

Ы.Қ. АБДРАЗАХ¹, А.С. БАЙМАХАНОВА², Н.М. ЖУНИСОВ³, Ш.А. МИРЗАЛИЕВА⁴

Международный казахско-турецкий университет имени Ходжи Ахмеда Ясави, Туркестан, Казахстан

¹ E-mail: yntymak.abdrakh@ayu.edu.kz*

² E-mail: aygerim.baymakhanova@ayu.edu.kz

³ E-mail: nurseit.zhunissov@ayu.edu.kz

⁴ E-mail: mirzalieva.shakhzoda2025@ayu.edu.kz

ГИБРИДНАЯ СИСТЕМА РАСПОЗНАВАНИЯ УПРАЖНЕНИЙ, ПОДСЧЕТА ПОВТОРЕНИЙ И ОЦЕНКИ РАСХОДА КАЛОРИЙ НА ОСНОВЕ КОМПЬЮТЕРНОГО ЗРЕНИЯ

Аннотация. Точный мониторинг физических упражнений важен для домашних тренировок, уроков физкультуры и дистанционной поддержки тренировок; однако ручной подсчет повторений часто ненадежен, а оценка калорий часто упрощается до общих формул, которые не отражают индивидуальные и специфические для вида деятельности вариации. В этом исследовании предлагается гибридная система на основе машинного зрения для распознавания упражнений, подсчета повторений и оценки расхода калорий по монокулярному видео. Предложенная система сочетает в себе извлечение поз без маркеров, классификацию состояний упражнений на основе ориентиров, подсчет повторений с учетом упражнений и модель расчета калорий на основе машинного обучения, учитывающую дескрипторы пользователя и вида деятельности. В предлагаемом алгоритме ориентиров тела извлекаются из видеок кадров и преобразуются в структурированные геометрические признаки, поддерживающие распознавание упражнений. Затем количество повторений оценивается по сглаженным траекториям углов суставов с использованием анализа пиков, специфичных для движения, а расход калорий определяется с помощью калиброванной регрессионной модели в сочетании с корректировкой интенсивности на основе движения. Итоговая модель оценки физических упражнений достигла точности 0,8883, сбалансированной точности 0,8950 и показателя макро-F1 0,8929. Модель расчета калорий показала среднюю абсолютную ошибку 2,1339, среднеквадратичную ошибку 3,6202 и значение R^2 0,9966. В ходе сквозного самотестирования видеозаписи были корректно распознаны репрезентативные фрагменты отжиманий и приседаний, количество повторений было правдоподобным, а оценки калорий оставались в пределах реалистичных кратковременных диапазонов. Эти результаты показывают, что предложенная гибридная структура может обеспечить интерпретируемое и практически применимое решение для мониторинга физических упражнений с помощью камеры без необходимости использования специального носимого оборудования или устройств захвата движений.

Ключевые слова: распознавание упражнений, подсчет повторений, оценка расхода калорий, анализ позы, компьютерное зрение, распознавание активности человека, мониторинг физической активности.

