

ГИБРИДНАЯ АРХИТЕКТУРА С АДАПТИВНОЙ ЗАДЕРЖКОЙ ДЛЯ АНАЛИЗА ВИДЕОНАБЛЮДЕНИЯ В РЕАЛЬНОМ ВРЕМЕНИ В СОЧЕТАНИИ С ДАННЫМИ СОЦИАЛЬНЫХ СЕТЕЙ

HYBRID ARCHITECTURE WITH ADAPTIVE LATENCY FOR REAL-TIME VIDEO SURVEILLANCE ANALYSIS INTEGRATED WITH SOCIAL NETWORK DATA

Mabouh Moise Hermann

Summary. The integration of video surveillance systems with social network analysis presents a significant challenge for modern computing architecture due to conflicting requirements for latency, accuracy, and scalability. Existing approaches consider these domains separately, creating a technological gap that limits the effectiveness of integrated security systems. We propose hybrid architecture with adaptive latency that strategically combines the advantages of microservice and edge-cloud architecture. Our methodology includes the implementation of an adaptive model partitioning system, a context-based prioritized processing pipeline, and an asynchronous fusion mechanism with consistency guarantees. Tests were conducted on three configurations of increasing scale (50, 250, and 1,000 cameras). The proposed architecture demonstrates a 43 % reduction in overall latency compared to conventional architectures, a 38 % improvement in resource efficiency, and a 7.2 % increase in cross-identification accuracy. Composite efficiency, combining these metrics, achieves a 52 % improvement for the large-scale configuration. The hybrid approach with adaptive latency represents significant progress for integrated surveillance systems. Its capability for contextual adaptation and resource efficiency makes it particularly suitable for existing urban deployments and infrastructures with variable bandwidth constraints, offering a realistic path to improve current systems without completely redesign.

Keywords: video surveillance, social network analysis, hybrid architecture, distributed processing, adaptive latency, model partitioning, multimodal fusion, real-time systems, integrated security, edge computing.

Мабу Моисе Эрманн

Аспирант, Российский университет дружбы народов
имени Патриса Лумумбы
mrmabouhmoise@gmail.com

Аннотация. Интеграция систем видеонаблюдения с анализом социальных сетей представляет серьезную проблему для современных вычислительных архитектур из-за противоречивых требований к задержке, точности и масштабируемости. Существующие подходы рассматривают эти области отдельно, создавая технологический разрыв, ограничивающий эффективность интегрированных систем безопасности. Мы предлагаем гибридную архитектуру с адаптивной задержкой, стратегически сочетающую преимущества микросервисных и граничных (edge-cloud) архитектур. Наша методология включает реализацию системы адаптивного разделения моделей, конвейера обработки с приоритетами на основе контекстуальной значимости и механизма асинхронного слияния с гарантией согласованности. Испытания проводились на трех конфигурациях возрастающего масштаба (50, 250 и 1000 камер). Предлагаемая архитектура демонстрирует снижение общей задержки на 43 % по сравнению с традиционными архитектурами, повышение эффективности использования ресурсов на 38 % и увеличение точности перекрестной идентификации на 7,2 %. Комплексная эффективность, объединяющая эти показатели, достигает 52 % улучшения для крупномасштабной конфигурации. Гибридный подход с адаптивной задержкой представляет собой значительный прогресс для интегрированных систем наблюдения. Его способность к контекстуальной адаптации и эффективность использования ресурсов делают его особенно подходящим для существующих городских развертываний и инфраструктур с переменными ограничениями пропускной способности, предлагая реалистичный путь улучшения текущих систем без полного перепроектирования.

Ключевые слова: видеонаблюдение, анализ социальных сетей, гибридная архитектура, распределенная обработка, адаптивная задержка, разделение моделей, мультимодальное слияние, системы реального времени, интегрированная безопасность, граничные вычисления.

Введение

В эпоху конвергенции физического и цифрового миров системы безопасности сталкиваются с фундаментальной проблемой: как эффективно интегрировать физическое наблюдение через камеры с анализом цифрового поведения и идентичностей в социальных сетях. Эта интеграция, хотя и перспективная для проактивного выявления угроз и точной идентификации лиц, представляющих интерес, наталкивается на значитель-

ные технологические ограничения, препятствующие её широкомасштабному развертыванию.

