

10. Bulat I.V. i dr. Sovremennoe sostoyanie metodov mashinnogo obucheniya v prilozheniyakh obrabotki odnomernykh signalov [Current state of machine learning methods in one-dimensional signal processing applications], *Elektronika, fotonika i kiberfizicheskie sistemy* [Electronics, Photonics and Cyber-Physical Systems], 2022, Vol. 2, No. 3, pp. 86-95.
11. Ivanov D.V. i dr. Novye vozmozhnosti sistem shirokopolosnoy kognitivnoy svyazi, rabotayushchikh v ionosfernykh KV-radiokanalakh s vnutrimodovoy dispersiey [New capabilities of broadband cognitive communication systems operating in ionospheric HF radio channels with intramode dispersion], *Radiotekhnika* [Radio Engineering], 2022, Vol. 86, No. 11, pp. 162-177.
12. Konkin N.A. Metodika i algoritm opredeleniya periodov operativnogo prognozirovaniya dinamiki maksimal'no primenimykh chastot KV-svyazi na osnove algoritma mashinnogo obucheniya XGBoost [Methodology and algorithm for determining the periods of operational forecasting of the dynamics of the maximum applicable frequencies of HF communication based on the XGBoost machine learning algorithm], *Vestnik Povolzhskogo gosudarstvennogo tekhnologicheskogo universiteta. Seriya: radiotekhnicheskie i infokommunikatsionnye sistemy* [Bulletin of the Volga Region State Technological University. Series: Radio Engineering and Infocommunication Systems], 2022, No. 3, pp. 6-16.
13. Ivanov D.V. i dr. Obespechenie predel'noy shirokopolosnosti sistem kognitivnoy nazemnoy i trasionosfernoy radiosvyazi, rabotayushchikh v usloviyakh dispersionnykh iskazheniy [Ensuring the maximum bandwidth of cognitive terrestrial and transionospheric radio communication systems operating under conditions of dispersion distortions], *Rasprostraneniye radiovoln* [Radio Wave Propagation], 2023, pp. 5-13.
14. Kazantseva A.A. Primeneniye metodov glubokogo mashinnogo obucheniya dlya prognozirovaniya dostupnosti radiokanalov KV-svyazi [Application of deep machine learning methods for predicting the availability of HF radio channels], *Sb. dokladov I Vserossiyskoy molodezhnoy nauchnoy shkoly-konferentsii* [Collection of reports of the 1st All-Russian Youth Scientific School-Conference], 2024, pp. 107-110.
15. Kirillov A.S. Algoritmy demodulyatsii dlya kanala svyazi Vattersona s izvestnymi parametrami [Demodulation algorithms for the Watterson communication channel with known parameters], *Radiotekhnika* [Radio Engineering], 2025, Vol. 89, No. 4, pp. 43-50.
16. Kirillov A.S. Machine Learning Algorithms for Signal Demodulation and Determination of HF Communication Channel Parameters, *IEEE Radiation and Scattering of Electromagnetic Waves (RSEMW)*, 2023, pp. 268-271.
17. Kirillov A.S. Opredeleniye parametrov modeli Vattersona s pomoshch'yu adaptivnykh fil'trov [Determining the parameters of the Watterson model using adaptive filters], *Radiotekhnika* [Radio Engineering], 2024, No. 10, pp. 5-13.
18. Kirillov A.S. Algoritmy mashinnogo obucheniya priem signala tsifrovyykh radiolinii [Machine learning algorithms for receiving digital radio lines], *Radiolokatsiya i radiosvyaz'* [Radar and Radio Communications], 2022, No. 15, pp. 265.
19. Recommendation ITU-R F.1487, Testing of HF modems with bandwidths of up to about 12 kHz using ionospheric channel simulators, 2000.
20. Recommendation ITU-R F.520, Use of high frequency ionospheric channel simulators, 1992.

Кириллов Андрей Сергеевич – МФТИ (Национальный исследовательский университет); e-mail: Andrey.kir1997@yandex.ru; г. Долгопрудный, Россия; тел.: +79161497179.

Kirillov Andrey Sergeevich – Moscow Institute of Physics and Technology (National Research University); e-mail: Andrey.kir1997@yandex.ru; Dolgoprudny, Russia; phone: +79161497179.

