

УДК 004.932

DOI: [10.26102/2310-6018/2020.29.2.010](https://doi.org/10.26102/2310-6018/2020.29.2.010)

Интеллектуальный анализ видеоданных в системе контроля соблюдения правил промышленной безопасности

А.М. Вульфин

ФГБОУ ВО «Уфимский государственный авиационный технический университет»
Уфа, Российская Федерация

Резюме: Использование интеллектуальных камер и датчиков в системах видеоаналитики в сочетании с человеком-оператором, с которого снята большая часть аналитической и зрительной нагрузки, позволяет увеличить эффективность видеонаблюдения и, как результат, повысить безопасность и результативность труда на производстве в целом. Анализ существующих методов обработки данных в системах видеонаблюдения промышленного объекта показал, что применение бесконтактных способ анализа позы и действий человека в поле зрения камеры встречается редко, однако может иметь критическое значение в определенных ситуациях (человек в спецодежде лежит в зоне видимости камеры, но система на него не реагирует, ведь он находится не в запрещенной зоне). Рассмотрено совершенствование алгоритмов интеллектуального анализа видеоданных в системе контроля соблюдения правил промышленной безопасности (анализ типа динамики и контроль «свой-чужой») за счет использования нейросетевых технологий обработки. Оценка эффективности программной реализации алгоритмов анализа натуральных видеоданных показала корректность классификации в 97 % случаев. Оценка эффективности классификации 5 субъектов на два класса «свой» и «чужой» производилась методом перекрестной проверки и показала точность 99 % на тестовой выборке.

Ключевые слова: видеоаналитика, интеллектуальный анализ, распознавание типа динамики, нейронная сеть, классификатор, определение позы

Для цитирования: Вульфин А.М. Интеллектуальный анализ видеоданных в системе контроля соблюдения правил промышленной безопасности. *Моделирование, оптимизация и информационные технологии*. 2020;8(2). Доступно по: https://moit.vivt.ru/wp-content/uploads/2020/05/Vulfin_2_20_1.pdf DOI: 10.26102/2310-6018/2020.29.2.010

Intelligent analysis of video data in system for monitoring compliance with industrial safety rules

A.M. Vulfin

Ufa State Aviation Technical University, Ufa, Russian Federation

Abstract: The use of intelligent cameras and sensors, in combination with the human operator in video analytics systems, from which most of the analytical and visual load has been removed, allows you to increase the efficiency of video surveillance and, as a result, increase the safety and productivity of work in production as a whole. Analysis of the existing data processing methods in the video surveillance systems of industrial facility showed that the use of a non-contact method for analyzing person's posture and actions in the camera's field of vision is rare, but it can be critical in certain situations (person in overalls is in the camera's field of view, but the system is on him does not respond, because he is not in the forbidden zone). The improvement of algorithms for the intellectual analysis of video data in the system for monitoring compliance with industrial safety rules (analysis of the type of dynamics and control "friend or foe") using neural network processing technologies is considered. Effectiveness evaluation of algorithms for analyzing full-scale video data software implementation showed the correctness of classification in 97% of cases. Effectiveness evaluation of the 5 subjects into two classes

of “own” and “alien” classification was carried out by cross-validation and showed an accuracy of 99% on the test sample.

Keywords: video analytics, intelligent analysis, dynamics type recognition, neural network, classifier, pose determination

For citation: Vulfin A.M. Intelligent analysis of video data in system for monitoring compliance with industrial safety rules. *Modeling, optimization and information technology*. 2020;8(2). Available from: https://moit.vivt.ru/wp-content/uploads/2020/05/Vulfin_2_20_1.pdf DOI: 10.26102/2310-6018/2020.29.2.010 (In Russ).

Введение

Совершенствование систем видеоаналитики – систем, использующих методы компьютерного зрения для автоматизированного анализа данных, получаемых из последовательности изображений, поступающих с видеокамер в режиме реального времени, – позволит повысить эффективность контроля соблюдения правил техники безопасности и снизить вероятность возникновения инцидентов на производстве, обусловленных человеческим фактором.

Для обеспечения эффективного функционирования системы контроля соблюдения правил промышленной безопасности при эксплуатации опасных производственных объектов актуальной является разработка подсистемы интеллектуального анализа видеоданных. Основные задачи подсистемы:

- обнаружение человека в запрещенной зоне;
- анализ позы и типа активности человека (не двигается, лежит, бежит);
- оперативный контроль за соблюдением сотрудниками должностных инструкций и технологического регламента, соглашения по охране труда, других локальных нормативных правовых актов организации (например, выполнение регламента обслуживания оборудования, обязательное использование спецодежды);
- оперативный контроль за применением средств индивидуальной и коллективной защиты согласно установленному регламенту.

Цель создания подсистемы – совершенствование алгоритмов интеллектуального анализа видеоданных в системе контроля соблюдения правил промышленной безопасности (анализ типа динамики и контроль «свой-чужой») за счет использования нейросетевых технологий обработки.

Для достижения указанной цели решены следующие задачи:

- 1) анализ существующих систем видеоаналитики промышленного объекта;
- 2) разработка структурной схемы системы интеллектуального анализа видеоданных;
- 3) разработка алгоритмов анализа видеоданных на основе методов интеллектуальной обработки;
- 4) программная реализация разработанных алгоритмов и оценка эффективности предложенного решения на натуральных данных.