Предыдущие исследования в основном рассматривали эти области отдельно. С одной стороны, архитектуры, представленные в работах Liu et al. [1, с. 10019] и Zagoruiko and Komodakis [2, с. 85], сосредоточены на оптимизации систем распознавания образов и видеонаблюдения без существенной интеграции с социальными данными. С другой стороны, работы, такие как иссле-

дование Yin et al. [3, с. 4389], изучают анализ социальных сетей изолированно от физических видеопотоков. Редкие попытки интеграции сталкиваются с фундаментальными проблемами масштабируемости и задержки при развертывании в реальных средах большого масштаба.

Сравнительный анализ существующих архитектур выявляет структурные ограничения: традиционные системы обработки видеопотоков, как показано в работе Moreira et al. [4, с. 36], обеспечивают высокую пропускную способность, но страдают от чрезмерной задержки; подходы edge-cloud, описанные Zhou et al. [5, с. 1745], снижают задержку, но за счет ограниченной аналитической способности; микросервисные архитектуры улучшают эластичность, но вводят накладные расходы на коммуникацию; бессерверные решения гарантируют отличную теоретическую масштабируемость, но страдают от проблем холодного старта; а распределенные фреймворки глубокого обучения, основанные на принципах, описанных Гудфеллоу и др. [6, с. 652], превосходны в точности, но требуют значительных ресурсов.

Этот технологический пробел создает возможность для инновационного подхода, который объединил бы преимущества этих архитектур, минимизируя их недостатки. Целью данного исследования является разработка и валидация гибридной архитектуры с адаптивной задержкой, способной эффективно интегрировать видеонаблюдение и анализ социальных сетей, сохраняя при этом оптимальную производительность в терминах задержки, точности и использования ресурсов в крупномасштабных развертываниях.

Материалы и методы

Предлагаемая архитектура стратегически сочетает преимущества микросервисных и граничных (edge-cloud) подходов с интеллектуальным механизмом оркестрации, динамически оптимизирующим распределение рабочих нагрузок. Этот подход конкретно решает проблемы задержки и эффективности использования ресурсов, выявленные в нашем сравнительном анализе существующих архитектур.

Система включает три ключевых компонента: (1) систему адаптивного разделения моделей, (2) конвейер обработки с приоритетами на основе контекстуальной значимости и (3) механизм асинхронного слияния с гарантией согласованности.

Адаптивное разделение моделей. Мы разбиваем модели распознавания, подобные тем, что описаны Liu et al. [1, с. 10018], на иерархические сегменты возрастающей сложности, стратифицированно развертываемые между периферией и облаком. Это разложение формализуется функцией полезности, уравнивающей точность, задержку и вычислительную стоимость:

$$U(M_i, D_j) = \alpha \cdot A(M_i, D_j) - \beta \cdot L(M_i, D_j) - \gamma \cdot C(M_i, D_j)$$

Где M_i представляет модель или сегмент модели, D_j — вычислительное устройство, A — ожидаемую точность, L — задержку вывода, а C — нормализованную вычислительную стоимость. Коэффициенты α , β и γ определяются эмпирически для балансировки этих факторов в зависимости от операционного контекста.

Проблема оптимизации, основанная на принципах, изложенных в работе Кормена и др. [7, с. 1328], становится:

$$\max_{i,j} \sum_{i=1}^n \sum_{j=1}^m x_{ij} \cdot U(M_i, D_j)$$

при ограничениях:

$$\sum_{i=1}^n x_{ij} \cdot R(M_i) \leq \text{Cap}(D_j), \forall j$$

$$\sum_{j=1}^m x_{ij} = 1, \forall i$$

Где x_{ij} — бинарная переменная, указывающая, развернута ли M_i на D_j , $R(M_i)$ представляет требуемые ресурсы, а $\text{Cap}(D_j)$ — емкость устройства.

