

Министерство образования Республики Беларусь
Учреждение образования
Белорусский государственный университет информатики и
радиоэлектроники

УДК 004.021:159.946.2

РЯЗАНЦЕВ
НИКИТА ДМИТРИЕВИЧ

АЛГОРИТМ РАСПОЗНАВАНИЯ ДВИЖЕНИЙ ЧЕЛОВЕКА

АВТОРЕФЕРАТ
на соискание степени магистра технических наук
по специальности 7-06-0612-03— Системы управления информацией

Научный руководитель
к.т.н., доц.
Жиляк Надежда
Александровна

Минск, 2025

ВВЕДЕНИЕ

Современные технологии все чаще требуют точного и эффективного распознавания движений человека. Эта задача является ключевой в таких областях, как виртуальная и дополненная реальность, робототехника, медицинская диагностика, спортивный анализ и системы взаимодействия человека с компьютером. Распознавание движений позволяет создавать интуитивно понятные интерфейсы, улучшать качество реабилитационных программ, анализировать технику выполнения спортивных упражнений и обеспечивать безопасность в системах видеонаблюдения. Однако, несмотря на значительные успехи в этой области, существующие алгоритмы часто сталкиваются с проблемами низкой точности, высокой вычислительной сложности и недостаточной адаптивности к различным условиям эксплуатации. Это делает актуальной разработку новых подходов, которые могли бы преодолеть указанные ограничения.

Актуальность темы обусловлена растущим спросом на технологии, способные точно и быстро анализировать движения человека в реальном времени. С развитием виртуальной реальности, робототехники и систем мониторинга здоровья потребность в эффективных алгоритмах распознавания движений только возрастает. Например, в медицинской сфере такие алгоритмы могут использоваться для анализа походки пациентов с нарушениями опорно-двигательного аппарата, а в спорте — для оценки техники выполнения упражнений. Кроме того, в игровой индустрии и системах виртуальной реальности точное распознавание движений необходимо для создания эффекта полного погружения, который бы максимально точно отражал действия пользователя. Таким образом, разработка алгоритмов, способных работать с высокой точностью и минимальными задержками, является важной научной и практической задачей.

Целью данной магистерской диссертационной работы является разработка и тестирование эффективного алгоритма распознавания движений человека, который бы обеспечивал высокую точность классификации движений при сокращении времени обработки данных.

Для достижения заданной цели поставлены и решены следующие задачи:

- проведен анализ существующих алгоритмов распознавания движений человека и оценить применимость различных подходов;
- исследованы методы машинного обучения, включая классические и современные (глубокое обучение, ансамбли, transfer learning);

- разработан комбинированный алгоритм, объединяющий CNN, RNN и механизм внимания, обеспечивающий как извлечение пространственных признаков, так и учет временных зависимостей;
- реализована архитектуру модели и проведено обучение на открытых видеодатасетах;
- осуществлено тестирование и сравнение производительности разработанного решения с существующими методами;
- продемонстрировано применимость алгоритма в реальных условиях, включая обработку видеопотока в реальном времени.

Достоверность научных результатов подтверждается соответствием теоретических выкладок и результатов разработки программного средства.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Связь работы с крупными научными программами и темами.

Диссертационное исследование выполнено в соответствии указа Президента Республики Беларусь от 7 апреля 2022 г. №136 «Об органе государственного регулирования в сфере цифрового развития и вопросах цифровизации».

Цель и задачи исследования.

Целью данной магистерской диссертационной работы является разработка и тестирование эффективного алгоритма распознавания движений человека, который бы обеспечивал высокую точность классификации движений при сокращении времени обработки данных

Для достижения заданной цели поставлены и решены следующие задачи:

- проведен анализ существующих алгоритмов распознавания движений человека и оценить применимость различных подходов;
 - исследованы методы машинного обучения, включая классические и современные (глубокое обучение, ансамбли, transfer learning);
 - разработан комбинированный алгоритм, объединяющий CNN, RNN и механизм внимания, обеспечивающий как извлечение пространственных признаков, так и учет временных зависимостей;
 - реализована архитектуру модели и проведено обучение на открытых видеодатасетах;
 - осуществлено тестирование и сравнение производительности разработанного решения с существующими методами;
- продемонстрировано применимость алгоритма в реальных условиях, включая обработку видеопотока в реальном времени В данной диссертационной работе объектом исследования являются движения человека, которые могут быть представлены в виде данных, полученных с различных сенсоров (например, видеокамер, акселерометров, гироскопов). Предметом исследования выступает алгоритм распознавания движений человека, который анализирует и классифицирует эти данные.

