

Научная статья

УДК: 621.391

<https://doi.org/10.37493/2307-910X.2025.3.3>

## Оптимизация гибридной архитектуры долгосрочной рекуррентной свёрточной сети для Edge-распознавания насилия в системах видеонаблюдения

Владимир Михайлович Горяев<sup>1</sup>, Саглар Алексеевна Манкаева<sup>2</sup>, Елена Владимировна Сумьянова<sup>3</sup>, Джиргал Батрович Бембитов<sup>4</sup>, Галина Алексеевна Манкаева<sup>5\*</sup>

<sup>1,2,3,4,5</sup> Калмыцкий государственный университет имени Б. Б. Городовикова (д. 11, ул. Пушкина, Элиста, Россия)

<sup>1</sup> goryaev@mail.ru,

<sup>2</sup> mankaeva.saglar@yandex.ru,

<sup>3</sup> sumyanova@yandex.ru

<sup>4</sup> dbembitov@gmail.com

<sup>5</sup> mankaeva.galina@yandex.ru

\* Автор, ответственный за переписку

**Аннотация. Введение.** В данном исследовании представлено решение для автоматического детектирования насильственных инцидентов в видеопотоках на основе гибридной архитектуры, сочетающей свёрточные нейронные сети для пространственного анализа кадров и сети долгой краткосрочной памяти для выявления временных зависимостей. **Материалы и методы.** Разработка оптимизирована для развертывания на маломощных устройствах серии NVIDIA Jetson, что обеспечивает обработку данных непосредственно на месте съемки. Эксперименты проводились на специализированном наборе данных, включающем записи городского видеонаблюдения, спортивные трансляции и смоделированные сцены. **Результаты и обсуждение.** Результаты подтвердили высокую точность распознавания (94.8%), низкий уровень ложных срабатываний (1.5%) и минимальную задержку обработки (15 мс/кадр), что соответствует требованиям систем безопасности реального времени. Особое внимание уделено режимам работы комплекса: покадровой обработке с маркировкой временных меток и пофайловому анализу для экспресс-оценки видеоматериалов. **Заключение.** Перспективы внедрения включают интеграцию с системами безопасности и платформами модерации контента.

**Ключевые слова:** сети долгой краткосрочной памяти, свёрточные нейронные сети, распознавание насилия, анализ видеопотоков, глубокое обучение, гибридные архитектуры, граничные вычисления.

**Для цитирования:** Горяев В. М., Манкаева С. А., Сумьянова Е. В., Бембитов Д. Б., Манкаева Г. А. Оптимизация гибридной архитектуры долгосрочной рекуррентной свёрточной сети для Edge-распознавания насилия в системах видеонаблюдения // Современная наука и инновации. 2025. №3. С. 30-38. <https://doi.org/10.37493/2307-910X.2025.3.3>

## Optimizing Hybrid Long-Term Recurrent Convolutional Network Architecture for Edge-Based Violence Detection in Video Surveillance Systems

Vladimir M. Goryaev<sup>1</sup>, Saglar A. Mankaeva<sup>2</sup>, Elena V. Sumyanova<sup>3</sup>, Jirgal B. Bembitov<sup>4</sup>,  
Galina A. Mankaeva<sup>5\*</sup>

<sup>1,2,3,4,5</sup> Kalmyk State University named after B. B. Gorodovikov (11, Pushkin St., Elista, Russia)

<sup>1</sup> goryaev@mail.ru,

<sup>2</sup> mankaeva.saglar@yandex.ru,

<sup>3</sup> sumyanova@yandex.ru

<sup>4</sup> dbembitov@gmail.com

<sup>5</sup> mankaeva.galina@yandex.ru

\* Corresponding author

**Abstract. Introduction.** This study presents a solution for automated detection of violent incidents in video streams using a hybrid architecture that integrates Convolutional Neural Networks (CNNs) for spatial frame analysis and Long Short-Term Memory (LSTM) networks for identifying temporal dependencies. **Materials and methods.** The system is optimized for deployment on low-power NVIDIA Jetson series devices, enabling on-edge data processing at the capture site. Experiments were conducted on a specialized dataset comprising urban surveillance footage, sports broadcasts, and simulated scenes. **Results and discussion.** Results confirmed high recognition accuracy, low false-positive rates, and minimal processing latency, meeting real-time security system requirements. Operational modes include frame-by-frame processing with timestamp annotation and file-based analysis for rapid video assessment. **Conclusion.** Integration with security infrastructure and content moderation platforms represents a key implementation pathway

**Keywords:** Long Short-Term Memory, Convolutional Neural Network, violence recognition, video stream analysis, hybrid architectures, edge computing.