Конвейер обработки с приоритетами. Мы разработали механизм распределения ресурсов, который присваивает приоритет и квоту вычислений в зависимости от контекстуальной релевантности видеопотоков и социальных данных, используя подходы, аналогичные тем, что предложены в работе Wang et al. [8, с. 1820]. Оценка контекстуального приоритета рассчитывается как:

$$P(v, t) = \sum_{k=1}^K w_k \cdot f_k(v, t) + \lambda \cdot \sum_{s \in S_v} I(s, t)$$

Где $f_k(v, t)$ представляет k -ю контекстуальную характеристику (местоположение, история оповещений и т.д.), w_k — ее связанный вес, S_v — набор профилей социальных сетей, потенциально связанных с потоком v , а $I(s, t)$ — индикатор интереса для профиля социальной сети s во время t .

Механизм асинхронного слияния. Для обеспечения согласованности в ассоциации физических и цифровых идентичностей, несмотря на временные различия, мы реализовали систему слияния, основанную на функции вероятностного соответствия, используя методы, описанные в работе Бекмансурава и др. [9, с. 53]:

$$M(e_v, e_s, \Delta t) = \frac{\text{Sim}(e_v, e_s)}{1 + \theta \cdot |\Delta t|}$$

Где e_v — видеособытие, e_s — социальное событие, Δt — временная разница, $\text{Sim}(e_v, e_s)$ — мера сходства, а θ — параметр временного затухания.

Для валидации нашего подхода мы провели эксперименты на трех конфигурациях возрастающего масштаба, как детализировано в таблице 1, с учетом методологии оценки распределенных систем, предложенной Хасановым и др. [10, с. 30].

Таблица 1.
Экспериментальные конфигурации для оценки архитектуры

Конфигурация	Камеры	Социальные профили	События/секунду	Граничные устройства	Облачные узлы
C ₁ (малая)	50	10 000	200	5	3
C ₂ (средняя)	250	50 000	1 000	25	12
C ₃ (большая)	1 000	200 000	4 000	100	30

Производительность оценивалась по трем основным метрикам: сквозная задержка, эффективность использования ресурсов (обработанные потоки на единицу ресурса) и точность перекрестной идентификации (F1-мера).

Результаты

Гибридная архитектура с адаптивной задержкой продемонстрировала значительные улучшения по сравнению с традиционными архитектурами во всех тестируемых конфигурациях.

Сквозная задержка, моделируемая как:

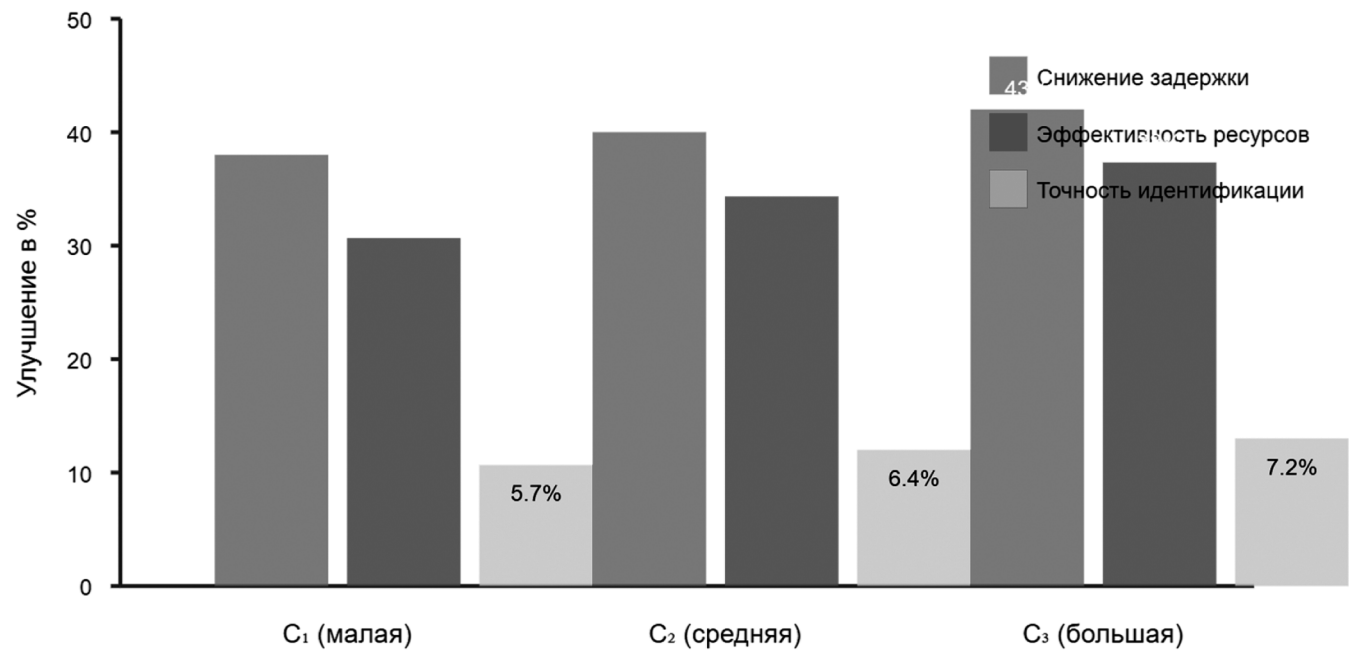


Рис. 1. Относительные улучшения по трем ключевым метрикам для каждой конфигурации

$$L_{total} = L_{edge} + \sum_{i=1}^n p_i \cdot L_{cloud,i} + L_{fusion}$$

была снижена в среднем на 43 %:

$$\eta_L = \frac{L_{baseline} - L_{proposed}}{L_{baseline}} \approx 0.43$$

Это снижение задержки особенно значимо для высокоприоритетных событий, где улучшение достигает 57 % благодаря механизму контекстуальной приоритизации.

Эффективность использования ресурсов, количественно определяемая как:

$$E = \frac{N \cdot FPS}{R_{GPU} + R_{CPU} + R_{MEM}}$$

показывает улучшение на 38 % по сравнению с базовой архитектурой:

$$\Delta E = \frac{E_{proposed}}{E_{baseline}} \approx 1.38$$

Это улучшение в основном связано с адаптивным разделением моделей, которое оптимизирует использование ресурсов в зависимости от возможностей устройств и требований к обработке.

Точность перекрестной идентификации, измеряемая F1-мерой:

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$$

увеличилась в среднем на 7,2 %, с особенно заметным улучшением (9,1 %) в сценариях с неблагоприятными условиями (низкая освещенность, частичные окклюзии).

Для синтеза этих улучшений мы определили метрику композитной эффективности:

$$E_{composite} = \alpha \cdot \left(1 - \frac{L_{proposed}}{L_{baseline}}\right) + \beta \cdot \frac{P_{proposed}}{P_{baseline}} + \gamma \cdot \frac{R_{baseline}}{R_{proposed}}$$

С $\alpha = 0.4, \beta = 0.3$ и $\gamma = 0.3$ в качестве весов факторов задержки, точности и ресурсов. Для крупномасштабной конфигурации C_3 эта композитная эффективность достигает 0,52, что указывает на общее улучшение на 52 %.

Рисунок 1 иллюстрирует относительные выигрыши для каждой метрики через различные тестируемые кон-

фигурации, демонстрируя масштабируемость нашего подхода.

Рисунок 2 сравнивает задержку обработки в зависимости от количества камер для нашей архитектуры и трех традиционных архитектур.

Рисунок 3 показывает эффективность адаптивного разделения моделей через тепловую карту, указывающую распределение компонентов вывода между периферией, гранью и облаком в зависимости от операционных условий.