Апробация результатов диссертации.

Результаты, полученные в ходе выполнения диссертационной работы, а так же материалы, полученные в ходе исследования темы диссертационной работы докладывались и обсуждались на: международной научной конференции (Минск: БГУиР, 2023) [1–А], международной научно-технической конференции (Минск: БГТУ, 2023) [2–А], международной научной конференции (Минск: БГУиР, 2024) [3–А], 61-й студенческой научно-

технической конференции аспирантов, магистрантов и студентов БГУИР (Минск: БГУиР, 2025) [4–A], международной научно-технической конференции (Минск: БГТУ, 2025) [5–A], и опубликованы в виде тезисов в материалах к перечисленным выше конференциям. Имеется справка о внедрении НИР в учебный процесс, в частности результаты НИР внедрены в курс лабораторных работ по дисциплине «Игровые платформы» для специальности 1-40 05 01 12 «Информационные системы и технологии (в игровой индустрии)» по кафедре ВМиП.

Личный вклад соискателя.

Все этапы работы, включая анализ литературы, выбор архитектуры, реализацию и тестирование алгоритма, а также программную реализацию модели с применением библиотек PyTorch и OpenCV, выполнены автором лично. Автор самостоятельно разработал код обучения, предобработки данных, систему мониторинга метрик и систему распознавания движений в реальном времени. Научный руководитель принимал участие в постановке задачи, определений возможных путей из решений, в предварительном анализе, обсуждении результатов и практических исследований, проведенных автором лично.

Научная новизна.

Научная новизна исследования заключается в разработке комбинированного алгоритма, который интегрирует возможности сверточных нейросетей (CNN) для пространственного анализа, рекуррентных сетей (LSTM) для временного анализа, а также механизма внимания и ансамбля моделей для повышения точности. Предложенное решение дополнено рядом оптимизаций: transfer learning, label smoothing, аугментация данных и регуляризация, что обеспечивает устойчивую работу модели на разнородных датасетах и видеопотоках.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Диссертационная работа состоит из введения, общей характеристики, пяти глав основной части, заключения и списка использованных источников, оформленного в соответствии с установленными требованиями. Общий объём диссертации составляет 70 страницы, включая 2 страницы библиографического списка.

Во введении раскрыта актуальность темы, сформулированы цель и задачи исследования, обоснована научная новизна, определена практическая значимость, а также уточнены объект и предмет исследования. Также представлены общие проблемы, связанные с автоматическим распознаванием движений человека, и обозначены направления, в которых предполагается научный вклад.

Первая глава «Анализ существующих алгоритмов распознавания движений человека» содержит обзор современных методов и архитектур, используемых в задачах анализа движений по видеоданным. Рассматриваются подходы с использованием скелетных моделей, видеопотоков, акселерометров, а также методы на основе классического машинного обучения и нейронных сетей. Представлен сравнительный анализ по критериям точности, адаптивности и вычислительной эффективности.

Вторая глава «Разработка алгоритма распознавания движений на основе машинного обучения» посвящена построению архитектуры алгоритма. Подробно рассмотрены компоненты модели: свёрточные нейронные сети (CNN), рекуррентные слои (LSTM), механизм внимания и ансамбли моделей. Описан процесс подготовки данных, стратегия обучения с использованием *transfer learning*, методы регуляризации и настройка гиперпараметров.

Третья глава «Тестирование алгоритма» описывает организацию процедуры тестирования, включая структуру выборок, условия проведения, сценарии проверки и используемые метрики. Приведены результаты тестирования на независимых видеоданных и проанализирована устойчивость алгоритма к различным условиям съёмки, шумам и ошибкам классификации.

Четвёртая глава «Анализ результатов» включает всестороннюю оценку точности и эффективности разработанного алгоритма. Выполнено сравнение с существующими методами распознавания движений, представлены сильные и слабые стороны модели, даны рекомендации по её доработке. Также рассмотрены перспективные направления дальнейших исследований и возможные улучшения архитектуры.

Пятая глава «Применение в реальных условиях (Апробация результатов)» демонстрирует практическую реализацию алгоритма.

Представлены примеры использования системы в условиях, приближенных к реальным. Подтверждена работоспособность модели в режиме, близком к реальному времени, а также её применимость на устройствах с ограниченными вычислительными ресурсами.

В заключении подведены итоги выполненной работы, сформулированы основные результаты, подтверждена практическая ценность разработанного решения и обозначены направления для дальнейшего развития предложенного подхода.