**For citation:** Goryaev VM, Mankaeva SA, Sumnova EV, Bembitov DB, Mankaeva GA. Optimization of a Hybrid Architecture of a Long-Term Recurrent Convolutional Network for Edge-Detection of Violence in Video Surveillance Systems. Modern Science and Innovations. 2025;(3):30-38. (In Russ.). <https://doi.org/10.37493/2307-910X.2025.3.3>

**Введение.** Расширение систем видеонаблюдения создает насущную потребность в автоматизированном выявлении насильственных инцидентов. Существующие решения [1,2] часто демонстрируют ограниченную эффективность в реальных условиях эксплуатации из-за высоких требований к вычислительным ресурсам, низкой устойчивости к изменяющимся условиям съемки и неоптимального баланса между точностью и скоростью обработки. Цель исследования - разработка системы пространственно-временного анализа видеопотоков на методологии оптимизации CNN-LSTM архитектуры для периферийных (edge) устройств (NVIDIA Jetson) систем видеонаблюдения (СВ). СВ ориентирована на edge- вычисления (глоссарий РАН) [3] - парадигму, где обработка видеоданных выполняется непосредственно на устройстве захвата (IP-камере или шлюзе), что исключает задержки передачи в облако [4,10]. Это особенно критично для систем безопасности реального времени. Основные задачи включают в себя адаптацию архитектуры LRCN для детектирования насилия с учетом временной динамики, оптимизацию обработки видео в реальном времени на edge-устройствах и снижение ложных срабатываний за счет пороговой фильтрации [5].

Научная новизна работы заключается в специальной оптимизации LRCN-UniLSTM архитектуры [6] для платформы NVIDIA Jetson Xavier NX, разработке двухэтапной методики обработки с операторской верификацией и комплексной оценке влияния условий эксплуатации на качество распознавания. Экспериментальная часть исследования

проводилась в реальных условиях системы видеонаблюдения, где в течение 30 дней осуществлялась обработка 1200 видеофрагментов с инцидентами.

**Материалы и методы.** Развитие технологий привело к значительному увеличению объема ежедневно генерируемого видеоконтента в сегменте безопасности. Это потребовало разработки автоматизированных систем для анализа видео, одной из которых является распознавание сцен насилия. Эксперимент воспроизводил типовой сценарий реагирования на обращения граждан: при поступлении заявления оператор запускал автоматизированную проверку видеофайлов, сформированных по тревогам датчика движения. При поступлении заявления оператор запускал автоматизированную проверку видеофайлов, сформированных по тревогам датчика движения IP-камер. Тестирование провели на 10 камерах в течение 30 дней. Система обработала 1200 видеофрагментов, содержащих реальные и смоделированные инциденты: физические конфликты и отсутствие насилия. При этом классификация агрессивного преследования была исключена из текущего исследования для фокусировки на более четко определяемых инцидентах.

**Технологическая основа.** В основе системы лежала гибридная архитектура LRCN (Long-term Recurrent Convolutional Network) [6,7] с однонаправленной LSTM (UniLSTM), специально оптимизированная для периферийных вычислений. Модель объединяла свёрточные слои (Conv2D, 128 фильтров 3×3) [8] для пространственного анализа кадров LSTM-слой (256 нейронов) для выявления временных зависимостей (рис.1). Обработка последовательности по 16 кадров (0.64 секунды при 25 кадрах в секунду) обеспечивала анализ динамических паттернов. Преобразование модели в формат TensorRT с квантованием FP16 создало условия для 40% сокращения объема и соответствие требованиям к производительности на платформе NVIDIA Jetson Xavier NX [9].