1. Анализ возможностей существующих систем видеоаналитики в задачах мониторинга промышленной безопасности на предприятии

Автоматизация анализа видеоизображения позволяет предприятию значительно увеличить продуктивность видеоконтроля за счет снижения нагрузки на человека-оператора системы, избегая чрезвычайных ситуаций, связанных с снижением концентрации оператора при длительном контроле происходящего на множестве

экранов. Особенно актуальны подобные системы на производственных объектах, технологический процесс на которых неразрывно связан с потенциальными рисками для жизни и здоровья человека.

1.1. Сравнительная характеристика существующих вариантов анализа видеопотока

Варианты анализа видеопоследовательности [1, 2] в существующих системах можно разделить на три основных направления:

- классическая видеоаналитика;
- ситуационная видеоаналитика;
- сервисная видеоаналитика.

В классической видеоаналитике чаще всего применяются алгоритмы сравнения изображений с заранее созданными паттернами. Применяются детекторы различной сложности. Примеры детекторов:

- детекторы движения;
- детекторы направления;
- детекторы пересечения линии;
- детекторы лиц;
- счетчики каких-либо объектов.

В основе ситуационной аналитики лежат алгоритмы, которые позволяют обеспечивать контроль за:

- использованием персоналом средств технической защиты;
- ограничением прохода в опасные зоны;
- вхождение в зону / выход;
- появление в зоне / исчезновение;
- нахождение, остановка, задержка в зоне;
- превышение количества объектов в зоне;
- одновременный проход;
- оставленные предметы;
- акустический мониторинг (крик, сирена, ударная волна, хлопок).

Сервисная видеоаналитика позволяет контролировать работу видеокамер и видеопотока с применением детекторов, генерирующих соответствующие сигналы:

- при потере сигнала от камеры регистрируется нарушение соединения между сервером и камерой;
- при потере резкости изображения (грязь на объективе, конденсат на оптике);
- при регистрации направленного в объектив луча яркого света с целью засветки матрицы фоточувствительных элементов;
- регистрация затемнения изображения в случаях закрытия объектива, выхода из строя устройства освещения;
- регистрация несанкционированного поворота камеры (изменение поля зрения).

1.2. Анализ возможностей существующих систем видеоаналитики для мониторинга промышленной безопасности «CenterVisio» от компании «Центр 2М» [3]

Система промышленной видеоаналитики «CenterVisio» компании Центр 2М. Система базируется на комплексном подходе к видеоаналитике, активно используются возможности нейросетевого анализа:

- использование обученной нейросети с точность распознавания до 98%;

- среднее время отклика системы 2,4 с позволяет оперативно реагировать на нарушения техники безопасности на производстве;
- незначительное влияние погодных условий;
- точность распознавания позволяет уменьшить количество видеокамер, необходимых для аналитики.

Комплекс промышленной безопасности от компании «ISS» [4]

Решение от ISS (Интеллектуальные системы безопасности) предоставляет комплексный подход по обеспечению безопасности на промышленном объекте, включающий в себя видеоконтроль за территорией предприятия, обнаружение событий, потенциально приводящих к возникновению угроз, а также инициацию протоколов, позволяющих минимизировать ущерб при выявлении инцидентов.

«Video Analytics & Artificial Intelligence» от компании «IntelliView Technologies Inc» [5]

Система основана на использовании технологий видеоанализа и применения методик глубокого обучения и искусственного интеллекта. Система позволяет выполнять удаленную проверку оборудования, наблюдение за периметром, обнаружение утечек жидкостей и так далее.

Несмотря на разнообразие представленных на рынке систем, все они реализуют схожий функционал (Таблица 1).

Таблица 1 – Сравнительная характеристика систем анализа видеоданных
Table 1 – Comparative characteristics of video analysis systems

Система	Использование нейронных сетей	Контроль территории	Верификация персонала по изображению лица	Наличие спецодежды на работнике предприятия	Акустический мониторинг	Детализация настроек параметров системы	Ограничение влияния погодных условий
Center Vision	+	+	+	+	-	-	+
ISS SecurOS	+	+	+	+	+	-	+
VA and AI	+	+	+	-	-	+	+

Описанные системы могут обнаруживать работников в запрещенных зонах, определять наличие на них средств спецзащиты, но при этом не в полной мере реализован механизм оценки типа движения человека (спит на рабочем месте, не двигается на рабочем месте, находится в специфической позе), что может иметь критическое значение в ситуациях, когда, например, человек в спецодежде лежит в зоне видимости камеры, но система на него не реагирует, так как он находится не в запрещенной зоне. Таким образом, анализ позы и типа движения представляет собой возможность улучшить стандартный подход к реализации алгоритмов интеллектуальной видеоаналитики.

1.3. Декомпозиция работы системы видеоаналитики

В контексте машинного зрения, исходя из предложенной Д. Марром [6] парадигмы, принято выделять следующие ступени процесса анализа изображения:

- 1) преобработка изображений;

- 2) сегментация (выделение признаков);
- 3) выделение геометрической структуры (селекция признаков);
- 4) определение относительной структуры и семантики (классификация).

Анализ позы человека с помощью алгоритмов компьютерного зрения нетривиален ввиду анатомической сложности строения человека. Тело человека состоит из множества сочленений, суставов (как больших, так и едва заметных), что обуславливает сложность геометрического анализа позы [7, 8]. В [7] предложено решение, позволяющее задействовать в анализе позы глубокие нейронные сети, рассматривая эту задачу как задачу многомерной регрессии.