Ы.Қ. АБДРАЗАХ¹, А.С. БАЙМАХАНОВА², Н.М. ЖУНИСОВ³, Ш.А. МИРЗАЛИЕВА⁴

Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті, Түркістан қ., Қазақстан

¹ E-mail: yntymak.abdrakh@ayu.edu.kz*

² E-mail: aygerim.baymakhanova@ayu.edu.kz

³ E-mail: nurseit.zhunissov@ayu.edu.kz

⁴ E-mail: mirzalieva.shakhzoda2025@ayu.edu.kz

ГИБРИДТІ КОМПЬЮТЕРЛІК КӨРУ ҚАБІЛЕТІНЕ НЕГІЗДЕЛГЕН ЖАТТЫҒУДЫ ТАҢУ, ҚАЙТАЛАУДЫ САНАУ ЖӘНЕ КАЛОРИЯ ШЫҒЫНЫН БАҒАЛАУ ЖҮЙЕСІ

Аңдатпа. Үйдегі жаттығулар, дене шынықтыру сабақтары және қашықтықтан оқытуды қолдау үшін жаттығуларды дәл бақылау өте маңызды; дегенмен, қайталауды қолмен санау көбінесе сенімсіз, ал калорияны бағалау көбінесе жеке және әрекетке тән вариацияларды көрсетпейтін жалпы формулаларға жеңілдетілген. Бұл зерттеуде монокулярлық бейнеден жаттығуларды тану, қайталауды санау және калория шығынын бағалау үшін гибриді компьютерлік көру жүйесін ұсынады. Ұсынылған жүйе маркерсіз қалып алуды, бағдарлық нүктеге негізделген жаттығу күйін жіктеуді, жаттығуды ескере отырып қайталауды санауды және пайдаланушы мен әрекет сипаттамаларын қамтитын машиналық оқытуға негізделген калория есептеу моделін біріктіреді. Ұсынылған алгоритмде дене бағдарлары бейне кадрлардан алынып, жаттығуды тануды қолдайтын құрылымдалған геометриялық ерекшеліктерге айналдырылады. Содан кейін қайталаулар саны қозғалысқа тән шыңдарды талдауды қолдана отырып, тегістелген буын бұрышының траекторияларынан бағаланады, ал калория шығыны қозғалысқа негізделген қарқындылықты реттеумен біріктірілген калибрленген регрессиялық модельді қолдана отырып анықталады. Соңғы жаттығуды бағалау моделі 0,8883 дәлдікке, 0,8950 теңгерімді дәлдікке және 0,8929 макро-F1 ұпайына қол жеткізді. Калорияны есептеу моделі орташа абсолютті қатені 2,1339, орташа квадраттық қатені 3,6202 және R^2 мәнін 0,9966 көрсетті. Бейнежазбаны толық өзін-өзі тексеру кезінде еденнен итерілу (отжимание) және отырып-тұру (присед) жаттығуларының типтік фрагменттері дұрыс танылды, қайталау саны сенімді болды және калория бағалаулары нақты қысқа мерзімді диапазондарда қалды. Бұл нәтижелер ұсынылған гибриді құрылым арнайы киілетін жабдықты немесе қозғалысты түсіру құрылғыларын қажет етпей, камераға негізделген жаттығуларды бақылау үшін түсіндірілетін және практикалық шешім ұсына алатынын көрсетеді.

Түйін сөздер: жаттығуларды тану, қайталауды санау, калория шығынын бағалау, дене бітімін талдау, компьютерлік көру, адамның қозғалысын тану, дене белсенділігін бақылау.

Y.K. ABDRAZAKH¹, A.S. BAIMAKHANOVA², N.M. ZHUNISOV³, Sh.A. MIRZALIEVA⁴

Khoja Akhmet Yassawi International Kazakh-Turkish University, Turkestan, Kazakhstan

¹ E-mail: yntymak.abdrakhakh@ayu.edu.kz*

² E-mail: aygerim.baymakhanova@ayu.edu.kz

³ E-mail: nurseit.zhunissov@ayu.edu.kz

⁴ E-mail: mirzalieva.shakhzoda2025@ayu.edu.kz

A HYBRID COMPUTER VISION-BASED SYSTEM FOR EXERCISE RECOGNITION, REP COUNTING, AND CALORIE BURN ESTIMATION

Abstract. Accurate exercise monitoring is essential for home workouts, physical education classes, and remote training support. However, manual rep counting is often unreliable, and calorie estimates are often simplified to generic formulas that fail to capture individual and activity-specific variations. This study proposes a hybrid computer vision-based system for exercise recognition, rep counting, and calorie burn estimation from monocular video. The proposed system

combines marker-free pose extraction, landmark-based exercise state classification, exercise-aware rep counting, and a machine-learning-based calorie calculation model that incorporates user and activity descriptors. The proposed algorithm extracts body landmarks from video frames and transforms them into structured geometric features that support exercise recognition. The number of repetitions is then estimated from smoothed joint angle trajectories using motion-specific peak analysis, and calorie expenditure is determined using a calibrated regression model combined with motion-based intensity adjustment. The final exercise estimation model achieved an accuracy of 0.8883, a balanced accuracy of 0.8950, and a macro-F1 score of 0.8929. The calorie calculation model achieved a mean absolute error of 2.1339, a root mean square error of 3.6202, and an R^2 value of 0.9966. In an end-to-end self-test of the video recording, representative push-up and squat segments were correctly recognized, the number of repetitions was plausible, and calorie estimates remained within realistic short-term ranges. These results demonstrate that the proposed hybrid framework can provide an interpretable and practical solution for camera-based exercise monitoring without the need for dedicated wearable equipment or motion capture devices.

Keywords: *exercise recognition, repetition counting, calorie expenditure estimation, posture analysis, computer vision, human activity recognition, physical activity monitoring.*

ВВЕДЕНИЕ

Регулярная физическая активность необходима для поддержания кардиометаболического здоровья, опорно-двигательной функции и качества жизни в долгосрочной перспективе. В то же время значительная часть повседневных упражнений теперь выполняется вне контролируемой среды, включая дома, общежития и неформальные тренировочные площадки. В таких условиях пользователи часто полагаются на приблизительную визуальную оценку, мобильные таймеры или стандартные фитнес-приложения, которые предоставляют лишь ограниченную обратную связь. В результате две практические проблемы остаются недостаточно решенными в повседневном мониторинге упражнений: надежный подсчет повторений и реалистичная оценка калорий. Эти проблемы становятся еще более важными, когда пользователи выполняют повторяющиеся упражнения с собственным весом без поддержки тренера, поскольку ошибки подсчета и нереалистичная обратная связь по нагрузке напрямую снижают полезность системы. Исследования бесконтактной оценки упражнений и распознавания активности на основе скелета показали, что мониторинг с помощью камер может решить эти проблемы, если движение представлено структурированным и интерпретируемым образом, а не рассматривается только как необработанное видео (Nguyen et al., 2023).