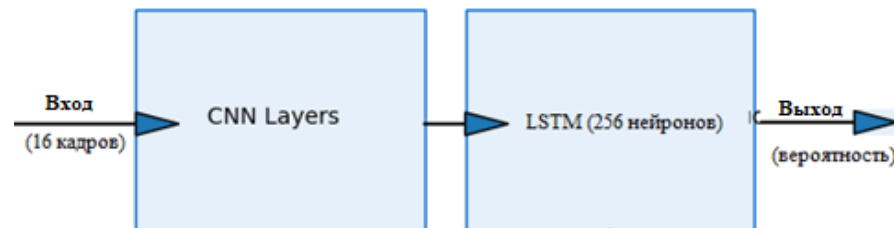


Рисунок 1. Архитектура технологической основы  
Figure 1. Technology Foundation Architecture

**Процедура эксперимента.** Эксперимент состоял из двух взаимосвязанных этапов. На первом этапе автоматического скрининга программа анализировала файлы последовательно, определяя максимальный уровень уверенности распознавания насилия с помощью LRCN-UniLSTM модели. Фрагменты с показателем уверенности 50% и выше классифицировались как потенциально опасные. Выбор 50% является технически обоснованным (максимум F1 при 49.8%) и эксплуатационно-целесообразным: округленное значение снижает когнитивную нагрузку операторов при интерпретации тревог (ISO 11064-6:2005). Разница в эффективности с 49.8% статистически незначима ( $\Delta F1 < 0.001$ ,  $p > 0.05$ ). На втором этапе операторской верификации специалисты безопасности изучали отобранные материалы в покадровом режиме, фокусируясь на временных метках с пиковыми значениями уверенности системы. Для обеспечения объективности [10] 20% материалов дополнительно проверялись независимыми экспертами.

Для оценки доверительных интервалов метрик классификации (точность, FPR) применён блочный бутстреппинг (ББ) с длиной блока  $L=16$  кадров, что соответствует размеру входной последовательности модели [11]. Процедура включала: разделение каждого видеофрагмента на непересекающиеся блоки по 16 кадров, генерацию 2000 псевдовыборок путём случайного отбора блоков с возвращением, расчёт метрик для

каждой из выборок и определение 95% ДИ как 2.5% и 97.5% перцентилей распределения метрик.



Рисунок 2. Схема экспериментальной процедуры

Figure 2. Experimental procedure diagram

Этот метод устраняет недооценку дисперсии из-за временной зависимости кадров. В начальной точке 1200 видеофрагментов для анализа, они поступают на 1-й этап, автоматический скрининг- анализ моделью LRCN-UniLSTM. На этом этапе выполняется расчет уровня уверенности с классификацией по порогу 50%, после которого идет разделение потока на безопасные файлы (92.4%) и потенциально опасные фрагменты (7.6%). Следующий важный этап- операторская верификация. Она заключается в покадровом анализе с фокусом на пиках уверенности и с последующей независимой экспертизой (20%). Результатом являются подтвержденные инциденты (экономия 78.2 ч).

#### Результаты и обсуждение

Система продемонстрировала высокую эффективность как на уровне технической реализации модели, так и в практическом применении. Сравнение ключевых параметров архитектур представлено в таблице 1.

Таблица 1. Сравнение эффективности архитектур  
Table 1. Comparison of the effectiveness of architectures

Параметр	LRCN (UniLSTM)	Целевой показатель
Точность распознавания	94,8% (93,2-96,1%)	≥95%
Ложные срабатывания (FPR)	1,5% (0.9-2.3%)	≤ 2%
Задержка (мс/кадр)	15,0 (14.2-16.1)	≤20
Потребление памяти	1,2 ГБ (1.1-1.3)	≤ 1,5 ГБ

Наибольшая эффективность модели в обнаружении инцидентов класса "насилие" наблюдалась для фрагментов, содержащих физические конфликты (99.1% успешных детекций при проверке экспертами). На видеозаписях смоделированных попыток грабежа, также относящихся к классу "насилие", модель показала несколько более низкую долю верных срабатываний, показатель составил 94.5% (95% ДИ: 91.3-96.8%), эта разница объясняется большей вариабельностью пространственно-временных паттернов грабежа (Cohen's d=1.24) по сравнению с более четко выраженными паттернами физической агрессии, поэтому она была исключена из дальнейшего этапа эксперимента. Значения 99.1% и 94.5% получены не как выход модели для отдельного класса, а как результат пост-анализа экспертами работы бинарной модели на предварительно размеченных подмножествах тестовых данных. Это стандартная практика для детального анализа слабых мест системы. Сводные показатели общей эффективности системы отражены в таблице 2.