Входное изображение представлено в виде матрицы, а вектор позы как координаты положения суставов. Скелет человека представляется в виде графа из k вершин. На вход нейросетевого блока поиска ключевых точек изображения – суставов – подается трехканальное изображение (RGB формат) фиксированного размера, а на выходе получается k пар координат суставов:

Сверточная нейронная сеть принимает на вход изображение фиксированного размера (220 x 220 пикселей). Существенная часть информации при подобном сжатии теряется, вместе с этим теряется и возможность анализировать мелкие детали изображения. Увеличение размера входного изображения приводит к увеличению количества настраиваемых параметров нейронной сети и проблемам переобучения [9].

Компромиссом между подачей на вход сети изображения в высоком разрешении и адекватным числом параметров стало использование каскада нейронной сетей:

1. По исходному изображению, сжатому до 220 x 220 пикселей, предсказываются приблизительные координаты суставов.
2. Каждая последующая нейронная сеть уточняет координаты по новому изображению.
3. Новое изображение – часть предыдущего изображения (ограничивающая область вокруг предсказания с предыдущего этапа), приведенная к размеру 220 x 220 пикселей.

2. Разработка системы обработки видеоданных

Предлагается следующая обобщенная структурная схема системы анализа видеоданных в задаче распознавания типа динамики человека, представленная на Рисунке 1.

Приняты следующие обозначения для структурных элементов схемы: 1 – видеоряд с привязкой к отметкам времени; 2 – нормализованный видеоряд с двумя цветовыми палитрами (RGB и Greyscale); 3 – силуэт человека; лицо человека (ROI – region of interest); ROI силуэта человека в видеоряде с привязкой к метке времени; ROI лица человека с привязкой к метке времени; 5 – компактный вектор признаков (ключевые точки скелета + ключевые точки лица); 6 – специалист по извлечению и представлению знаний; 7 – вектор оценок принадлежности объекта к возможным классам состояний, полученный набором классификаторов (RF – Random Forest, классификатор на основе «случайного леса», CNN – Convolution Neural Network, классификатор на основе сверточной нейронной сети, SVM – Support Vector Machine, классификатор на основе машины опорных векторов, KNN – k-Nearest Neighbor, классификатор k-ближайших соседей); 8 – лицо, принимающее решение; 9 – системный администратор; БПР – блок принятия решения; НМИ – Humane Machine Interface, человеко-машинный интерфейс; PCA – Primary Component Analysis, метод главных компонент; HOG – Histogram of Oriented Gradients, гистограмма ориентированных градиентов; LBP – local binary patterns, локальные бинарные шаблоны.

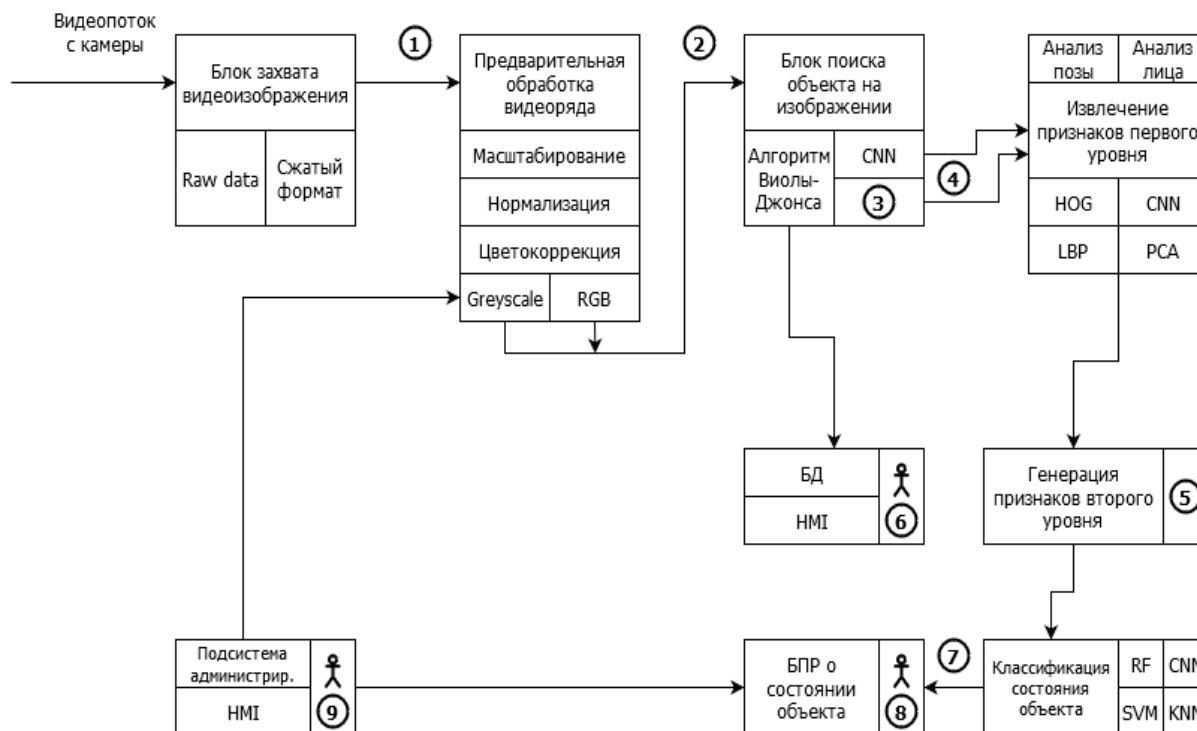


Рисунок 1 – Структурная схема системы анализа видеоданных в задаче распознавания типа динамики человека

Figure 1 – Block diagram of video data analysis system in the problem of recognizing the type of human dynamics