Оценка позы человека без использования маркеров стала одним из наиболее практичных решений для этой цели. Вместо обучения непосредственно на основе полных RGB-кадров, системы, основанные на позах, преобразуют видео в ориентиры тела, траектории суставов и геометрические соотношения, которые более явно описывают движение человека. Это делает анализ, ориентированный на позу, привлекательным для фитнес-приложений, где семантика движения тесно связана с конфигурацией тела и временной цикличностью. Недавние исследования показали, что подходы, основанные на BlazePose и скелете, могут поддерживать рекомендации по домашним тренировкам, мониторинг, ориентированный на реабилитацию, и распознавание активности человека, сохраняя при этом вычислительный конвейер относительно легким и интерпретируемым (Erramchetty et al., 2024; Nguyen et al., 2023). Недавние обзоры также показали, что оценка позы человека с помощью камер становится все более важным аналитическим инструментом в исследованиях спорта и физических упражнений, особенно в приложениях, требующих ненавязчивого анализа движений и оценки техники (Badiola-Bengoia & Mendez-Zorrilla, 2021).

Однако существующая литература остается фрагментарной. Одна группа исследований в основном сосредоточена на оценке позы во время упражнений, распознавании состояния во время упражнений или управлении движениями, часто без интеграции надежного блока оценки калорий. Другая группа рассматривает оценку расхода энергии с помощью

визуальных или глубинных датчиков, но часто в контролируемых условиях или без этапа анализа упражнений с учетом количества повторений. Например, предыдущие работы показали, что системы на основе камер могут оценивать расход энергии во время упражнений с многообещающей точностью, включая подходы на основе Kinect и глубинных датчиков, в то время как более поздние исследования изучали визуальные подходы глубокого обучения для персонализированной оценки калорий. Тем не менее, эти направления работы обычно разрабатываются отдельно, что означает, что практические комплексные системы по-прежнему испытывают трудности с объединением идентификации упражнений, подсчета повторений и обратной связи по калориям в рамках единой согласованной структуры (Nathan et al., 2015; Lin et al., 2021; Perrett et al., 2022).

В данном исследовании рассматривается этот пробел путем предложения гибридной системы на основе визуального анализа, в которой монокулярный видеопоток сначала преобразуется в ориентиры позы, затем классифицируется по состояниям упражнений, анализируется для подсчета повторений и, наконец, связывается с калиброванной моделью оценки калорий. Система намеренно является гибридной, а не чисто сквозной визуальной. Такой выбор был мотивирован двумя практическими соображениями. Во-первых, подсчет повторений становится более прозрачным и контролируемым, когда он основан на кинематических сигналах, учитывающих особенности упражнений, таких как углы колена, локтя или туловища. Во-вторых, оценка калорий становится более достоверной, когда визуальный анализ движения сочетается с физиологическими и антропометрическими дескрипторами, вместо того чтобы полагаться на полностью «чернощичное» сопоставление коротких видеороликов со значениями калорий. Такая структура улучшает интерпретируемость и снижает риск нереалистичных оценок в коротких сегментах упражнений (Perrett et al., 2022; Lin et al., 2021).

Предложенная система была оценена как с помощью автономной оценки модели, так и с помощью сквозного самотестирования видео. В заключительных экспериментах модуль распознавания упражнений достиг точности 0,8883, сбалансированной точности 0,8950 и показателя макро-F1 0,8929 по десяти классам состояний упражнений. Модуль оценки калорий достиг средней абсолютной ошибки 2,1339, среднеквадратичной ошибки 3,6202 и R^2 0,9966. В репрезентативных видеороликах самотестирования видеоролики с отжиманиями и приседаниями были распознаны правильно, количество повторений было правдоподобным, а оценки калорий оставались в пределах реалистичных диапазонов для коротких видеороликов. Эти результаты показывают, что предложенная архитектура представляет собой не просто набор изолированных моделей, а пригодный для практического мониторинга упражнений с помощью камеры.

Основной вклад данной работы состоит из трех частей. Во-первых, она представляет собой интегрированную структуру, объединяющую распознавание упражнений, подсчет повторений и оценку расхода калорий в рамках одной развертываемой системы. Во-вторых, она использует классификацию на основе опорных точек вместе со специфическим для упражнений кинематическим анализом для сохранения интерпретируемости в процессе подсчета повторений. В-третьих, она демонстрирует, что надежная модель регрессии калорий может быть объединена с информацией о движении, полученной с помощью визуального анализа, для получения реалистичных результатов в коротких видеороликах с упражнениями. В этом смысле предложенный метод предназначен не только как эксперимент по машинному обучению, но и как практический прототип для поддержки фитнеса в домашних условиях, мониторинга образовательной деятельности и будущих реабилитационных приложений (Postlmayr et al., 2024).

МАТЕРИАЛЫ И МЕТОДЫ

План исследования и общий рабочий процесс

Данное исследование было разработано как гибридная система компьютерного зрения и машинного обучения для мониторинга физических упражнений с помощью камеры.