УДК 004.93

DOI 10.18522/2311-3103-2026-3-105-113

А.Е. Колоденкова, М.О. Бочкарев

КОМПЛЕКСНЫЙ ПОДХОД К РАСПОЗНАВАНИЮ НАРУШИТЕЛЯ ПО ВИДЕОИЗОБРАЖЕНИЯМ

Распознавание нарушителей в условиях неконтролируемой среды является важнейшей функцией систем контроля безопасности (СКБ), основанных на биометрии и обеспечивающих защиту на объектах с массовым скоплением людей. Повысить точность подобных систем возможно за счет объединения нескольких различных биометрических признаков, полученных по видеоизображениям. Однако при работе с видеоизображениями возникает ряд проблем, связанных с изменением точки обзора, контрастности, освещения, окклюзией, посторонних объектов на заднем плане, с выбором уровня объединения биометрических признаков. Для решения данной проблемы предложен комплексный подход к распознаванию нарушителя, основанный на предварительной обработке видеоизображений, а также интеграции двух сверточных нейронных сетей (СНС), теории Демпстера-Шеффера, метода случайного леса для

объединения поведенческих биометрических признаков на уровне оценок соответствия и классификации видеоизображений. В качестве биометрических признаков рассматриваются походка и жестикация, поскольку их можно наблюдать без прямого взаимодействия с человеком. Для решения задачи классификации определено три класса видеоизображений: класс 1 – человек не является нарушителем, класс 2 – потенциальный нарушитель, класс 3 – нарушитель. В работе также представлена обобщенная схема комплексного подхода к распознаванию нарушителя с подробным описанием этапов. Для оценки эффективности предложенного комплексного подхода были проведены эксперименты с набором данных KTH Dataset, который содержит шесть типов простых действий человека, совершенных разными людьми на различных фонах. Результаты исследований показали, что предложенный подход повышает точность распознавания нарушителя по видеоизображениям в условиях неконтролируемой среды (точность классификации составила 87%).

Объединение биометрических признаков; видеоизображения; сверточная нейронная сеть; теория Демпстера-Шефера; метод случайного леса.

A.E. Kolodenkova, M.O. Bochkarev

COMPREHENSIVE APPROACH TO INTRUDER RECOGNITION BASED ON VIDEO IMAGERY

Intruder recognition in uncontrolled environments is a critical function of biometric-based security control systems (SCS), ensuring protection at facilities with large crowds. The accuracy of such systems can be improved by combining multiple biometric traits extracted from video imagery. However, processing video data involves several challenges, including viewpoint variation, occlusion, and selecting an appropriate feature fusion level. To resolve these challenges, a complete approach to intruder recognition is proposed. The approach is based on video preprocessing, the combination of two convolutional neural networks (CNNs), Dempster-Shafer theory, and the random forest method. These techniques fuse behavioral biometric features at the score level to classify video sequences. Gait and gestural behavior are selected as biometric modalities, as they can be captured without direct subject interaction. For the classification task, three classes of video data are defined: class 1 – person is not intruder, class 2 – a potential intruder, class 3 – intruder. The study also presents a general architecture for the proposed approach, along with a detailed description of its processing stages. The effectiveness of the approach is measured through experiments performed on the KTH dataset, which comprises six types of simple human actions performed by different subjects under different background conditions. Experimental results show that the proposed approach improves intruder recognition accuracy in uncontrolled environments, achieving an 87 % classification rate.

Biometric feature fusion; video imagery; convolutional neural network; Dempster-Shafer theory; random forest method.

Введение. Задача распознавания нарушителя по видеоизображениям в настоящее время является популярной, сложной и широко используется в системах контроля безопасности (СКБ) во всем мире. Однако большинство данных систем далеки от совершенства, поскольку используют один биометрический признак для распознавания людей и соответственно имеют ряд ограничений при их использовании (неуниверсальность биометрических признаков, зашумленные данные, чувствительность к атакам) [1, 2]. Для преодоления ограничений в последнее время в СКБ распознавание человека осуществляется за счет объединения нескольких биометрических признаков с использованием различных алгоритмов объединения признаков. Подобные системы обеспечивают максимальную независимость между биометрическими шаблонами, следовательно, низкое качество одного биометрического признака никак не влияет на распознавание нарушителя с помощью другого признака, и точность распознавания остается при этом высокой [3–6].

При распознавании нарушителя используются такие специфические биометрические признаки, как изображение лица, походка, жестикация, тембр голоса, отпечатки пальцев и даже рисунок радужной оболочки глаза. Однако одними из самых популярных и весьма значимыми динамическими признаками для криминологии и сотрудников полиции являются движение тела, а именно походка и жестикация [7, 8]. Это связано с тем, что распознавание нарушителя по походке и жестикации можно выполнять по видеоизображениям дистанционно, без какого-либо физического контакта с человеком.

Несмотря на большое количество работ отечественных и зарубежных ученых [9–11], проблема распознавания человека по видеоизображениям в условиях неконтролируемой среды до сих пор остается открытой. Во-первых, большинство методов распознавания человека ориентированы на использование лишь одного биометрического признака, что может привести к ложному распознаванию. Во-вторых, многие методы ориенти-

рованы на видеоизображения хорошего качества, тем самым отклоняются видеоизображения, полученные в условиях неконтролируемой среды, что может привести к потере важной информации о нарушителе. В-третьих, большинство методов ориентированы на объединение статических биометрических признаков, поскольку поведенческие биометрические признаки человека непостоянны и вариативны.

