УДК 004.8

ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА ЛОКАЛЬНОГО АНАЛИЗА ПРОМЫШЛЕННОГО ВИДЕО

А. Е. АРХИПОВ, А. С. ЛОПАТКО, И. И. МАКСИМОВ
Тамбовский государственный технический университет
Тамбов, Россия

Автоматический контроль соблюдения требований безопасности по видеоданным затруднён одновременным присутствием нескольких типов неопределённости: изменением ракурса и освещения, перекрытием людей и оборудования, временной протяжённостью действий и неоднозначностью отдельного кадра. Применение только детектора объектов позволяет локализовать человека, но не определяет смысл производственной ситуации [1]. В работе предложена локальная интеллектуальная система MU Video Control, объединяющая детекцию, трекинг, видеоклассификацию, формализованные правила и дополнительный анализ локальной мультимодальной моделью.

Система выполнена как модульное web-приложение на Python и Flask. Видеоданные поступают из файла, web-камеры либо RTSP-потока. Для детекции людей применяется модель семейства YOLO [2]. При отсутствии весов предусмотрен CPU-режим на основе OpenCV HOG. Центроидный трекер присваивает объектам идентификаторы, а MediaPipe формирует позные ориентиры [3]. Параллельно специализированная VideoMAE обрабатывает окно из 16 кадров и относит ситуацию к одному из восьми производственных классов [4]. Независимые правила обнаруживают падение, скачок скорости, длительное присутствие и нераспознанное действие.

В отличие от схемы, в которой большая языковая модель анализирует каждый кадр, в предложенной архитектуре локальная VLM подключается только к сформированным событиям. Инциденты помещаются в неблокирующую очередь, поэтому задержка ответа VLM не останавливает видеоконвейер. Модель LM Studio либо Ollama получает evidence-кадр и исходный триггер, после чего уточняет уровень риска, формирует русскоязычное описание и рекомендацию оператору. При недоступности VLM сохраняется первичная оценка VideoMAE или правила, что обеспечивает функциональную деградацию без полного отказа.

Операторский интерфейс включает разделы мониторинга, инцидентов, аналитики, правил и администрирования. В интерфейсе отображаются активный источник, детектор, прогноз VideoMAE, состояние очереди VLM и число обработанных событий. Для инцидента доступны evidence, подтверждение, отметка ложного срабатывания и ручная повторная AI-проверка. Аналитический раздел строит временной график, распределения типов, рисков и статусов, экспортирует таблицы и формирует AI-резюме выбранного периода. Журнал сохраняется в CSV, JSON или TXT. Действия пользователей фиксируются в аудите.

Практический эксперимент включал последовательное открытие локальных видеозаписей, распознавание небезопасных фрагментов, сохранение evidence и автоматическую обработку локальной VLM. Подтверждена работа OpenAI-совместимого сервера LM Studio с моделью qwen/qwen3.6-35b-a3b: для реального кадра получены описание события и рекомендация без передачи изображения во внешнее облако.

Разработанный прототип реализует полный цикл – от получения видеопотока до решения оператора. Каскадное соединение специализированной видеомодели, интерпретируемых правил и локальной VLM снижает вычислительную нагрузку и сохраняет работоспособность при отказе отдельных компонентов. Система может быть использована как основа пилотного мониторинга производственного участка. Для промышленного внедрения требуются калибровка на целевых камерах, испытания RTSP-reconnect, регламент хранения evidence и оптимизация inference.

Работа выполнена при финансовой поддержке Министерства образования и науки Тамбовской области в рамках проекта № МУ2025-02/23 областного конкурса «Гранты для поддержки прикладных научных исследований молодых учёных в 2025 году».

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. **Глазкова, И. А.** К вопросу формализации функционирования сложных систем распознавания образов / И. А. Глазкова, М. А. Ивановский, Б. Х. М. Эль Эиссави // Вестник Тамбовского государственного технического университета. – 2024. – Т. 30, № 1. – С. 42–52.

2. YOLOv10: Real-Time End-to-End Object Detection / A. Wang, H. Chen, L. Liu [et al.] // Advances in Neural Information Processing Systems. – 2024. – Vol. 37.

3. **Никитников, Ю. В.** Применение компьютерного зрения при разработке системы опорно-двигательной реабилитации / Ю. В. Никитников, В. С. Круглов, М. А. Шильцын // Виртуальное моделирование, прототипирование и промышленный дизайн : материалы IX Междунар. науч.-практ. конф., Тамбов, 24–25 окт. 2023 г. – Тамбов : ТГТУ, 2023. – Вып. 9. – С. 177–179.

4. VideoMAE V2: Scaling Video Masked Autoencoders With Dual Masking / L. Wang, B. Huang, Z. Zhao [et al.] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. – 2023. – P. 14549–14560.