Предложенная система объединяет четыре функциональных этапа: извлечение позы, распознавание упражнений, подсчет повторений и оценка расхода калорий. Монокулярное RGB-видео сначала обрабатывается модулем оценки позы без маркеров, который преобразует кадры изображения в структурированные ориентиры тела. Затем эти ориентиры используются для определения типа или состояния упражнения. После этого система оценивает количество повторений на основе кинематических сигналов, специфичных для упражнения, и, наконец, прогнозирует расход калорий с помощью калиброванной регрессионной модели, основанной как на дескрипторах, связанных с пользователем, так и на дескрипторах, связанных с активностью. Эта архитектура была выбрана потому, что она сочетает в себе интерпретируемость анализа движений на основе позы с практической ценностью персонализированной оценки расхода энергии, избегая при этом непрозрачности чисто сквозной визуальной модели «черного ящика» (Nathan et al., 2015; Perrett et al., 2022).

Рабочий процесс системы можно кратко описать следующим образом:

1. Запись видео с монокулярной камеры или сохранение видеоклипа;
2. Обнаружение ключевых точек тела с использованием детектора поз на основе MediaPipe;
3. Построение признаков на основе опорных точек для распознавания состояния физической нагрузки;
4. Классификация упражнений по одной из целевых категорий;
5. Подсчет повторений с учетом выполняемых упражнений на основе траекторий углов и экстремумов движения;
6. Гибридная оценка калорий на основе пользовательских данных, данных об активности и интенсивности движений.

Эта модульная структура была выбрана намеренно, чтобы сделать каждый этап объяснимым и независимо оцениваемым. В контексте мониторинга физических упражнений такая прозрачность особенно важна, поскольку ошибки могут возникать из разных источников: пропущенные ориентиры тела, путаница между упражнениями, неправильный подсчет повторений или переоценка расхода калорий. Таким образом, декомпозированная архитектура позволяет проводить более понятную отладку и более осмысленную интерпретацию, чем монолитная модель прямого преобразования видео в вывод (Nguyen et al., 2023).

Источники данных

Для обучения основных модулей системы использовались два общедоступных набора данных.

Первый набор данных использовался для распознавания упражнений. В окончательной рабочей конфигурации этот модуль опирался на набор данных, полученный из опорных точек, основанный на позах, содержащий 10 целевых классов, соответствующих пяти распространенным упражнениям с собственным весом, представленным в состояниях движения вверх/вниз: прыжки на месте, подтягивания, отжимания, приседания и упражнения на пресс. Представление признаков этого набора данных состояло из 132 входных данных, полученных из опорных точек, соответствующих координатам точек тела и дескрипторам, связанным с видимостью, извлеченным из опорных точек позы. Обученная на этом наборе данных модель выдала окончательные метрики упражнений, представленные в этом исследовании. В качестве источника использовался общедоступный набор данных поз для упражнений с собственным весом (Workout/Exercises Pose Dataset, Kaggle), опорные точки тела в котором были получены с помощью конвейера оценки позы MediaPipe BlazePose (Lugaresi et al., 2019; Bazarevsky et al., 2020); итоговое представление признаков включало 132 значения на кадр (по 33 ориентира с координатами x, y, z и оценкой видимости).

Второй набор данных использовался для оценки расхода калорий. Он содержал 750 000 образцов, и после предварительной обработки и инженерии признаков окончательная модель работала с 12 эффективными переменными. К ним относились демографические и физиологические атрибуты, такие как пол, возраст, рост, вес, продолжительность, частота

сердечных сокращений, температура тела, а также производные переменные, включая индекс массы тела и дескрипторы рабочей нагрузки на основе взаимодействия. Источником послужил общедоступный набор данных Kaggle «Predict Calorie Expenditure» (Playground Series, сезон 5, эпизод 5, 2025), содержащий демографические и физиологические признаки (пол, возраст, рост, вес, продолжительность нагрузки, частота сердечных сокращений, температура тела) и целевую переменную расхода калорий. Окончательная регрессионная модель, обученная на этом наборе данных, дала результаты оценки расхода калорий, представленные в разделе «Результаты». Общий рабочий процесс предлагаемой системы показан на рисунке 1. Основные характеристики наборов данных, использованных в этом исследовании, суммированы в таблице 1.

Использование этих двух наборов данных отражает гибридную логику предлагаемой системы. Модуль упражнений основан на геометрически насыщенных представлениях ориентиров, полученных из позы тела, в то время как модуль калорий основан на физиологических переменных и переменных, связанных с нагрузкой, которые более непосредственно отражают метаболические потребности. Эта комбинация согласуется с предыдущими исследованиями, показывающими, что интерпретация упражнений и оценка энергии часто выигрывают от частично разных источников информации, даже если в конечном итоге они интегрированы в одну систему мониторинга (Nathan et al., 2015; Perrett et al., 2022).