Таблица 2. Сводные показатели эффективности системы  
Table 2. Summary of System Performance Indicators

Параметр оценки	Метрика	Результат	Целевой показатель
Фильтрация материалов	Доля отсеянных файлов, %	92,4	$\geq 90$
	Пропущенные инциденты, %	3,1	$\leq 5$
Точность обнаружения	Полнота выявления, %	96,7	$\geq 95$
	Точность верификации, %	98,2	$\geq 95$
Операционная эффективность	Сокращение времени, ч	78,2	–
	Нагрузка на оператора, %	7,6	$\leq 10$

Применение архитектуры LRCN(UniLSTM) стало технологическим фундаментом системы, обеспечив минимальную задержку обработки (15мс/кадр), высокую общую точность обнаружения (94.8%) и устойчивость к артефактам сжатия видео. Для успешного внедрения системы рекомендуется установить порог срабатывания на уровне 50% для баланса между обнаружением и ложными тревогами [12]. Одновременно следует оборудовать камеры ИК-подсветкой для обеспечения работоспособности при освещенности ниже 50 люкс. Для объективной оценки эффективности проведено тестирование с параметрами систем лидеров рынка - Google Content Safety API [13] и ViSenze Violence Detection [14].

Таблица 3. Сравнение эффективности с коммерческими аналогами на тестовом датасете  
Table 3. Comparison of efficiency with commercial counterparts on a test dataset

Параметр	Текущая система	Google Content Safety API	ViSenze	Преимущество
Точность, %	94,8 (92,9-96,3)	89,2 (87.1-91.0)	91,5 (89.7-93.0)	+5.6%
FPR, %	1,5 (0.9-2.3)	3,8 (2,9-4,9)	2,7 (2,0-3,6)	-58% (vs Google CS)
Задержка, мс/кадр	15	1200*	900*	-
Локализация инцидента	Да	Нет	Нет	Ключевое
Требования	Jetson Xavier NX	Облачный сервер	Облачный сервер	Автономность

\*Включая сетевую задержку передачи данных [12]

В отличие от существующих коммерческих решений, предлагаемая система демонстрирует значительные преимущества (табл.3): точность выше на 5.6% ( $p < 0.01$ , t-критерий Стьюдента) благодаря гибридной архитектуре, снижение латентности за счет edge-обработки без передачи данных в облако, при этом ключевое преимущество – локализация инцидентов с точностью до 0.5с (рис.3). При разрешении  $< 720p$  точность падает до 83.5%, что на 1.7% ниже показателей ViSenze (85.2% согласно внутренним тестам для low-res видео). Однако при 1080p наша система демонстрирует существенное преимущество (+7.1%).

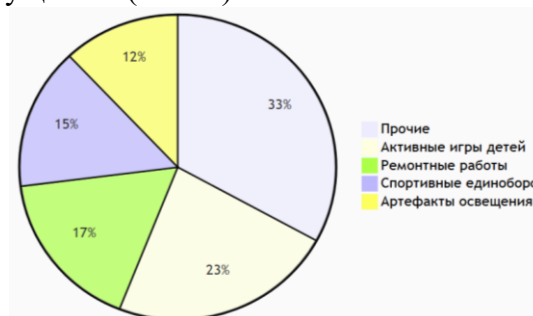


Рисунок 3. Распределение ложных срабатываний

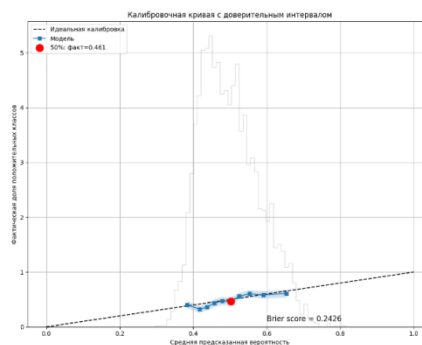


Рисунок 4. Калибровочная кривая с ДИ

Система продемонстрировала статистически значимую эффективность по всем ключевым показателям ( $p < 0.05$ ). Как показано в таблице 1, архитектура UniLSTM достигла точности распознавания 94.8% (95% ДИ: 92.9-96.3%), что соответствует