Выраженный системный характер проблемы определяет необходимость в разработке комплексного подхода к распознаванию нарушителя по походке и жестикуляции, основанный на предварительной обработке видеоизображений, а также интеграции сверточных нейронных сетей (СНС), теории Демпстера-Шефера (Dempster-Shafer theory), метода случайного леса для объединения поведенческих биометрических признаков и классификации видеоизображений, полученных в условиях неконтролируемой среды.

Обобщенная схема комплексного подхода к распознаванию нарушителя. В настоящее время существует множество определений понятия и типологии нарушителя. Однако в работе под нарушителем понимается человек, предпринявший попытку выполнения запрещенных действий по ошибке, незнанию (например, шумное поведение) или осознанно со злым умыслом (например, устрашение, насилие) и находящийся в состоянии повышенного эмоционального напряжения (например, нервозность, частое осматривание по сторонам и др.) [12].

Наборы данных, используемые при распознавании человека по видеоизображениям, обычно подразделяются на два типа: получены в контролируемой и неконтролируемой среде.

В наборах данных, полученных в условиях контролируемой среды, видеоизображения хорошего качества (высокий уровень освещенности, яркости). При этом в кадре может находиться один человек, например, на пунктах досмотра или контрольно-пропускных пунктах. В результате добиться высокой точности распознавания нарушителя относительно просто, даже без применения сложных моделей, поэтому СКБ работают эффективно.

И наоборот, если наборы данных получены в условиях неконтролируемой среды, например, при низком разрешении, контрастности, освещении, наличии шума и т.д. При этом в кадре может находиться много людей, хаотично исполняющих необычные движения (например, зоны выдачи багажа в аэропортах, а также выходы и входы на станциях метро) и меняющих мимику лица. Все это приводит к тому, что достижение высокой точности распознавания нарушителя становится более сложной задачей.

В связи с этим для работы с наборами данных, полученных в условиях неконтролируемой среды, предлагаются различные методы машинного обучения, которые значительно повышают точность и способность к обобщению [13, 14].

В основе предлагаемого подхода к распознаванию нарушителя по походке и жестикуляции в условиях неконтролируемой среды лежит использование обобщенной схемы, состоящей из восьми этапов и представленной на рис. 1.

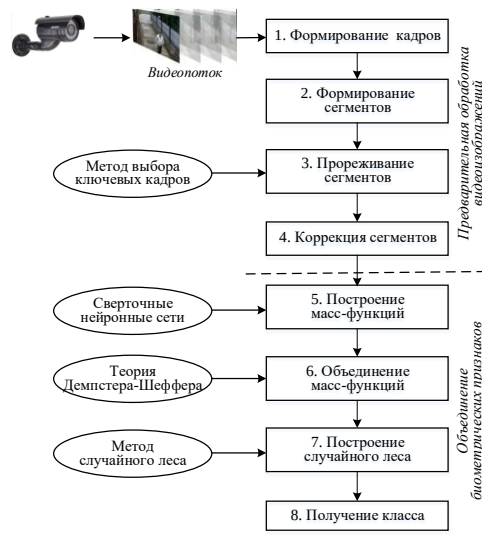


Рис. 1. Обобщенная схема распознавания нарушителя по походке и жестикуляции

Этап 1. Формирование кадров (кадрирование). Как правило, человек присутствует в поле зрения камеры видеонаблюдения в течение некоторого времени, поэтому его изображение можно отследить в последовательности кадров (видеоряд, видеопоток) (рис. 2)

$$V = \{f_1, f_2, \dots, f_k\},$$

где k – число кадров; f_i – текущий кадр видео; $H \times W \times C$ – размерность кадра, H – высота изображения в пикселях, W – ширина изображения в пикселях, C – количество каналов, которые подвергаются первичной буферизации данных для последующей обработки.

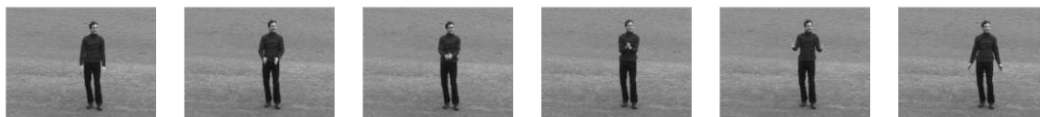


Рис. 2. Пример кадрирования действий человека в видеопотоке

Этап 2. Формирование сегментов. Поскольку кадры из разных сцен имеют разные временные и пространственные характеристики [15], то сегментация видеопотока V на Q сегментов (смысловые, важные сцены) является важной частью анализа видеoinформации

$$Q(V) = \{q_1, q_2, \dots, q_h\},$$

где $q_j = \{f_{j1}, f_{j2}, \dots, f_{jg}\}$ – сегмент (отрезок) с g кадрами, каждый сегмент содержит фиксированное число кадров (длина сегмента) (рис. 3).