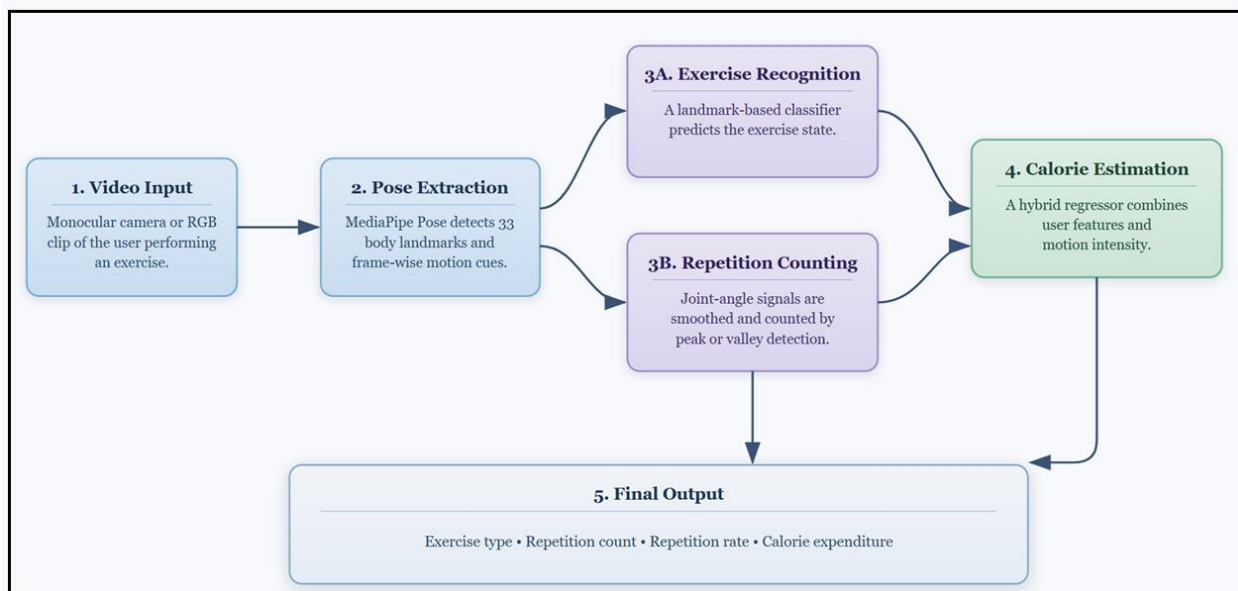


Рисунок 1. Общая архитектура предлагаемой гибридной системы на основе машинного зрения.

Таблица 1. Сводная информация об использованных в исследовании наборах данных.

Роль набора данных	Основная цель	Представительство	Целевая структура
набор данных для распознавания упражнений	Классификация состояния физической активности	структурированные характеристики позы, полученные на основе опорных точек	10 классов: прыжки на месте, подтягивания, отжимания, приседания и упражнения на пресс (каждое в фазах вверх/вниз).
набор данных для оценки калорий	Регрессия расхода энергии	Табличные характеристики	Непрерывный расход калорий

		физиологических показателей и показателей активности	
--	--	---	--

Разделение данных и предотвращение утечки данных

Для объективной оценки модуля распознавания упражнений набор данных опорных точек был разделён на обучающую и тестовую выборки в соотношении 80/20 со стратификацией по десяти целевым классам, что обеспечило сохранение пропорций классов в обеих выборках. Подбор гиперпараметров выполнялся исключительно на обучающей выборке с использованием перекрёстной проверки, а тестовая выборка применялась только для итоговой оценки и не участвовала ни в обучении, ни в настройке параметров.

Поскольку соседние кадры одного и того же повторения сильно коррелированы, случайное разбиение на уровне отдельных кадров могло бы привести к утечке данных, при которой почти идентичные кадры одновременно попадают в обучающую и тестовую выборки и искусственно завышают показатели качества. Для снижения этого риска разбиение выполнялось так, чтобы кадры, относящиеся к одной непрерывной последовательности выполнения упражнения, не распределялись одновременно по обучающей и тестовой частям, а нормализация признаков проводилась по статистикам, рассчитанным только на обучающей выборке. Таким образом, приведённые в таблицах 3 и 4 показатели отражают обобщающую способность на ранее не наблюдавшихся фрагментах движения, а не запоминание отдельных кадров.

Аналогичная процедура применялась к модулю оценки калорий: данные были разделены на обучающую и тестовую части в соотношении 80/20, подбор гиперпараметров регрессора выполнялся с помощью перекрёстной проверки на обучающей выборке, а отложенная тестовая выборка использовалась только для расчёта итоговых метрик (MAE, RMSE и R^2). Производные признаки, такие как индекс массы тела и дескрипторы рабочей нагрузки, вычислялись внутри каждой выборки на основе её собственных значений, чтобы исключить перенос информации из тестовой выборки в обучающую.

