inference [Электронный ресурс] // Ultralytics. – URL: <https://docs.ultralytics.com/ru/guides/sahi-tiled-inference> (дата обращения: 01.03.2026).
10. Person Detection Object Detection Dataset [Электронный ресурс] // Roboflow Universe. – URL: <https://universe.roboflow.com/robolab-tnbci/person-detection-dizkz-6z0j5> (дата обращения: 24.03.2026).
11. SH17 — Safety Helmets Detection Dataset [Электронный ресурс] // Kaggle. – URL: <https://www.kaggle.com/datasets/mugheesahmad/sh17-dataset-for-ppe-detection> (дата обращения: 27.04.2026).
12. Hard Hat Workers Computer Vision Dataset [Электронный ресурс] // Roboflow Universe. – URL: <https://universe.roboflow.com/robolab-tnbci/hard-hat-workers-esvt4> (дата обращения: 05.05.2026).

DEVELOPMENT AND EXPERIMENTAL TESTING OF A COMPUTER-BASED GRAIN-BASED PERSONAL PROTECTIVE EQUIPMENT AND INDUSTRIAL SAFETY MONITORING SYSTEM

Mikhail A. Stulishenko,
Graduate, MTUCI, Moscow, Russia,
Stulishenkom@gmail.com

Mohammad Navar,
Associate Professor at the Department of ISU&A MTUCI
Ph.D., Docent, Moscow, Russia
navar@mtuci.ru

Abstract

This article presents a developed system for automatically monitoring compliance with personal protective equipment (PPE) wearing regulations and personnel presence in hazardous areas at industrial facilities. The system is built on a modular architecture and utilizes the YOLOv11m neural network model for real-time object detection. Tile processing based on the SAHI library is used to improve the recognition accuracy of small PPE components. Mechanisms for spatially associating detected PPE with specific workers, dual-mode hazardous area generation (manual and automatic based on equipment detection), and temporal violation filtering to reduce false alarms are implemented. The results of experimental testing of the system in various scenarios, including operation on a local computer and server hardware, are presented. The quality metrics of the trained model ($mAP@0.5 = 76.4\%$, Recall by violation classes up to 91.2%) confirm the effectiveness of the proposed approach for industrial safety tasks.

Keywords: computer vision, YOLOv11, personal protective equipment, object detection, industrial safety, SAHI, temporal filtering.