2.1 Разработка алгоритма анализа видеопоследовательности

Алгоритм распознавания позы человека представлен на Рисунке 2:

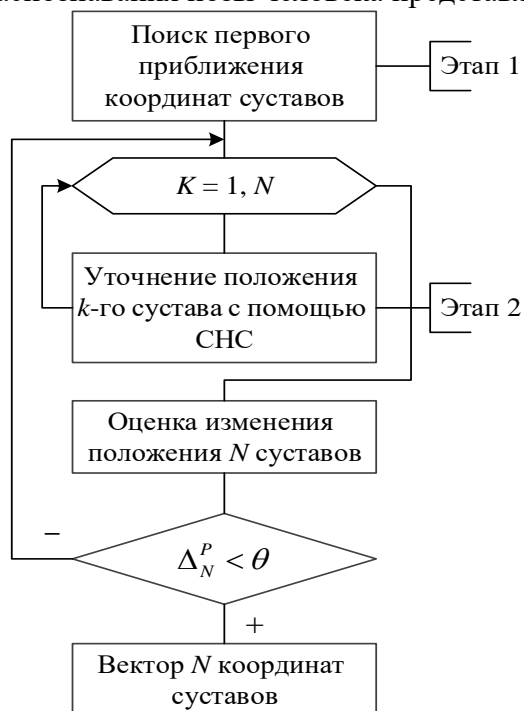


Рисунок 2 – Схема алгоритма распознавания позы

N – количество ключевых точек – суставов; Δ_N^P – область локализации k -го сустава, сравниваемая с пороговым значением θ

Figure 2 – Pose Recognition Algorithm Diagram

Схема алгоритма классификации человека по выделенному изображению лица на два класса «свой-чужой» представлена на Рисунке 3.

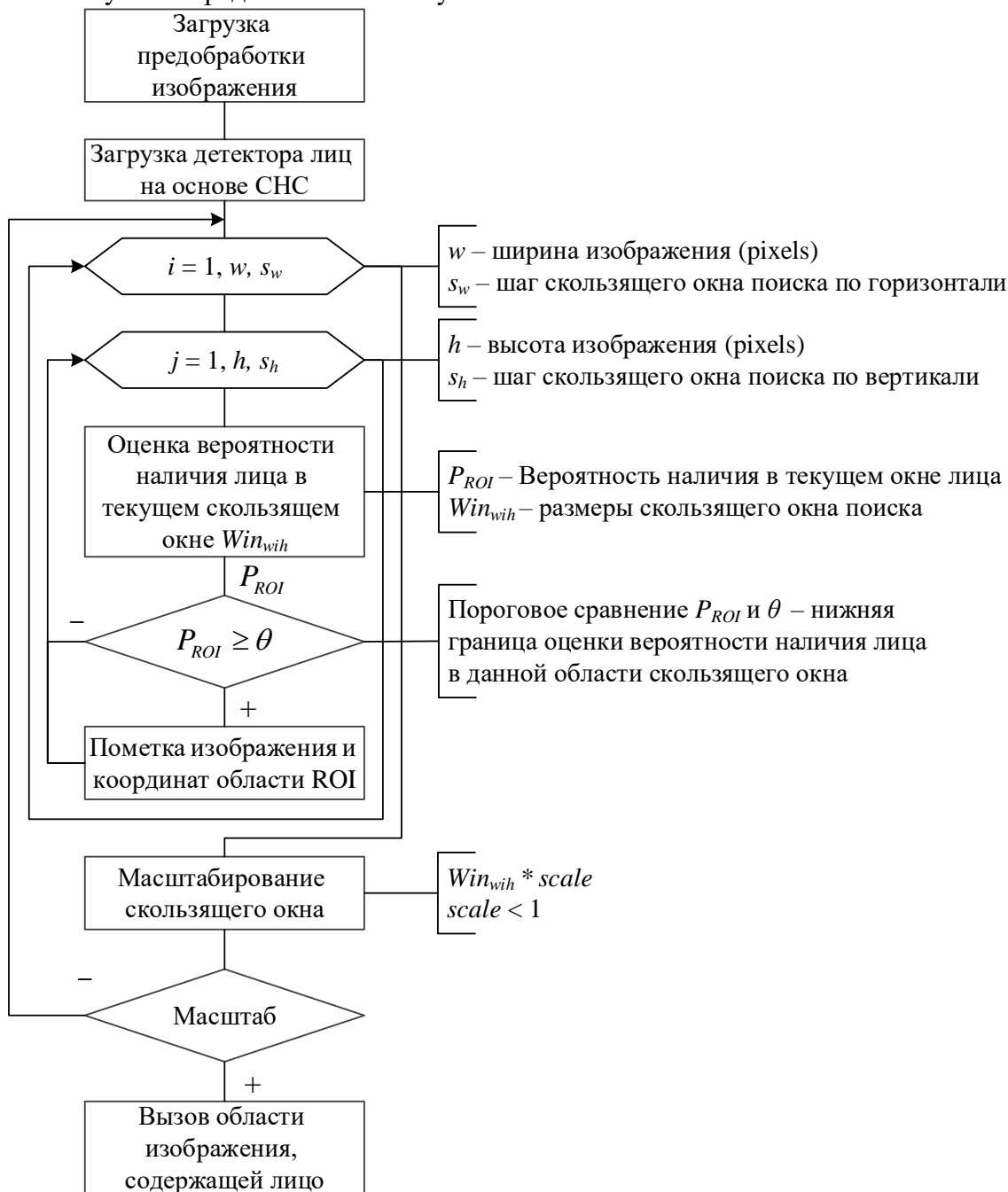


Рисунок 3 – Схема алгоритма распознавания лица
 СНС – сверточная нейронная сеть

Figure 3 – Face Recognition Algorithm

Подготовка набора данных для обучения нейросетевого классификатора

Видеоизображения взяты из источника Berkeley Multimodal Human Action Database (MHAD) [10] и обработаны системой OpenPose, позволяющей определять 135 ключевых точек человеческого тела.