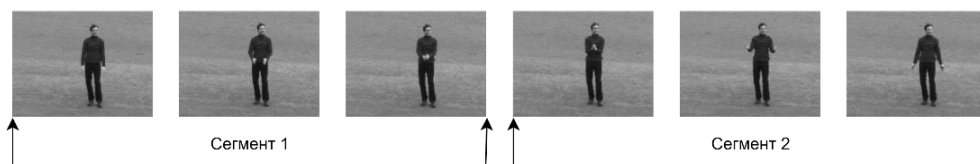


Рис. 3. Формирование сегментов

Сегментация видеопотока позволяет значительно снизить вычислительные требования, а также ускорить и упростить процесс вычислений.

Этап 3. Прореживание сегментов. Поскольку методы обработки видеоизображений обычно требуют фиксированной длины видеопотока и размеров кадра, а обучающие видеоизображения имеют переменную длину, то крайне важно выбирать ключевые кадры, которые охватывают максимум информации, связанной с действиями человека [16, 17]. Следовательно, оптимальный выбор ключевых кадров позволяет свести к минимуму избыточность информации в системах обработки видеоизображений.

На данном этапе происходит исключение некоторых кадров из сегмента с применением метода выбора ключевых кадров, которое позволяет одновременно снизить вычислительную нагрузку и уменьшить объем потребляемой памяти. Предлагаемый метод сочетает в себе обнаружение изменений на кадрах с использованием яркости и адаптивной фильтрации с пороговыми значениями и будет подробно рассмотрен в следующих работах авторов.

На рис. 4 приведен пример прореживания сегментов.

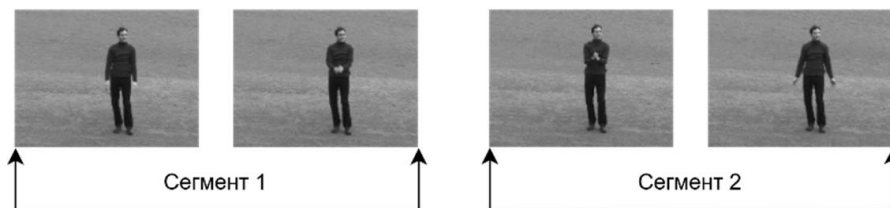


Рис. 4. Пример прореживания сегментов

Этап 4. Коррекция сегментов. На данном этапе осуществляется коррекция яркости и контраста изображений с использованием алгоритма [18], который был адаптирован для коррекции сегментов, поскольку видеопоток можно легко представить в виде набора изображений.

На рис. 5 приведен пример коррекции сегментов, на котором силуэт человека значительно выделяется на общем фоне, что достигается за счет изменений яркости и контраста, по сравнению с рис. 4.

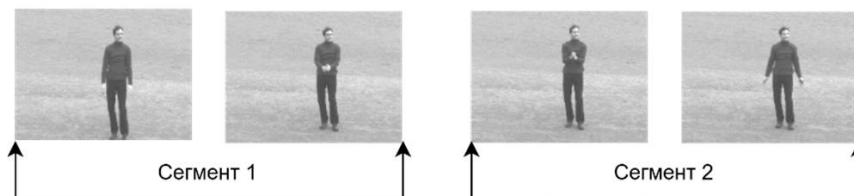


Рис. 5. Пример коррекции сегментов

Этап 5. Построение масс-функций. После предварительной обработки видеоизображений сформированный набор данных подается на вход СНС для изучения и извлечения биометрических признаков (походка и жестикация). Сверточная нейронная сеть обучается отдельно для каждого биометрического признака, после чего выходные данные СНС преобразуются в распределения вероятностей по трем классам (класс 1 – человек не является нарушителем, класс 2 – потенциальный нарушитель, класс 3 – нарушитель) с помощью метода softmax, которые могут рассматриваться с некоторым допущением как массы-функций $m_1(S_{\text{походка}})$ и $m_2(S_{\text{жесты}})$.

$$m_p(\{S_l\}) = \frac{\exp(Z_{lp})}{\sum_{l=1}^C Z_{lp}}, l = 1, \dots, C,$$

где Z_{lp} – l -е апостериорное значение вероятности, полученное от p -ой СНС, C – количество классов ($l = 3$).

Для стабилизации обучения используется *Batch Normalization*, для предотвращения переобучения используется *Dropout*.

Этап 6. Объединение масс-функций. В предлагаемом подходе в качестве экспертов для распознавания нарушителя и получения первоначальных результатов распознавания выбрана СНС. Поскольку у каждого эксперта свой взгляд на классификацию видеоизображения, то могут возникнуть некоторые разногласия (проблема конфликта). Для решения данной проблемы предложено использовать теорию Демпстера-Шефера.

Для объединения масс-функций используется правило объединения Демпстера-Шефера [19]

$$m_{12}(S) = \frac{1}{1 - K} \sum_{S_{\text{походка}} \cap S_{\text{жесты}} = \emptyset} m_1(S_{\text{походка}}) \cdot m_2(S_{\text{жесты}}),$$

где K – мера конфликта, которая интуитивно измеряет, насколько сильно $m_1(S_{\text{походка}})$ и $m_2(S_{\text{жесты}})$ противоречат друг другу.