ЗАКЛЮЧЕНИЕ

В рамках данной магистерской работы была решена комплексная научно-прикладная задача: разработка, обучение, тестирование и апробация алгоритма распознавания движений человека на основе современных методов машинного обучения. Целью исследования являлось создание точного, устойчивого и практически применимого алгоритма, способного эффективно распознавать широкий спектр действий по видеоданным.

В процессе работы были реализованы следующие этапы:

проанализированы существующие подходы к распознаванию движений и определены их ключевые ограничения;

разработана архитектура модели, объединяющая возможности свёрточных нейронных сетей для извлечения пространственных признаков, рекуррентных сетей для анализа временной динамики, а также механизма внимания и ансамблирования для повышения устойчивости;

проведена тщательная настройка и обучение модели с использованием открытых видеодатасетов, реализованы методы регуляризации, балансировки классов и оптимизации гиперпараметров;

выполнено тестирование алгоритма на независимых данных, подтверждающее его высокую точность, способность к обобщению и стабильность в нестандартных условиях;

реализована прикладная система, демонстрирующая работу алгоритма в реальном времени и подтверждающая его готовность к внедрению в практические решения.

Новизна работы заключается в интеграции нескольких подходов — transfer learning, bidirectional LSTM, attention-механизмов и ансамбли моделей — в единую архитектуру, обеспечивающую высокую точность распознавания при оптимальной вычислительной нагрузке. Предложенный алгоритм успешно сочетает научную обоснованность с прикладной ценностью, позволяя использовать его в широком спектре задач.

Практическая значимость подтверждена экспериментальной апробацией, в ходе которой алгоритм продемонстрировал способность к адаптации, масштабируемости и корректной работе на реальных видеопотоках. Это даёт основание рассматривать разработанную систему как технологическую базу для дальнейших прикладных проектов в области анализа движений.

Таким образом, поставленные цели и задачи работы были полностью достигнуты. Перспективы дальнейшего развития включают расширение архитектуры за счёт более современных моделей, работу с мультиканальными

данными и реализацию механизмов обучения без разметки, что позволит ещё более эффективно применять алгоритм в нестандартизированных средах.

Работа представляет собой завершённое научно-техническое исследование, обладающее теоретической новизной и высокой практической ценностью. Её результаты могут служить основой для дальнейших разработок в области распознавания движений человека.

СПИСОК ПУБЛИКАЦИЙ СОИСКАТЕЛЯ

1. Н. Д. Рязанцев, Н. А. Жиляк// Информационные технологии и системы 2023 (ИТС 2023): материалы Международной научной конференции, Минск, БГУИР, 2023 - С. 161–162.
2. Н. Д. Рязанцев, Д. Д. Рязанцев, Н. А. Жиляк // Минские научные чтения-2023. Технологическая независимость и конкурентоспособность Союзного государства, стран СНГ, ЕАЭС и ШОС : Сборник статей VI Международной научно-технической конференции, Минск, БГТУ, 2023 г. : в 3 т. Т. 2 – С. 370–375
3. Д. Д. Рязанцев, Н. Д. Рязанцев, Н. А. Жиляк // Минские научные чтения-2023. Технологическая независимость и конкурентоспособность Союзного государства, стран СНГ, ЕАЭС и ШОС : сборник статей VI Международной научно-технической конференции, Минск, БГТУ, 2023 г. : в 3 т. Т. 1 – С. 179–183.
4. Д. Д. Рязанцев, Н. Д. Рязанцев // Информационные технологии и системы 2024 (ИТС 2024): материалы международной научной конференции, Минск, БГУИР, 2024 г. – С. 115–118.
5. Д. Д. Рязанцев, Н. Д. Рязанцев, Н. А. Жиляк // 61-ая студенческая научно-техническая конференция аспирантов, магистрантов и студентов БГУИР, Минск, БГУИР, 2025 г. – в печати
6. Н. Д. Рязанцев, Д. Д. Рязанцев, Н. А. Жиляк // 61-ая студенческая научно-техническая конференция аспирантов, магистрантов и студентов БГУИР, Минск, БГУИР, 2025 г. – в печати
7. Д. Д. Рязанцев, Н. Д. Рязанцев, Н. А. Жиляк // Международной научно-технической конференции, Минск, БГТУ, 2025 г, Минск, БГУИР, 2025 г. – в печати
8. Н. Д. Рязанцев, Д. Д. Рязанцев, Н. А. Жиляк // Международной научно-технической конференции, Минск, БГТУ, 2025 г, Минск, БГУИР, 2025 г. – в печати