Набор данных MHAD включает видеоизображения 12 субъектов, выполняющих 6 действий по 5 повторений, снятых с четырех углов зрения с разного расстояния (Рисунок 4).



Рисунок 4 – Кадр из видео, иллюстрирующего действие «Подъем двух рук»
Figure 4 – Scene from a video illustrating the “WAVING_2HANDS” action

Типы основных действий субъекта, представленные в наборе данных:

- прыжки (JUMPING);
- прыжки с подъемом рук (JUMPING JACKS);
- боксирование (BOXING);
- подъем двух рук (WAVING_2HANDS);
- подъем одной руки (WAVING_1HAND);
- хлопки в ладоши (CLAPPING_HANDS).

Всего было использовано 1438 видеофрагментов, содержащих 211200 кадров.

Набор данных состоял из:

- 211200 изображений;
- длина фрейма – 32 кадра (≈ 1.5 с движения человека при частоте дискретизации 22 Гц);
- Количество изображений с ошибками обнаружения поз (обнаружение более двух человек в кадре) – 5132;
- Коэффициент разбиения выборки на обучающую и тестовую 0,8;
- Перекрывание соседних последовательностей изображений с коэффициентом 0,8125 (26/32);
- Количество примеров в обучающей выборке – 22625 * 32 кадра;
- Количество примеров в тестовой выборке – 5751 * 32 кадра.

Разработка классификатора для распознавания типа динамики

Используется классификатор на основе LSTM-сети (классификатор на основе долгой краткосрочной памяти, LSTM – long short-term memory) [11].

Исходные видеопоследовательности были обработаны с помощью библиотеки OpenPose [12] для выделения 18 ключевых точек скелета человека. На один кадр приходится N опорных выделенных точек скелетной анимации (Рисунок 5):

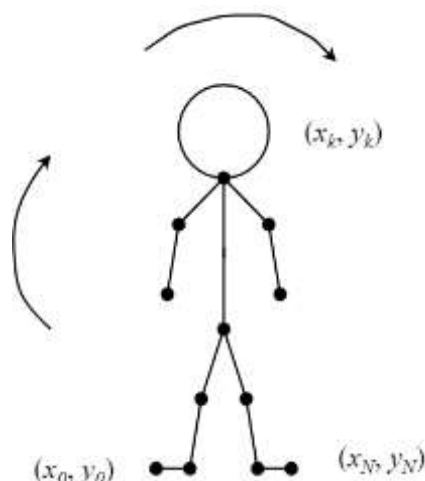


Рисунок 5 – Схематичное обозначение опорных точек
 Figure 5 – Schematic designation of control points

Ключевые точки скелета представлены множеством пар:

$$P = \{(x_k, y_k)\},$$

где $k = \overline{1, N}$, $N = 18$ (количество выделенных опорных точек).

Выделенные данные сохранены в текстовых структурированных файлах формата CSV и содержат X и Y координаты каждого сустава в определенном типе движения в последовательности кадров для формирования обучающей и тестовой выборки.

Для того чтобы распознавать статичные позы и движения, необходимо учитывать изменения положения ключевых опорных точек во времени и пространстве. Для этого необходимо рассматривать фреймы – последовательность кадров видеопотока с выделенными опорными точками [13].

Задание параметров обучающей и тестовой выборки для распознавания типа динамики

Последовательности фрейма фиксированной длины, описывающего тип активности, формируются из исходных видеоданных с взаимным перекрытием с коэффициентом 0,8, что позволяет корректно обрабатывать начало и завершение цикла движения и плавную смену движений.

Получим следующие параметры выборки:

- типов распознаваемых классов – 6;
- обучающая выборка содержит 22625 последовательностей 32-х кадров;
- тестовая выборка содержит 5751 последовательностей 32-х кадров.

Таблица 2 – Параметры классификатора типа динамики

Table 2 – Dynamics classifier parameters

Параметры классификатора	Значение параметра
Адаптивный коэффициент скорости обучения	Да
Начальное значение коэффициента скорости обучения	0,0025
Начальное значение коэффициента экспоненциального уменьшения (затухания) коэффициента скорости обучения	0,96
Количество итераций затухания коэффициента скорости обучения (decay_steps)	100000
Количество эпох обучения (training_iters)	300
Размер пакета в режиме пакетного обучения (batch_size)	512

Подготовка данных для классификатора типа динамики

Матрица попарных расстояний между опорными точками внутри одного кадра строится следующим образом:

$$\forall(p_i, p_j) \rightarrow d_{i,j}$$

Далее каждый элемент матрицы нормализуется с помощью деления на $\max_{i,j} d$ в данном кадре, и вся матрица построчно разворачивается в вектор $D = N(N - 1) = 306$.

Далее для каждого кадра D находится «центр тяжести» $d_j = \text{dist}(p_j, \{\bar{x}, \bar{y}\})$, $j = 1, N$ и нормируется вектор d на $\max_j d$.

Процедура нормализации позволяет учесть масштабирование изображения при разном расстоянии от объекта съемки до камеры.