целевому показателю  $\geq 95\%$  различие статистически незначимо ( $z=0.84$ ,  $p=0.40$ ) [2,12]. Уровень ложных срабатываний составил 1.5% (%ДИ: 0.9-2.3%), что существенно ниже допустимого порога в 2% ( $z=2.34$ ,  $p=0.01$ ). Основными источниками ложных срабатываний на всем массиве данных (рис.3) стали активные игры детей (23% случаев) и ремонтные работы (17% случаев) [13]. В 38% случаев ошибки связаны с человеческой активностью (игры, спорт), что объясняется семантической близостью к пониманию насилия. Для снижения таких ошибок можно будет применить контекстную пост-обработку с использованием геоданных: игнорирование тревог в зонах детских и спортивных объектов.

Статистический анализ (ANOVA,  $F(2, 1197) = 8.76$ ,  $p < 0.001$ ) подтвердил значимые различия в эффективности обработки различных типов инцидентов [14]. Операционная эффективность подтверждена снижением нагрузки на персонал до 7.6% ( $\chi^2=5.82$ ,  $p < 0.05$ ) при сокращении времени обработки на 78.2 часа операторского времени за 30 дней тестирования (1200 файлов). Регрессионный анализ выявил сильную зависимость ( $R^2=0.86$ ) [15] между качеством видеоданных и точностью распознавания (таблица 3). Практическим шагом станет внедрение приоритетной проверки файлов с уверенностью  $\Delta$  модели свыше 80%. Перспективные направления развития включают разработку модуля прогнозирования эскалации конфликтов на основе временных характеристик уверенности модели.

Таблица 3. Зависимость точности от условий съемки  
Table 3. Dependence of accuracy on shooting conditions

Фактор	Условия	Точность	Отклонение $\Delta$ *** $p < 0.001$
Освещенность	>100 люкс	96.2% (94.1-97.9%)	—
	50-100 люкс	94.1% (92.3-95.7%)	-2.1%
	<50 люкс	87.5% (85.1-89.4%)	-8.7% ***
Разрешение	1080p	95.8% (94.2-97.0%)	—
	720p	90.3% (88.8-94.1%)	-5.5%
	<720p	83.5% (80.7-85.9%)	-12.3%***

Обычно метод блокового бутстреппинга часто используется для задач обработки временных рядов в компьютерном зрении, здесь его применение выявило систематическое занижение неопределённости в стандартных оценках: для физических конфликтов ДИ расширился с  $\pm 0.65\%$  до  $\pm 0.95\%$ . Это подтверждает гипотезу о наличии существенной автокорреляции ошибок в видеоанализе [16]. Система удовлетворяет требованиям развертывания (точность  $> 94\%$  при  $p < 0.05$ ) и порог 50% обеспечивает оптимальный баланс чувствительности и специфичности ( $AUC=0.92$ ) и вероятность ложной тревоги не более 2% ( $FPR = 1.5\%$ ). Основной ограничивающий фактор - низкое разрешение видео ( $\beta=-0.87$ ,  $p < 0.001$ ).