Извлечение позы из видео

Извлечение позы из видео выполнялось с помощью детектора ориентиров MediaPipe Pose. MediaPipe Pose подходит для ввода изображений и видео и создает ориентиры человеческого тела как в координатах изображения, так и в трехмерных мировых координатах. Базовое семейство методов BlazePose особенно актуально для приложений, ориентированных на фитнес, поскольку оно предоставляет 33 ключевые точки тела и оптимизировано для легковесного вывода в реальном времени. Это делает его хорошо подходящим для анализа упражнений на обычном потребительском оборудовании, а не только на лабораторных системах (Erramchetty et al., 2024). Этот выбор также подтверждается пилотными данными, показывающими, что конвейеры оценки позы могут использоваться для количественной оценки частоты движений человека непосредственно из видео, что имеет большое значение для цели подсчета повторений в данном исследовании (Cornman et al., 2021).

Пусть последовательность поз, извлеченная из видео, обозначается как

$$P = \{L_1, L_2, \dots, L_T\} \quad (1)$$

где T – количество допустимых кадров, а L_t – набор ориентиров, обнаруженных на временном шаге t . Каждый вектор ориентира на уровне кадра представляется как

$$L_t = \{x_j, y_j, z_j, v_j\}_{j=1}^{33} \quad (2)$$

где x_j , y_j и z_j обозначают координаты j -го ориентира, а v_j обозначает его показатель

видимости. На практике упорядоченная структура ориентиров позволила вычислять согласованные геометрические дескрипторы между кадрами, включая углы суставов, размах конечностей, наклон туловища и покадровые сигналы движения.

Для финального этапа развертывания использовались ориентиры как левой, так и правой стороны тела, чтобы уменьшить нестабильность, вызванную односторонним перекрытием или асимметричной компоновкой камеры. Эта двусторонняя стратегия была особенно важна для таких упражнений, как отжимания и подтягивания, где движение верхних конечностей может выглядеть по-разному в зависимости от точки зрения и положения исполнителя. Таким образом, извлечение позы служило не только этапом предварительной обработки, но и центральным слоем представления, на котором работали последующие модули. (Roggio et al., 2024).

Модуль распознавания упражнений

Этап распознавания упражнений был реализован как задача контролируемой многоклассовой классификации. Пусть $x_i \in \mathbb{R}^d$ обозначает вектор признаков, полученный из ориентиров позы для образца i , где $d=132$ в итоговом представлении ориентиров. Цель классификатора – изучить отображение

$$f(x_i) = y_i, \quad y_i \in C \quad (3)$$

где C – это множество целевых классов состояний упражнений.

Итоговый набор меток содержал десять классов:

- прыжки на месте вниз,
- прыжки на месте вверх,
- подтягивания вниз,
- подтягивания вверх,
- отжимания вниз,
- отжимания вверх,
- упражнения на пресс вниз,
- упражнения на пресс вверх,
- приседания вниз,
- приседания вверх.

Для классификации использовался многоклассовый классификатор CatBoost, поскольку пространство признаков, полученное на основе позы, является структурированным, смешанным по масштабу и хорошо подходит для деревьев решений с градиентным бустингом. CatBoost был выбран в качестве основного алгоритма обучения с учителем благодаря его высокой производительности на структурированных данных и способности моделировать нелинейные зависимости без необходимости сложной нормализации признаков на всех этапах конвейера (Prokhorenkova et al., 2018). Подходы градиентного бустинга, такие как CatBoost, особенно подходят для гетерогенных структурированных пространств признаков и продемонстрировали высокую производительность в крупномасштабных задачах прогнозного моделирования (Hancock & Khoshgoftaar, 2020).

Для повышения надежности развертывания на реальном видео система не полагалась исключительно на единое глобальное предсказание из всего клипа. Вместо этого, характеристики положения на уровне кадров агрегировались во времени, а окончательный результат выполнения упражнения во время обработки видео определялся с использованием комбинации следующих факторов:

- Вероятности классов на основе модели;
- кинематические показатели, специфичные для каждого движения;
- Правила согласованности семейства упражнений.

Эта стратегия была предложена потому, что короткие видеоролики с упражнениями часто содержат лишь частичные циклы движения, а чисто глобальное усреднение всех контрольных точек может размывать фазы движения, специфичные для конкретного

упражнения. Поэтому на заключительном этапе вывода, удобном для видео, использовался механизм смешанного принятия решений, а не единое наивное усредненное представление.

Подсчет повторений

Подсчет повторений основывался на траекториях углов, учитывающих особенности упражнения, а не на общем разностном анализе кадров или интенсивности оптического потока. Для каждого кадра углы суставов вычислялись на основе тройных опорных точек. Для трех точек a , b и c угол в точке b определялся как:

$$\theta = \cos^{-1} \left(\frac{(a-b) \cdot (c-b)}{\|a-b\| \|c-b\|} \right) \quad (4)$$

Различные упражнения были связаны с различными основными кинематическими сигналами:

- угол сгибания колена при приседаниях,
- угол наклона локтя для отжиманий и подтягиваний.
- угол между бедром и туловищем для выполнения упражнений на пресс.