Объединение биометрических признаков осуществляется на уровне оценок соответствия, что улучшает процесс принятия решений за счет интеграции различных источников информации, тем самым снижая вероятность ложного принятия или отклонения. Данное объединение используется в дальнейшем для классификации с использованием метода случайного леса.

Этап 7. Построение случайного леса. Для преобразования объединенных масс-функций в окончательное решение применяется метод случайного леса (ансамблевый метод), основанный на построении ансамбля деревьев решений T_n . Это связано с тем, что он позволяет обрабатывать тысячи входных переменных без их удаления, а также ответить на вопрос, какие переменные важны для классификации.

При обучении деревьев решений использовался алгоритм бутстрэп (*bootstrap*). Хотя количество классов сравнительно небольшое ($l = 3$), но деревья решений обеспечивают устойчивость к шуму при «переобучении», а метод случайного леса обеспечивает стабильную классификацию.

В качестве критерия останова деревьев решений рекомендуется использовать, например, ограничение на максимальную глубину дерева, максимальное число листьев, минимальное число объектов в листе.

Этап 8. Получение класса. При решении задачи классификации используется мажоритарное голосование: каждое дерево решений «голосует» за определенный класс, и итоговым классом становится класс, который наиболее часто встречается.

Объединение биометрических признаков, основанное на интеграции СНС, теории Демпстера-Шефера и методе случайного леса, далее именуется ансамблевой моделью.

Результаты исследования. Для дальнейшей оценки эффективности предложенного комплексного подхода были проведены эксперименты с набором данных *KTH Dataset* [20], который содержит видеопотоки с различными действиями людей (размахивание руками, боксирование, бег, ходьба и др.), записанных с разных ракурсов.

Для классификации видеоизображений были перераспределены метки набора данных следующим образом: ходьба – человек не является нарушителем; размахивание руками, пробежка – потенциальный нарушитель, боксирование, бег – нарушитель.

Для изучения и извлечения биометрических признаков (походка, жестикуляция) использовались СНС, обучающиеся индивидуально, на одинаковом наборе данных, который был разделен (70% – тестовая выборка, 15% – валидационная выборка, 15% – тестовая выборка).

Длина сегмента видеопотока была установлена – 16 кадров, в результате чего получилось следующее распределение сегментов по классам: класс 1 – 1507 сегментов; класс 2 – 959 сегментов; класс 3 – 3932 сегмента.

Поскольку полученный набор сегментов несбалансирован, целесообразно использовать технику *oversampling* (дополнение каждого класса дублирующими данными). После *oversampling* сбалансированный набор данных содержит 11796 сегментов, распределенных следующим образом: класс 1 – 3932 сегмента; класс 2 – 3932 сегмента; класс 3 – 3932 сегмента.

После обучения ансамблевой модели была проведена оценка с помощью тестовой выборки, на которой она не обучалась, содержащая 591 сегмент (класс 1 – 191 сегмент; класс 2 – 202 сегмента, класс 3 – 198 сегментов).

В табл. 1 представлены результаты распознавания нарушителя по походке и жестикуляции с использованием разработанного комплексного подхода и подхода, представленного в работе [21].

Таблица 1

Результаты распознавания нарушителя

Количество сегментов	Верно классифицированные сегменты с использованием комплексного подхода	Ошибочно классифицированные сегменты с использованием комплексного подхода	Верно классифицированные сегменты с использованием известного подхода	Ошибочно классифицированные сегменты с использованием известного подхода
591	514	77	443	148

Тестирование ансамблевой модели показало, что она с точностью до 87% правильно определяет человека по видеоизображениям.

Также был проведен сравнительный анализ точности различных моделей классификации видеоизображений, предложенных различными исследователями: CNN + Tree (Gait + Gesture) [21] – 75%, HOG + Classification Tree (Gait) [22] – 58%; CNN (Gait + Gesture) – 61%.

Сравнивая результаты тестирования можно сделать вывод о высоком качестве разработанной ансамблевой модели и эффективности предложенного подхода.

Заключение. В настоящей работе предлагается комплексный подход к распознаванию нарушителя по походке и жестике, основанный на предварительной обработке видеоизображений, а также интеграции СНС, теории Демпстера-Шефера, метода случайного леса с целью достижения высокой надежности распознавания нарушителя по видеоизображениям. Тестирование ансамблевой модели показало, что она с точностью до 87 % правильно определяет человека по видеоизображениям в условиях неконтролируемой среды. Предложенный подход позволит оперативно сотрудникам, обеспечивающим внутриобъектовый и пропускной режим (частным охранным предприятиям) осуществлять безошибочный поиск нарушителей в условиях неконтролируемой среды, тем самым предупреждая преступления.