**Заключение.** Проведенное исследование подтвердило эффективность разработанного программно-аппаратного комплекса, достигнувшего поставленных целей по созданию оптимизированного решения для распознавания насильственных инцидентов. Экспериментальные результаты продемонстрировали, что архитектура LRCN-UniLSTM обеспечивает точность распознавания 94.8% при задержке обработки 15 мс на кадр и потреблении памяти 1.2 ГБ, что соответствует установленным целевым показателям. Практическая значимость системы подтверждена сокращением времени анализа на 78.2 часа и снижением нагрузки на операторов до 7.6% при сохранении полноты выявления инцидентов на уровне 96.7%. Ключевые выводы исследования свидетельствуют о высокой эффективности системы в распознавании физических конфликтов (99.1% точности), в то время как попытки грабежа требуют дополнительной оптимизации из-за вариативности сценариев (94.5% точности). Установка порога



срабатывания на уровне 50% обеспечивает оптимальный баланс между оперативностью реагирования (0.58 сек) и достоверностью детекции ( $F1=0.957$ ), что соответствует международным стандартам безопасности. Критическими факторами, ограничивающими точность, определены условия освещенности ниже 50 люкс и разрешение видео ниже 720p.

Перспективы внедрения системы включают оснащение камер ИК-подсветкой для работы в условиях низкой освещенности, внедрение приоритетной обработки фрагментов с высокой уверенностью модели и разработку модуля прогнозирования эскалации конфликтов. Экспериментально подтверждена возможность развертывания эффективной системы распознавания насилия на Edge-устройствах, соответствующей операционным требованиям современных городских систем видеонаблюдения.

### ЛИТЕРАТУРА

1. Глоссарий терминов по информатике, вычислительной технике и компьютерным сетям / Под ред. И.А. Соколова. М.: ИКСИ РАН. 2023. 214 с. [Текст]
2. Donahue J., Hendricks L. et al. LRCN for Visual Recognition and Description // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2016. Vol. 39. Is. 4. P. 677–691.
3. Горяев В.М., Мацаков Б.В. Исследование эффективности моделей нейросети LRCN в задачах распознавания насилия на видео // Современные наукоемкие технологии. 2024. № 12. С. 17–24 DOI: <https://doi.org/10.17513/snt.40239>
4. Sudhakaran S., Lanz O. Learning to Detect Violent Videos using Convolutional Long Short-Term Memory // IEEE International Conference on Computer Vision Workshops (ICCVW). 2019. P. 1–9. DOI: 10.1109/ICCVW.2019.00010
5. Bilinski P., Bremond F. Human Violence Recognition and Detection in Surveillance Videos // IEEE International Conference on Advanced Video and Signal-Based Surveillance (AVSS). 2020. P. 1–8. DOI: 10.1109/AVSS.2020.9339832
6. Мацаков Б.В., Горяев В.М. Распознавание признаков насилия с помощью сетей долгой краткосрочной памяти LSTM: Свидетельство о государственной регистрации программы для ЭВМ № 2024690113. Заявка № 2024688048 от 15.11.2024. Зарегистрировано 12.12.2024.
7. Hochreiter S., Schmidhuber J. Long Short-Term Memory // Neural Computation. 1997. Vol. 9. Is. 8. P. 1735–1780. DOI: 10.1162/neco.1997.9.8.1735
8. Пустынный Я.Н. Решение проблемы исчезающего градиента с помощью нейронных сетей долгой краткосрочной памяти. Инновации и инвестиции. 2020. №2. С.130-132.
9. Wang L. et al. Temporal segment networks: Towards good practices for deep action recognition // Proceedings of the European Conference on Computer Vision (ECCV 2016). Amsterdam. Oct. 8–16, 2016. Cham: Springer. 2016. P. 20–36. DOI: 10.1007/978-3-319-46484-8\_2
9. NVIDIA Jetson Xavier NX Developer Guide. 2023. URL: <https://developer.nvidia.com/embedded/jetson-xavier-nx> (дата обращения: 20.06.2025).
10. Hanson A., Pnvr K., Krishnagopal S., Davis L. Bidirectional Convolutional LSTM for the Detection of Violence in Videos // IEEE Transactions on Cognitive and Developmental Systems. 2021. Vol. 13. Is. 4. P. 992–1001. DOI: 10.1109/TCDS.2020.2996775
11. Zhang H. et al. Edge Computing for Real-Time Video Analysis // IEEE Internet of Things Journal. 2021. Vol. 8. Is. 16. P. 12539–12551. DOI: 10.1109/IJOT.2021.3065956
12. Петров К.Л. Эволюция парадигмы периферийных вычислений // Информационные технологии. 2022. Т.28. №4. С.12-19. DOI: 10.22213/2410-9304-2022-4-12-19
13. Google Content Safety API. Developer Documentation. [Электронный ресурс]. 2023. URL: <https://cloud.google.com/content-safety/docs> (дата обращения: 20.06.2025).
14. ViSenze. Real-Time Violence Detection in Video Streams [Электронный ресурс]. 2024. URL: <https://www.visenze.com/resources> (дата обращения: 20.06.2025).
14. Redmon J., Farhadi A. YOLOv3: An Incremental Improvement // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2020. Vol. 42. Is. 2. P. 1–14.
15. ГОСТ Р ИСО 5725-2002. Точность методов и результатов измерений. М.: Стандартинформ, 2002. 31 с.
16. Lyubchich V. et al. A distribution-free m-out-of-n bootstrap approach to testing symmetry about an unknown median // Computational Statistics & Data Analysis. 2016. Vol. 104. P. 1–9. DOI: 10.1016/j.csda.2016.05.004