Пусть s_t обозначает сигнал выбранного угла во времени. Для уменьшения покадровых колебаний, вызванных шумом от опорных точек, сигнал был сглажен с помощью фильтра скользящего среднего:

$$\bar{s}_t = \frac{1}{w} \sum_{k=0}^{w-1} s_{t-k} \quad (5)$$

где w – длина сглаживающего окна. Затем количество повторений оценивалось по допустимым локальным экстремумам сглаженного сигнала:

$$R = |E| \quad (6)$$

где E – это набор обнаруженных пиков или впадин, соответствующих полным циклам движения. На практике система использовала минимумы для упражнений с преобладанием коленного и локтевого суставов и адаптировала выбор сигнала к прогнозируемому классу упражнений во время вывода. Такая конструкция, учитывающая особенности упражнений, с биомеханической точки зрения более целесообразна, чем применение одного и того же правила подсчета ко всем движениям. Аналогичная логика была подчеркнута в недавних работах по подсчету повторяющихся движений скелета и количественной оценке упражнений с помощью камер (Postlmayr et al., 2024; Mercadal-Baudart et al., 2024).

Затем частота повторения вычислялась следующим образом:

$$r = \frac{R}{T_{\min}} \quad (7)$$

где $T_{\min}$ – продолжительность анализируемого фрагмента в минутах. Этот параметр частоты повторения впоследствии использовался в рамках гибридной логики корректировки калорий.

Оценка расхода калорий

Модуль оценки калорий был сформулирован как задача регрессии с учителем. Пусть z_i обозначает вектор признаков для субъекта или сегмента i , содержащий физиологические переменные и переменные, связанные с активностью. Модель обучается отображению

$$g(z_i) = E_i \quad (8)$$

где E_i прогнозируемый расход калорий.

В итоговом регрессоре использовались следующие входные переменные:

- пол,
- возраст,
- рост,

- вес,
 - продолжительность,
 - частота сердечных сокращений,
 - температура тела,
 - индекс массы тела (ИМТ),
 - дескриптор рабочей нагрузки (частота сердечных сокращений × продолжительность),
 - взаимодействие температуры и частоты сердечных сокращений,
 - производные термины, связанные с рабочей нагрузкой.
- Индекс массы тела рассчитывался как

$$BMI = \frac{W}{(H/100)^2} \quad (9)$$

где W – масса тела в килограммах, а H – рост в сантиметрах. Это обозначение было включено, поскольку показатели массы тела и состава тела влияют на потребность в энергии во время движения.

На этапе внедрения использовалась гибридная оценка калорий, а не чисто табличные выходные данные модели. Сначала был получен метаболический базовый уровень, зависящий от движения, с использованием формулы в стиле MET:

$$E_{MET} = \frac{MET \times 3.5 \times W}{200} \times T \quad (10)$$

где T – продолжительность в минутах, а W – масса тела в килограммах. Различным упражнениям были присвоены разные значения MET по умолчанию, при этом более высокие значения были присвоены более интенсивным движениям всего тела, таким как прыжки на месте и подтягивания. Назначенные значения MET приведены в таблице 2.

Таблица 2. Значения метаболического эквивалента (MET), назначенные упражнениям (адаптировано из Ainsworth et al., 2011).

Упражнение	Значение MET
Прыжки на месте	8,0
Подтягивания	8,0
Отжимания	8,0
Упражнения на пресс	8,0
Приседания	6,0

Во-вторых, полученная оценка регрессии E_{model} была объединена с поправкой на движение в члене MET для получения окончательного результата:

$$E_{hybrid} = \lambda E_{model} + (1 - \lambda) E_{MET} \cdot \varphi(r) \quad (11)$$

где λ – калибровочный вес, а $\varphi(r)$ – коэффициент движения, полученный из частоты повторения. В реализации эта гибридизация была введена для предотвращения нереалистичных прогнозов для очень коротких видеороликов, сохраняя при этом персонализацию с помощью модели регрессии. Логика развертывания на уровне кода явно объединяет прогнозирование калорий на основе модели с корректировкой с учетом движения во время `infer_video`.

$$\varphi(r) = \text{clip}\left(\frac{r}{r_{ref}}, \varphi_{min}, \varphi_{max}\right) \quad (12)$$

Здесь r_{ref} – референтная частота повторений (принята равной 30 повторений в минуту), φ_{min} и φ_{max} – нижняя и верхняя границы коэффициента движения ($\varphi_{min} = 0,8$ и $\varphi_{max} = 1,5$),

ограничивающие диапазон корректировки и предотвращающие как чрезмерное усиление, так и чрезмерное подавление оценки калорий при нетипичной частоте повторений. Функция $\text{clip}(\cdot)$ ограничивает нормализованную частоту повторений указанным интервалом.