Дальнейшие исследования авторов будут направлены на разработку моделей объединения других биометрических признаков и собственного набора данных.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Jain Anil, Nandakumar Karthik, Ross Arun 50 years of biometric research: accomplishments, challenges, and opportunities // Pattern Recognition Letters. – 2016. – Vol. 79. – P. 80-105. – DOI: 10.1016/j.patrec.2015.12.013.
2. Колоденкова А.Е. Онтология идентификации нарушителя по движениям тела и лицу в видеонаблюдениях // Онтология проектирования. – 2023. – № 1. – С. 55-74. – DOI: <https://doi.org/10.18287/2223-9537-2023-13-1-55-74>.
3. Pahuja S., Goel N. Multimodal biometric authentication: A review // AI Communications. – 2024. – Vol. 37, No. 4. – P. 525-547. – DOI: 10.3233/AIC-220247.
4. Xia F., Wei X., Q. Zhao, Jin B., Yu Z., Zha Z. Exploration on the application of multimodal biometric feature recognition in system access control // 2024 IEEE 13th International Conference on Communication Systems and Network Technologies (CSNT). – 2024. – P. 1-6. – DOI: 10.1109/CSNT60213.2024.10545831.
5. Орлов А.А., Миронов М.И., Абрамова Е.С. Обзор и анализ подходов и практических областей применения распознавания эмоций человека // Вестник Южно-Уральского государственного университета. Серия: Компьютерные технологии, управление, радиоэлектроника. – 2023. – Т. 23, № 4. – С. 5-15. – DOI: 10.14529/ctcr230401.
6. Rangel G., Cuevas-Tello J.C., Nunez-Varela J., Puente C., Silva-Trujillo A.G. A survey on convolutional neural networks and their performance limitations in image recognition tasks // Journal of sensors. – 2024. – No. 1. – P. 1-29. – DOI: 10.1155/2024/2797320.
7. Сафонов А.А., Булгаков В.Г., Варченко И.А. Криминалистическое исследование динамических признаков человека: история и современное состояние // Общество и право. – 2010. – № 3. – С. 250-257.
8. Кобец П.Н. О комплексном изучении личности преступника в отечественной криминологии. Проблемы развития личности: Матер. междунар. науч.-практ. конф. Прага. – 2013. – С. 93-95.
9. Sprager S., Juric M.B. Inertial sensor-based gait recognition: A Review // Sensors. – 2015. – Vol. 15. – P. 22089-22127. – DOI: 10.3390/s150922089.
10. Hou L., Wan W., Han K., Muhammad R., Yang M. Human detection and tracking over camera networks: A review // 2016 International Conference on Audio, Language and Image Processing (ICALIP). – 2016. – P. 574-580. – DOI: 10.1109/ICALIP.2016.7846643.
11. Li W., Hou S., Zhang C., Cao C., Liu X., Huang Y., Zhao Y. An in-depth exploration of person re-identification and gait recognition in cloth-changing conditions // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. – 2023. – P. 13824-13833. – DOI: 10.1109/CVPR52729.2023.01328.
12. Федюнина А.П. Выявление характерологических признаков и составление психологического портрета возможного нарушителя и лояльного сотрудника в сфере информационной безопасности // Вестник АГТУ. – 2007. – № 4. – С. 231-236.
13. Al-Azawi R.J. Enhancing recognition accuracy and efficiency through intelligent frame selection in uncontrolled conditions // J Opt. – 2024. – P. 1-7. – DOI: <https://doi.org/10.1007/s12596-024-02191-4>.
14. Zhao L. A Hybrid deep learning-based intelligent system for sports action recognition via visual knowledge discovery // IEEE Access. – 2023. – Vol. 11. – P. 46541-46549. – DOI: 10.1109/ACCESS.2023.3275012.
15. Zhou J., Jiangbo Ai, Zheng Wang, Chen S., Wei Q. Discovering attractive segments in the user generated video streams // Lecture Notes in Computer Science. – 2019. – Vol. 11642. – P. 236-250. – DOI: https://doi.org/10.1007/978-3-030-26075-0_18.
16. Mingju Chen, Han Xiaofeng, Zhang Hua, Lin Guojun, Kamruzzaman M.M. Quality-guided key frames selection from video stream based on object detection // Journal of Visual Communication and Image Representation. – 2019. – Vol. 65. – P. 1-7. – DOI: <https://doi.org/10.1016/j.jvcir.2019.102678>.