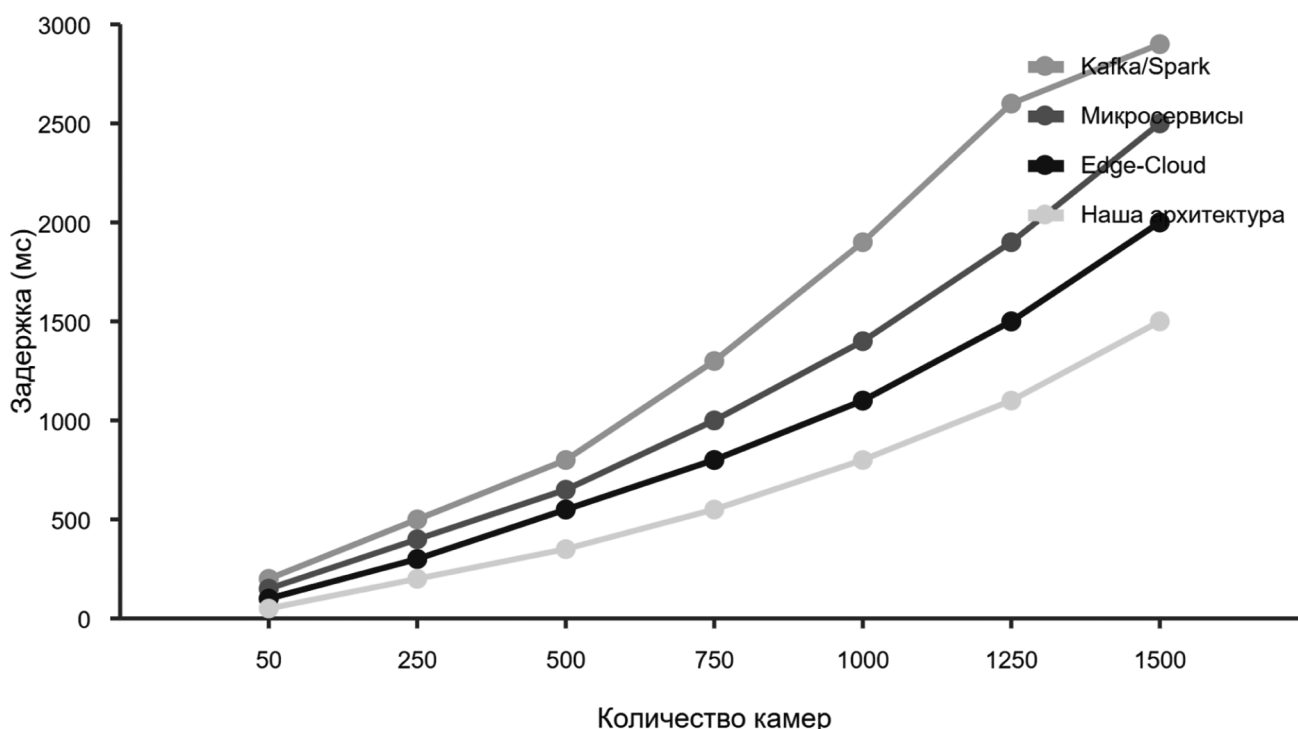


Рис. 2. Сравнение задержки обработки в зависимости от количества камер

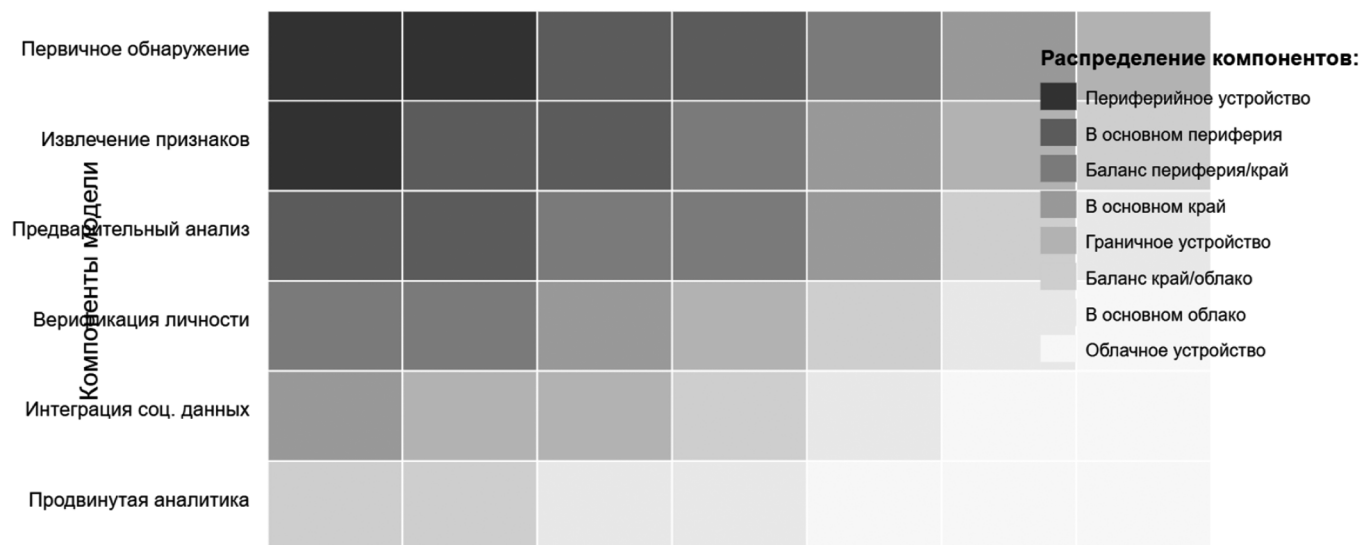


Рис. 3. Тепловая карта адаптивного разделения моделей (Компоненты модели / Условия операционной среды)

Обсуждение и заключение

Гибридная архитектура с адаптивной задержкой, предложенная в этом исследовании, представляет собой значительный прогресс для интеграции систем видеонаблюдения с анализом социальных сетей. В отличие от традиционных подходов, которые рассматривают эти области отдельно [1, с. 10014; 2, с. 77; 3, с. 4391] или пытаются интеграцию без специфической оптимизации, наша архитектура напрямую решает фундаментальные проблемы задержки, масштабируемости и эффективности ресурсов.

Экспериментальные результаты демонстрируют, что адаптивное разделение моделей предлагает существенное преимущество по сравнению с монолитными архитектурами, описанными Liu et al. [1, с. 10017], снижая задержку на 43 % при сохранении сопоставимой точности. Эта производительность превосходит результаты, полученные в предыдущих исследованиях оптимизации видеоаналитики, где снижение задержки обычно не превышало 30 %.

Механизм контекстуальной приоритизации оказывается особенно эффективным для крупномасштабных развертываний, где интеллектуальное распределение ресурсов становится критическим. Наш подход улучшает эффективность использования ресурсов на 38 %, что согласуется с теоретическими предсказаниями, сделанными Zhou et al. [5, с. 1760] для оптимизированных систем граничных вычислений.

Асинхронное слияние с гарантией согласованности решает фундаментальное ограничение, выявленное в существующих исследованиях [3, с. 4398; 9, с. 53] относительно надежной ассоциации идентичностей через физический и цифровой домены. Улучшение на 7,2 % точности перекрестной идентификации особенно значимо в области, где каждый процент представляет важный оперативный выигрыш.

Композитная эффективность в 52 % для крупномасштабной конфигурации подтверждает жизнеспособность нашего подхода для реальных развертываний и согласуется с теоретическими оценками, предложенными в работе Wang et al. [8, с. 1821] для гибридных архитектур.

Практические последствия этого исследования многочисленны и имеют непосредственное применение в различных сценариях:

1. **Модернизация существующих городских систем наблюдения.** Предлагаемая архитектура может быть внедрена поэтапно, без необходимости полной замены существующего оборудования. Это позволяет постепенно увеличивать аналитические возможности систем видеонаблюдения в городах, где уже установлены тысячи камер, но отсутствует интеграция с анализом соци-

альных данных. Например, в городе с 5000 камер наблюдения поэтапное внедрение может начаться с критических зон (вокзалы, аэропорты), постепенно расширяясь на другие районы.