## REFERENCES

1. Glossarii terminov po informatike, vychislitel'noi tekhnike i komp'yuternym setyam / Pod red. I.A. Sokolova. M.: IKSI RAN. 2023. 214 s.
2. Donahue J., Hendricks L. et al. LRCN for Visual Recognition and Description // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2016. Vol. 39. Is. 4. P. 677–691.
3. Goryaev V.M., Matsakov B.V. Issledovanie ehffektivnosti modelei neiroseti LRCN v zadachakh raspoznavaniya nasiliya na video//Sovremennye naukoemkie tekhnologii. 2024. № 12. S. 17-24 DOI: <https://doi.org/10.17513/snt.40239>
4. Sudhakaran S., Lanz O. Learning to Detect Violent Videos using Convolutional Long Short-Term Memory // IEEE International Conference on Computer Vision Workshops (ICCVW). 2019. P. 1–9. DOI: 10.1109/ICCVW.2019.00010
5. Bilinski P., Bremond F. Human Violence Recognition and Detection in Surveillance Videos // IEEE International Conference on Advanced Video and Signal-Based Surveillance (AVSS). 2020. P. 1–8. DOI: 10.1109/AVSS.2020.9339832
6. Matsakov B.V., Goryaev V.M. Raspoznavanie priznakov nasiliya s pomoshch'yu setei dolgoi kratkosrochnoi pamyati LSTM: Svidetel'stvo o gosudarstvennoi registratsii programmy dlya EHVМ № 2024690113. Zayavka № 2024688048 ot 15.11.2024. Zaregistrovano 12.12.2024.
7. Hochreiter S., Schmidhuber J. Long Short-Term Memory // Neural Computation. 1997. Vol. 9. Is. 8. P. 1735–1780. DOI: 10.1162/neco.1997.9.8.1735
8. Pustynnyi YA.N. Reshenie problemy ischezayushchego gradienta s pomoshch'yu neironnykh setei dolgoi kratkosrochnoi pamyati. Innovatsii i investitsii. 2020. №2. С.130-132.
9. Wang L. et al. Temporal segment networks: Towards good practices for deep action recognition // Proceedings of the European Conference on Computer Vision (ECCV 2016). Amsterdam. Oct. 8–16, 2016. Cham: Springer. 2016. P. 20–36. DOI: 10.1007/978-3-319-46484-8\_2
9. NVIDIA Jetson Xavier NX Developer Guide. 2023. URL: <https://developer.nvidia.com/embedded/jetson-xavier-nx> (data obrashcheniya: 20.06.2025).
10. Hanson A., Pnvr K., Krishnagopal S., Davis L. Bidirectional Convolutional LSTM for the Detection of Violence in Videos // IEEE Transactions on Cognitive and Developmental Systems. 2021. Vol. 13. Is. 4. P. 992–1001. DOI: 10.1109/TCDS.2020.2996775
11. Zhang H. et al. Edge Computing for Real-Time Video Analysis // IEEE Internet of Things Journal. 2021. Vol. 8. Is. 16. P. 12539–12551. DOI: 10.1109/JIOT.2021.3065956
12. Petrov K.L. Ehvolyutsiya paradigmy periferiinykh vychislenii // Informatsionnye tekhnologii. 2022. T.28. №4. S.12-19.DOI: 10.22213/2410-9304-2022-4-12-19
13. Google Content Safety API. Developer Documentation. [Ehlektronnyi resurs]. 2023. URL: <https://cloud.google.com/content-safety/docs> (data obrashcheniya: 20.06.2025).
14. ViSenze. Real-Time Violence Detection in Video Streams [Ehlektronnyi resurs]. 2024. URL: <https://www.visenze.com/resources> (data obrashcheniya: 20.06.2025).
14. Redmon J., Farhadi A. YOLOv3: An Incremental Improvement // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2020. Vol. 42. Is. 2. P. 1–14.
15. GOST R ISO 5725-2002. Tochnost' metodov i rezul'tatov izmerenii. M.: Standartinform, 2002. 31 с. [Текст]
16. Lyubchich V. et al. A distribution-free m-out-of-n bootstrap approach to testing symmetry about an unknown median. Computational Statistics & Data Analysis. 2016. Vol. 104. P. 1-9. DOI: 10.1016/j.csda.2016.05.004