Разработка классификатора для анализа человека по изображению лица

Выборка содержит по 1000 изображений 5 субъектов, которые разделены на два класса: «свой» и «чужой». Сначала фотографии (128 x 128 пикселей) сжимаются методом главных компонент (PCA) до 32-х компонент, которые впоследствии подаются в качестве компактного вектора признаков в многослойный персептрон [14].



Рисунок 6 – Примеры изображений, анализируемых с помощью НС
 Figure 6 – Examples of images analyzed using a neural network

Структура подготовки данных для создания обучающей выборки и построения классификатора приведена в [14, 15] (Рисунок 7).



Рисунок 7 – Структура подсистемы распознавания человека по изображению лица
 Figure 7 – The structure of the human recognition subsystem by face image

Структура нейросетевого классификатора на основе многослойного персептрона представлена в Таблице 3.

Таблица 3 – Структура нейросетевого классификатора на основе многослойного персептрона
 Table 3 – The structure of the neural network classifier based on the multilayer perceptron

Наименование слоя	Количество нейронов	Функции активации нейронов
Входной слой	32	-
Скрытый слой	12	гиперболический тангенс
Выходной слой	5	гиперболический тангенс

3. Оценка эффективности алгоритма и анализ результатов

Процесс обучения нейросетевого классификатора для распознавания типа динамики занял 679 тыс. итераций (Рисунок 8). Показано изменение среднеквадратичной ошибки классификатора на обучающей и тестовой выборках в зависимости от номера текущей эпохи в процессе обучения.

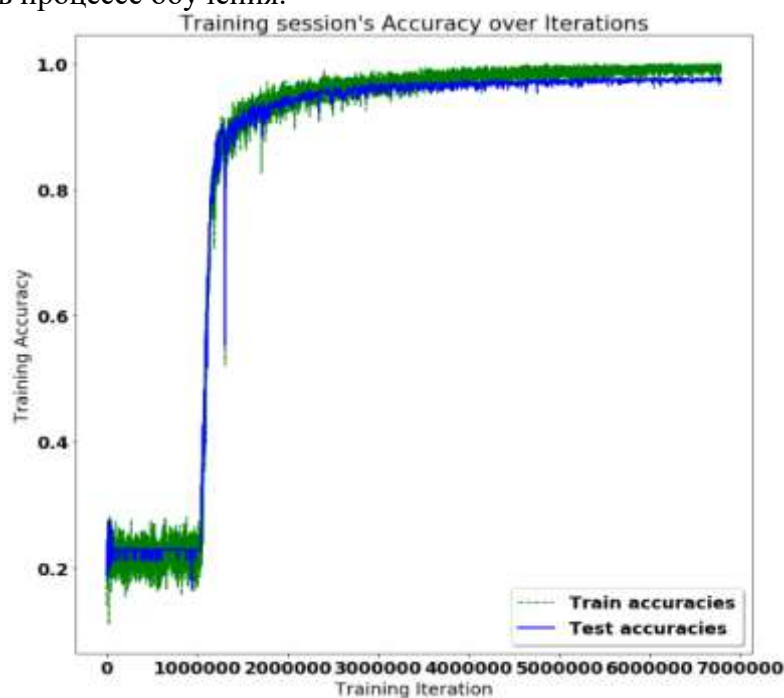


Рисунок 8 – График изменения суммарной квадратичной ошибки нейросетевого классификатора. По оси абсцисс – итерации обучения, по оси ординат – среднеквадратичная ошибка. Представлены графики изменения ошибки на обучающем (Train accuracies) и тестовом (Test accuracies) множествах.

Figure 8 – Schedule of changes in the total quadratic error of the neural network classifier.

Итоговая оценка точности классификации на обучающей выборке: 97.39 %, на тестовой выборке: 97.40 %.

На Рисунке 9 приведена матрица неточности обученного нейросетевого классификатора, что позволяет оценить ошибки классификации по каждому из анализируемых классов.

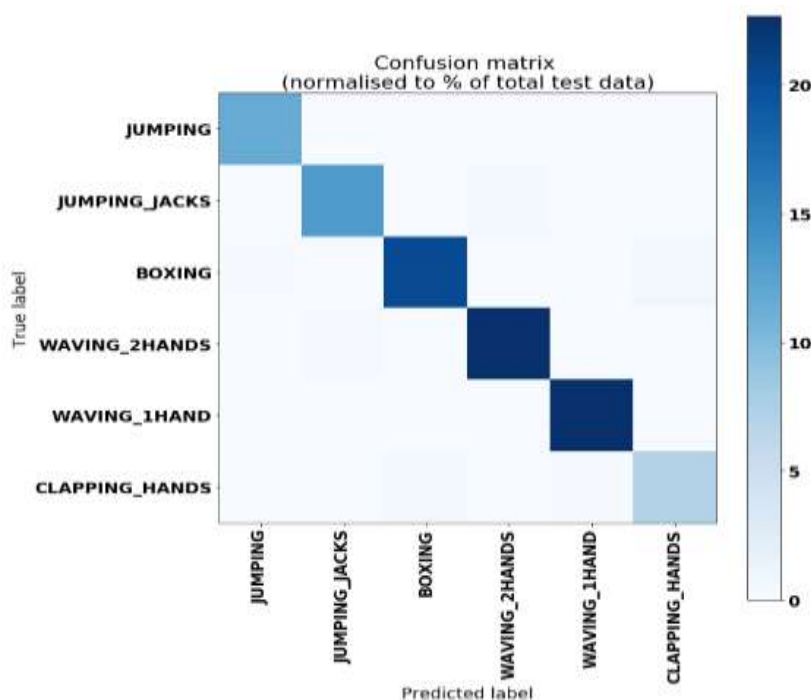


Рисунок 9 – Матрица неточностей нейросети на тестовой выборке
 Figure 9 – The matrix of inaccuracies of the neural network in the test samples

График изменения суммарной квадратичной ошибки на обучающей, тестовой и проверочной выборках для нейросетевого классификатора, предназначенного для отнесения человека по подготовленному изображению лица к классам «свой-чужой», представлен на Рисунке 10.