17. *Gharabagh Abdorreza, Hajhashemi Vahid, Ferreira Marta, Machado José, Tavares Joao* Best Frame Selection to Enhance Training Step Efficiency in Video-Based Human Action Recognition // *Applied Sciences*. – 2022. – Vol. 12. – P. 1-12. – DOI: 10.3390/app12041830.
18. *Трофименко Я.М., Виноградова Л.Н., Ершов Е.В.* Алгоритмы предобработки изображений в системе идентификации лиц в видеопотоке // *Вестник Череповецкого государственного университета*. – 2019. – № 4. – С. 21-29. – DOI: 10.23859/1994-0637-2019-4-91-2.
19. *Kumari P., Reddy S.R.N., Yadav R.* Application of Dempster-Shafer Theory in Sensor Data Fusion // *Proceedings of the International Conference on Machine Learning, Deep Learning and Computational Intelligence for Wireless Communication*. – 2024. – P. 91-102. – DOI: https://doi.org/10.1007/978-3-031-47942-7_9.
20. Recognition of human actions. – Режим доступа: <https://www.csc.kth.se/cvap/actions/> (дата обращения: 24.02.2026).
21. Computer-aided identification of degenerative neuromuscular diseases based on gait dynamics and ensemble decision tree classifiers. – Access mode: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0252380> (дата обращения: 24.02.2026).
22. *Zhang Z.* A gait recognition method for a moving target image in sports based on a decision tree // *International Journal of Biometrics*. – 2021. – Vol. 13, No. 2. – P. 165-179. – DOI: 10.1504/IJBM.2021.114642.

REFERENCES

1. *Jain Anil, Nandakumar Karthik, Ross Arun* 50 years of biometric research: accomplishments, challenges, and opportunities, *Pattern Recognition Letters*, 2016, Vol. 79, pp. 80-105. DOI: 10.1016/j.patrec.2015.12.013.
2. *Kolodenkova A.E.* Ontologiya identifikatsii narushitelya po dvizheniyam tela i lits v videonablyudeniakh [Ontology of human identification by face and body motions in video surveillance systems], *Ontologiya proektirovaniya* [Ontology of Designing], 2023, No. 1, pp. 55-74. DOI: <https://doi.org/10.18287/2223-9537-2023-13-1-55-74>.
3. *Pahuja S., Goel N.* Multimodal biometric authentication: A review, *AI Communications*, 2024, Vol. 37, No. 4, pp. 525-547. DOI: 10.3233/AIC-220247.
4. *Xia F., Wei X., Q. Zhao, Jin B., Yu Z., Zha Z.* Exploration on the application of multimodal biometric feature recognition in system access control, *2024 IEEE 13th International Conference on Communication Systems and Network Technologies (CSNT)*, 2024, pp. 1-6. DOI: 10.1109/CSNT60213.2024.10545831.
5. *Orlov A.A., Mironov M.I., Abramova E.S.* Obzor i analiz podkhodov i prakticheskikh oblastey primeneniya raspoznavaniya emotsiy cheloveka [Review and analysis of approaches and practical applications of human emotion recognition], *Vestnik Yuzhno-Ural'skogo gosudarstvennogo universiteta. Seriya: Komp'yuternye tekhnologii, upravlenie, radioelektronika* [Bulletin of the South Ural State University. Series: Computer Technologies, Automatic Control, Radio Electronics], 2023, Vol. 23, No. 4, pp. 5-15. DOI: 10.14529/ctcr230401.
6. *Rangel G., Cuevas-Tello J.C., Nunez-Varela J., Puente C., Silva-Trujillo A.G.* A survey on convolutional neural networks and their performance limitations in image recognition tasks, *Journal of sensors*, 2024, No. 1, pp. 1-29. DOI: 10.1155/2024/2797320.
7. *Safonov A.A., Bulgakov V.G., Varchenko I.A.* Kriminalisticheskoe issledovanie dinamicheskikh priznakov cheloveka: istoriya i sovremennoe sostoyanie [Criminalistic investigation of dynamic human characteristics: history and current state], *Obshchestvo i pravo* [Society and Law], 2010, No. 3, pp. 250-257.
8. *Kobets P.N.* O kompleksnom izuchenii lichnosti prestupnika v otechestvennoy kriminologii. Problemy razvitiya lichnosti [On the complex study of the personality of a criminal in domestic criminology. Problems of personality development], *Mater. mezhdunarod. nauch.-prakt. konf. Praga, 2013* [Mater. international scientific-practical. conf. Prague, 2013], pp. 93-95.
9. *Sprager S., Juric M.B.* Inertial sensor-based gait recognition: A Review, *Sensors*, 2015, Vol. 15, pp. 22089-22127. DOI: 10.3390/s150922089.
10. *Hou L., Wan W., Han K., Muhammad R., Yang M.* Human detection and tracking over camera networks: A review, *2016 International Conference on Audio, Language and Image Processing (ICALIP)*, 2016, pp. 574-580. DOI: 10.1109/ICALIP.2016.7846643.
11. *Li W., Hou S., Zhang C., Cao C., Liu X., Huang Y., Zhao Y.* An in-depth exploration of person re-identification and gait recognition in cloth-changing conditions, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 13824-13833. DOI: 10.1109/CVPR52729.2023.01328.
12. *Fedyunina A.P.* Vyyavlenie kharakterologicheskikh priznakov i sostavlenie psikhologicheskogo portreta vozmozhnogo narushitelya i loyального сотрудника в сфере информационной безопасности [Identification of characterological signs and drawing up a psychological portrait of a possible violator and a loyal employee in the field of information security], *Vestnik AGTU* [Bulletin of AGTU], 2007, No. 4, pp. 231-236.