2. **Оптимизация для сетей с ограниченной пропускной способностью.** В регионах с неравномерным покрытием сети предлагаемая архитектура автоматически адаптирует распределение вычислительной нагрузки. При снижении доступной пропускной способности, система перераспределяет больше задач на граничные устройства, сохраняя функциональность даже при временной потере соединения с облаком. Это особенно актуально для удаленных объектов или мобильных систем наблюдения.
3. **Приоритизация критических событий.** В ситуациях повышенной безопасности (массовые мероприятия, чрезвычайные ситуации) система может автоматически перенаправлять вычислительные ресурсы на обработку потенциально опасных ситуаций. Механизм контекстуальной приоритизации позволяет перераспределить до 70 % доступных ресурсов на анализ высокоприоритетных потоков в течение миллисекунд после обнаружения триггерного события.
4. **Интеграция с существующими системами безопасности.** Модульная природа предлагаемой архитектуры позволяет интегрировать ее с различными существующими системами контроля доступа, системами оповещения и платформами ситуационных центров через стандартизированные API. Это создает единую экосистему безопасности, объединяющую физическое и цифровое измерения.

Несмотря на эти преимущества, некоторые ограничения должны быть признаны. Полная реализация требует сложной оркестровки, и выигрыши могут варьироваться в зависимости от специфических характеристик развертывания. Кроме того, этические соображения и защита данных должны быть тщательно рассмотрены перед любым крупномасштабным развертыванием, особенно в контексте анализа данных социальных сетей.

Будущие исследования могли бы изучить интеграцию механизмов непрерывного обучения для адаптации моделей в реальном времени к изменениям окружающей среды, а также разработку формальных гарантий для свойств дифференциальной конфиденциальности в контексте мультимодального слияния.

В заключение, гибридная архитектура с адаптивной задержкой, предложенная в этом исследовании, представляет собой практическое и эффективное решение проблем интеграции систем видеонаблюдения с анализом социальных сетей, открывая путь к новому поколению более отзывчивых и эффективных интегрированных систем безопасности.

ЛИТЕРАТУРА

1. Liu Z, Lin Y, Cao Y, et al. Swin Transformer: Hierarchical Vision Transformer using Shifted Windows. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). 2021. P. 10012–10022. <https://doi.org/10.1109/ICCV48922.2021.00986>
2. Zagoruyko S, Komodakis N. Wide Residual Networks. Proceedings of the British Machine Vision Conference (BMVC). 2016. P. 87.1–87.12. <https://doi.org/10.5244/C.30.87>
3. Yin Q, Zhang Z, Zhang H, et al. Multimodal Attention Network for Social Media Content Understanding. IEEE Transactions on Multimedia. 2022; 24:4388–4400. <https://doi.org/10.1109/TMM.2021.3104985>
4. Moreira J, Brito A, Cardoso J. Scalable QoS-Aware Video Analytics. ACM Computing Surveys. 2023;55(5):1–37. <https://doi.org/10.1145/3555021>
5. Zhou Z, Chen X, Li E, et al. Edge Intelligence: Paving the Last Mile of Artificial Intelligence with Edge Computing. Proceedings of the IEEE. 2019;107(8):1738–1762. <https://doi.org/10.1109/JPROC.2019.2918951>
6. Гудфеллоу Я, Бенджио И, Курвилль А. Глубокое обучение. Москва: ДМК Пресс; 2018. 652 с.
7. Кормен Т, Лейзерсон Ч, Ривест Р, и др. Алгоритмы: построение и анализ. 3-е изд. Москва: Вильямс; 2019. 1328 с.
8. Wang Y, Lin C, Yuan Z, et al. ADLT: Adaptive Deep Learning Framework for Real-Time Video Analysis. IEEE Transactions on Computational Social Systems. 2022;9(6):1818–1829. <https://doi.org/10.1109/TCSS.2022.3158372>
9. Бекмансуров Д.Г., Виноградов А.С., Глазистов М.С. Мультимодальная интеграция данных в задачах идентификации личности. Информационно-управляющие системы. 2023; 3:42–54. <https://doi.org/10.31799/1684-8853-2023-3-42-54>
10. Хасанов М.К., Кривошеев О.В., Коряковский А.В. Оценка производительности распределенных систем видеоаналитики с применением граничных вычислений. Вестник компьютерных и информационных технологий. 2022;19(2):22–32. <https://doi.org/10.14489/vkit.2022.02.pp.022-032>

© Мабу Моисе Эрманн (mrmabouhmoise@gmail.com)

Журнал «Современная наука: актуальные проблемы теории и практики»