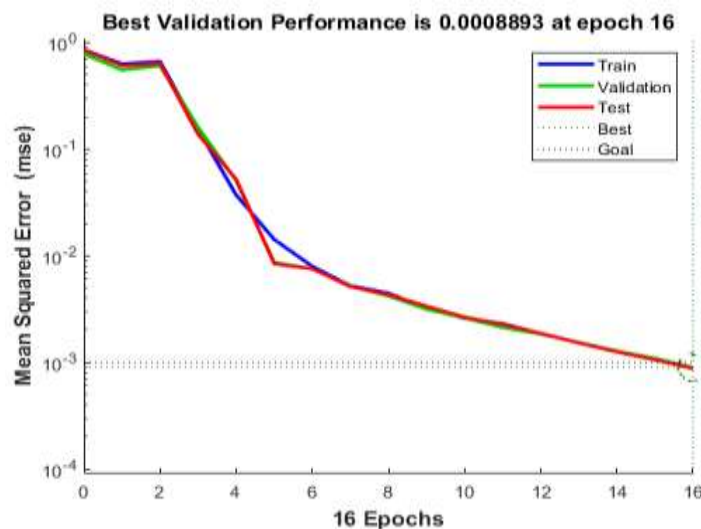


Рисунок 10 – График изменения квадратичной ошибки.
 По оси абсцисс – итерации обучения, по оси ординат – среднеквадратичная ошибка.
 Представлены графики изменения ошибки на обучающем (Train), тестовом (Test) и проверочном (Validation) множествах.

Figure 10 – Schedule of change of the quadratic error.

Для оценки точности классификации (в данном случае по классам: «свой-чужой») использована матрица неточностей.

Таблица 4 – Матрица неточностей

Table 4 – Inaccuracy matrix

Количество образцов	«свой»	«чужой»
«свой»	985	1
«чужой»	1	263

Итоговый классификатор обладает чувствительностью 0,9990 и специфичностью 0,9962.

4. Анализ результатов

Оценка эффективности программной реализации алгоритмов анализа видеоданных на натуральных данных показала корректность классификации в 97 % случаев после обобщения результатов перекрестной проверки.



Рисунок 11 – Распознавание позы с помощью библиотеки OpenPose

Figure 11 – Recognizing Poses Using the OpenPose Library




					
Тип динамики	Вероятность наличия в текущем фрейме, %	Тип динамики	Вероятность наличия в текущем фрейме, %	Тип динамики	Вероятность наличия в текущем фрейме, %
JUMPING	4,45	JUMPING	5,44	JUMPING	4,87
JUMPING_JACKS	4,14	JUMPING_JACKS	10,97	JUMPING_JACKS	8,97
BOXING	2,87	BOXING	2,43	BOXING	15,42
WAVING_2HANDS	9,78	WAVING_2HANDS	68,78	WAVING_2HANDS	7,62
WAVING_1HAND	77,79	WAVING_1HAND	11,14	WAVING_1HAND	5,87
CLAPPING_HANDS	0,97	CLAPPING_HANDS	1,24	CLAPPING_HANDS	57,25

Рисунок 12 – Примеры работы классификатора типа динамики по видеофрагментам

Figure 12 – Examples of the operation of a classifier of the type of dynamics by video fragments

Для снижения размерности вектора признаков в задаче идентификации по изображению лица применялся метод главных компонент для сжатия размерности пространства признаков. Оценка точности классификации 5 субъектов на два класса «свой» и «чужой» производилась методом перекрестной проверки и составила 99 % на тестовой выборке.

Заключение

Анализ существующих методов обработки данных в системах видеонаблюдения промышленного объекта показал, что применение бесконтактных способ анализа позы и действий человека в поле зрения камеры встречается редко, однако может иметь критическое значение в определенных ситуациях (человек в спецодежде лежит в зоне видимости камеры, но система на него не реагирует, ведь он находится не в запрещенной зоне).

Использование интеллектуальных камер и датчиков, в сочетании с человеком-оператором, с которого снята большая часть аналитической и зрительной нагрузки, позволяет увеличить эффективность видеонаблюдения и, как результат, повысить безопасность и результативность труда на производстве в целом.

Оценка эффективности программной реализации алгоритмов анализа натуральных видеоданных показала корректность классификации в 97 % случаев. Оценка эффективности классификации 5 субъектов на два класса «свой» и «чужой» производилась методом перекрестной проверки и показала точность 99 % на тестовой выборке.