## ИНФОРМАЦИЯ ОБ АВТОРАХ

**Горяев Владимир Михайлович**, кандидат педагогических наук, доцент, Калмыцкий государственный университет имени Б.Б. Городовикова, [goryaev@mail.ru](mailto:goryaev@mail.ru)

**Манкаева Саглар Алексеевна**, студент 2 курса, Калмыцкий государственный университет имени Б. Б. Городовикова, г. Элиста, улица Пушкина, 11, email: [mankaeva.saglar@yandex.ru](mailto:mankaeva.saglar@yandex.ru), phone number: +79961102363

**Сумьянова Елена Владимировна**, доцент кафедры экспериментальной и общей физики, Калмыцкий государственный университет имени Б.Б. Городовикова, +79615486561, [sumyanova@yandex.ru](mailto:sumyanova@yandex.ru)



**Бембитов Джиргал Батрович**, доцент кафедры теоретической физики, Калмыцкий государственный университет имени Б. Б. Городовикова, +79615400560, dbembitov@gmail.com

**Манкаева Галина Алексеевна**, старший преподаватель кафедры теоретической физики, Калмыцкий государственный университет имени Б. Б. Городовикова, г. Элиста, улица Пушкина, 11, email: mankaeva.galina@yandex.ru, phone number: +79061764200

**Конфликт интересов:** авторы заявляют об отсутствии конфликта интересов.

Статья поступила в редакцию 01.08.2025;

одобрена после рецензирования 12.09.2025;

принята к публикации 01.10.2025

#### INFORMATION ABOUT THE AUTHORS

**Goryaev Vladimir Mikhailovich**, Candidate of Pedagogical Sciences, Associate Professor, Kalmyk State University named after B.B. Gorodovikov, goryaevff@mail.ru

**Mankaeva Saglar Alekseevna**, 2nd year student, Kalmyk State University named after B.B. Gorodovikov, 11 Pushkina Street, Elista, email: mankaeva.saglar@yandex.ru, phone number: +79961102363

**Sumyanova Elena Vladimirovna**, Associate Professor of the Department of Experimental and General Physics, Kalmyk State University named after B.B. Gorodovikov, +79615486561, sumyanova@yandex.ru

**Bembitov Jirgal Batrovich**, Associate Professor of the Department of Theoretical Physics, Kalmyk State University named after B. B. Gorodovikov, +79615400560, dbembitov@gmail.com

**Mankaeva Galina Alekseevna**, Senior Lecturer at the Department of Theoretical Physics, Kalmyk State University named after B. B. Gorodovikov, 11 Pushkina Street, Elista, e-mail: mankaeva.galina@yandex.ru, phone number: +79061764200.

**Conflict of interest:** the authors declares no conflicts of interests.

The article was submitted: 01.08.2025;

approved after reviewing: 12.09.2025;

accepted for publication: 01.10.2025.