13. Al-Azawi R.J. Enhancing recognition accuracy and efficiency through intelligent frame selection in uncontrolled conditions, *J Opt.*, 2024, pp. 1-7. DOI: <https://doi.org/10.1007/s12596-024-02191-4>.
14. Zhao L. A Hybrid deep learning-based intelligent system for sports action recognition via visual knowledge discovery, *IEEE Acces.*, 2023, Vol. 11, pp. 46541-46549. DOI: 10.1109/ACCESS.2023.3275012.
15. Zhou J., Jiangbo Ai. Zheng Wang, Chen S., Wei Q. Discovering attractive segments in the user generated video streams, *Lecture Notes in Computer Science*, 2019, Vol. 11642, pp. 236-250. DOI: https://doi.org/10.1007/978-3-030-26075-0_18.
16. Mingju Chen, Han Xiaofeng, Zhang Hua, Lin Guojun, Kamruzzaman M.M. Quality-guided key frames selection from video stream based on object detection, *Journal of Visual Communication and Image Representation*, 2019, Vol. 65, pp. 1-7. DOI: <https://doi.org/10.1016/j.jvcir.2019.102678>.
17. Gharabagh Abdorreza, Hajhashemi Vahid, Ferreira Marta, Machado José, Tavares Joao Best Frame Selection to Enhance Training Step Efficiency in Video-Based Human Action Recognition, *Applied Sciences*, 2022, Vol. 12, pp. 1-12. DOI: 10.3390/app12041830.
18. Trofimenko Ya.M., Vinogradova L.N., Ershov E.V. Algoritmy predobrabotki izobrazheniy v sisteme identifikatsii lits v videopotoke [Algorithms for image preprocessing in the face identification system in the video stream], *Vestnik Cherepovetskogo gosudarstvennogo universiteta* [Bulletin of the Cherepovets State University], 2019, No. 4, pp. 21-29. DOI: 10.23859/1994-0637-2019-4-91-2.
19. Kumari P., Reddy S.R.N., Yadav R. Application of Dempster-Shafer Theory in Sensor Data Fusion, *Proceedings of the International Conference on Machine Learning, Deep Learning and Computational Intelligence for Wireless Communication*, 2024, pp. 91-102. DOI: https://doi.org/10.1007/978-3-031-47942-7_9.
20. Recognition of human actions. Available at: <https://www.csc.kth.se/cvap/actions/> (accessed 24 February 2026).
21. Computer-aided identification of degenerative neuromuscular diseases based on gait dynamics and ensemble decision tree classifiers. Access mode: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0252380> (accessed 24 February 2026).
22. Zhang Z. A gait recognition method for a moving target image in sports based on a decision tree, *International Journal of Biometrics*, 2021, Vol. 13, No. 2, pp. 165-179. DOI: 10.1504/IJBM.2021.114642.

Колоденкова Анна Евгеньевна – Самарский национальный исследовательский университет им. акад. С.П. Королева (Самарский университет); e-mail: anna82_42@mail.ru; г. Самара, Россия; д.т.н.; доцент; профессор кафедры «Информационные системы и технологии».

Бочкарев Михаил Олегович – Самарский национальный исследовательский университет им. акад. С.П. Королева (Самарский университет); e-mail: mikhail19997@gmail.com; г. Самара, Россия; ассистент кафедры «Информационные системы и технологии».

Kolodenkova Anna Evgenievna – Samara National Research University named after Academician S.P. Korolev (Samara University); e-mail: anna82_42@mail.ru; Samara, Russia; dr. of eng. sc.; associate professor; professor of the Department of Information Systems and Technologies.

Bochkarev Mikhail Olegovich – Samara National Research University named after Academician S.P. Korolev (Samara University); e-mail: mikhail19997@gmail.com; Samara, Russia; assistant of the Department of Information Systems and Technologies.

УДК 004.4

DOI 10.18522/2311-3103-2026-3-113-133

Д.А. Морозов, В.В. Гилка, А.С. Кузнецова

СОВРЕМЕННЫЕ ПОДХОДЫ РАСПОЗНАВАНИЯ ЛИЦ В УСЛОВИЯХ НИЗКОЙ ОСВЕЩЁННОСТИ: ОБЗОР И КОНЦЕПЦИЯ ГИБРИДНОЙ END-TO-END АРХИТЕКТУРЫ

Рассматривается проблема надёжного распознавания лиц в таких критически важных областях, как видеонаблюдение и биометрическая аутентификация в условиях недостаточной освещённости. Существующие подходы, как правило, разделяют задачи повышения качества изображения и собственно идентификации, что приводит к накоплению ошибок и потере информативных признаков. Целью работы является преодоление этого ограничения путём разработки и теоретического обоснования гибридной end-to-end архитектуры, в которой задачи улучшения изображения и распознавания лица решаются совместно. В исследовании проводится системати-