Благодарности

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 20-08-00668

ЛИТЕРАТУРА

1. Жукова П.Н., Насонова В.А., Прокопенко А.Н. Обеспечение безопасности на объектах транспортной инфраструктуры посредством использования систем видеонаблюдения и видеоаналитики. *Проблемы правоохранительной деятельности*. 2015;4:91-96.
2. Забашта А.Ю., Скорикова С.А. Функции видеоаналитики, анализ архитектур систем видеоаналитики. *Ростовский научный журнал*. 2017;7:194-200.
3. Центр 2М – российский информационный оператор [Электронный ресурс]. URL: <https://center2m.ru/> (дата обращения: 01.04.2020).
4. ISS – Интеллектуальные Системы Безопасности [Электронный ресурс]. URL: <https://iss.ru/> (дата обращения: 01.04.2020).
5. Automated Detection and Alerting in Seconds with Visual Confirmation of Events [Электронный ресурс]. URL: <https://intelliviewtech.com/solutions/video-analytics/> (дата обращения: 01.04.2020).
6. Марр Д. Зрение: Информационный подход к изучению представления и обработки зрительных образов. М.: «Радио и связь». 1987:2.
7. Toshev A., Szegedy C. Deeppose: Human pose estimation via deep neural networks. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2014:1653-1660.
8. Bulat A., Tzimiropoulos G. Human pose estimation via convolutional part heatmap regression. *European Conference on Computer Vision*. Springer, Cham. 2016:717-732.
9. Krizhevsky A., Sutskever I., Hinton G. E. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*. 2012:1097-1105.

10. Ofli F. et al. Berkeley mhad: A comprehensive multimodal human action database. *2013 IEEE Workshop on Applications of Computer Vision (WACV). IEEE.* 2013:53-60.
11. Greff K. et al. LSTM: A search space odyssey. *IEEE transactions on neural networks and learning systems.* 2016;28(10):2222-2232.
12. Cao Z. et al. OpenPose: realtime multi-person 2D pose estimation using Part Affinity Fields. *arXiv preprint arXiv:1812.08008.* 2018.
13. Буй Т.Т.Ч., Фан Н.Х., Спицын В.Г. Распознавание лиц на основе применения метода Виолы-Джонса, вейвлет-преобразования и метода главных компонент. *Известия Томского политехнического университета.* 2012;320(5):54-59.
14. Чуйков А.В., Вульфин А.М., Васильев В.И. Нейросетевая система преобразования биометрических признаков пользователя в криптографический ключ. *Доклады ТУСУР.* 2018;21(3):35-41.
15. Чуйков А.В., Вульфин А.М. Система распознавания жестов на основе нейросетевых технологий. *Вестник Уфимского государственного авиационного технического университета.* 2017;21(3 (77)):113-122.

REFERENCES

1. Zhukova P.N., Nasonova V.A., Prokopenko The transport infrastructure safety by the use of video surveillance and video analytics. *Problems of law-enforcement activity.* 2015;4:91-96. (In Russ)
2. Zabashta A.Ju., Skorikova S.A. Video analytics functions, analysis of video analytics systems architectures. *Rostov Scientific Journal.* 2017;7:194-200. (In Russ)
3. Center 2M - Russian information operator. Available at: <https://center2m.ru/> (accessed 01.04.2020). (In Russ)
4. ISS – Intelligent Security Systems. Available at: <https://iss.ru/> (accessed 01.04.2020). (In Russ)
5. Automated Detection and Alerting in Seconds with Visual Confirmation of Events. Available at: <https://intelliviewtech.com/solutions/video-analytics/> (accessed 01.04.2020).
6. Marr D. *Vision: An Informational Approach to Studying the Representation and Processing of Visual Images.* M.: «Radio i svjaz». 1987:2. (In Russ)
7. Toshev A., Szegedy C. Deeppose: Human pose estimation via deep neural networks. *Proceedings of the IEEE conference on computer vision and pattern recognition.* 2014:1653-1660.
8. Bulat A., Tzimiropoulos G. Human pose estimation via convolutional part heatmap regression. *European Conference on Computer Vision. Springer, Cham.* 2016:717-732.
9. Krizhevsky A., Sutskever I., Hinton G. E. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems.* 2012:1097-1105.
10. Ofli F. et al. Berkeley mhad: A comprehensive multimodal human action database. *2013 IEEE Workshop on Applications of Computer Vision (WACV). IEEE.* 2013:53-60.
11. Greff K. et al. LSTM: A search space odyssey. *IEEE transactions on neural networks and learning systems.* 2016;28(10):2222-2232.
12. Cao Z. et al. OpenPose: realtime multi-person 2D pose estimation using Part Affinity Fields. *arXiv preprint arXiv:1812.08008.* 2018.
13. Buoy T.T.Ch., Fan N.Kh., Spitsyn V.G. Face recognition based on the application of the Viola-Jones method, wavelet transform and principal component analysis. *News of Tomsk Polytechnic University.* 2012;320(5):54-59. (In Russ)

14. Chuikov A.V., Vulfin A.M., Vasilyev V.I. Neural network system for converting the user biometric characteristics into a cryptographic key. *TUSUR report*. 2018;21(3):35-41. (In Russ)
15. Chuikov A.V., Vulfin A.M. Gesture recognition system. *Vestnik UGATU*. 2017;21(3 (77)):113-122. (In Russ)

ИНФОРМАЦИЯ ОБ АВТОРЕ / INFORMATION ABOUT THE AUTHOR

Вульфин Алексей Михайлович, доцент кафедры вычислительной техники и защиты информации Уфимского государственного авиационного технического университета, Уфа, Российская Федерация.

e-mail: vulfin.alexey@gmail.com

ORCID: [0000-0002-9358-0651](https://orcid.org/0000-0002-9358-0651)

Aleksey M. Vulfin, Associate Professor, Department of Computer Engineering and Information Protection, Ufa State Aviation Technical University, Ufa, Russian Federation.