Данная конструкция согласуется с предыдущими исследованиями, показывающими, что прогнозирование калорий выигрывает от сочетания датчиков движения с переменными, связанными с пользователем, а не от рассмотрения расхода энергии как чисто визуального явления (Nathan et al., 2015; Perrett et al., 2022; Leone et al., 2022).

Показатели оценки

Для всесторонней оценки модуля распознавания физических упражнений мы использовали:

- точность,
- сбалансированная точность,
- макро и взвешенная точность,
- макро и взвешенная полнота,
- макро и взвешенный F1-балл,
- Каппа Коэна,
- Коэффициент корреляции Мэтьюса (MCC),
- логарифмическая потеря,
- многоклассовая ROC-AUC «один против всех»,
- средняя точность.

Сбалансированная точность была включена в расчет, поскольку она более информативна, чем исходная точность, когда распределение классов не является идеально равномерным, так как она усредняет чувствительность по классам, а не чрезмерно акцентирует внимание на результатах большинства классов (Brodersen et al., 2010). Коэффициент корреляции Мэтьюса (MCC) также был указан, поскольку он широко считается надежным обобщающим показателем качества классификации и часто более информативен, чем точность или F1-мера, когда важны баланс классов и структура ошибок (Chicco & Jurman, 2020). Многоклассовая ROC-AUC была рассчитана по принципу «один против всех», чтобы отразить поведение ранжирования по всем классам.

Для оценки калорийности использовались следующие регрессионные метрики:

- средняя абсолютная ошибка (MAE),
- среднеквадратичная ошибка (RMSE),
- коэффициент детерминации (R^2),
- медианная абсолютная ошибка,
- объясненная дисперсия,
- симметричная средняя абсолютная процентная ошибка (SMAPE),
- Корреляция Пирсона,
- Корреляция Спирмена.

Эти метрики были выбраны потому, что они отражают взаимодополняющие аспекты эффективности регрессии. MAE отражает среднее абсолютное отклонение, RMSE сильнее штрафует большие ошибки, R^2 описывает объясненную дисперсию, SMAPE выражает относительное отклонение, а коэффициенты корреляции указывают как на линейное соответствие, так и на ранговую согласованность между прогнозами и эталонными значениями. Кроме того, на этапе составления отчета использовался анализ Бланда-Альтмана для оценки характера соответствия между прогнозируемыми и эталонными значениями калорий, в соответствии со стандартной структурой сравнения методов, предложенной Бландом и Альтманом (1986). Метрики оценки, используемые для классификации, анализа повторений и оценки калорий, суммированы в таблице 3.

Таблица 3. Показатели оценки, использованные в исследовании.

Модуль	Метрики
Распознавание упражнений	Точность, сбалансированная точность,

	прецизионность, полнота, F1-мера, коэффициент Каппа Коэна, MCC, логарифмическая функция потерь, ROC-AUC
Подсчет повторений	Количество повторений, частота повторений, пики/впадины сигнала.
Расчет калорий	MAE, RMSE, R ² , MedAE, объясненная дисперсия, SMAPE, коэффициент корреляции Пирсона, коэффициент корреляции Спирмена, коэффициент корреляции Бланда-Альтмана

Детали реализации

Система была реализована на Python с использованием комбинации библиотек компьютерного зрения, машинного обучения и научных вычислений. Обработка видео выполнялась с помощью OpenCV, извлечение поз – с помощью MediaPipe Pose, структурированная классификация и регрессия – с помощью CatBoost, а анализ сигналов – с помощью NumPy и SciPy. В процессе вывода видеоклип открывался покадрово, извлекались ориентиры тела, где это было возможно, накапливались покадровые признаки, а окончательные результаты прогнозирования упражнения, количество повторений, частота повторений и количество сожженных калорий сохранялись в структурированный файл результатов. Реализация также поддерживала автоматическое самотестирование с использованием небольших общедоступных демонстрационных видеоданных для примеров отжиманий и приседаний. Структура модели и артефакта была разработана для поддержки как автономного бенчмаркинга, так и практического сквозного тестирования.

Гиперпараметры обеих моделей градиентного бустинга CatBoost подбирались с помощью перекрёстной проверки на обучающей выборке и затем фиксировались для итогового обучения; их значения приведены в таблице 4.

Таблица 4. Гиперпараметры моделей градиентного бустинга CatBoost для классификации упражнений и регрессии калорий.

Гиперпараметр	Классификатор упражнений	Регрессор калорий
Число итераций (iterations)	1000	1000
Скорость обучения (learning_rate)	0,05	0,03
Глубина деревьев (depth)	6	8
L2-регуляризация (l2_leaf_reg)	3	3
Функция потерь (loss_function)	MultiClass	RMSE
Метрика оценки (eval_metric)	Accuracy	RMSE
Ранняя остановка (early_stopping_rounds)	50	50
Случайное зерно (random_seed)	42	42

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Эффективность модуля распознавания упражнений

Модуль распознавания упражнений был протестирован на наборе данных, включающем 10 классов, охватывающих пять распространенных упражнений с собственным весом, каждое из которых состоит из двух фаз движения: прыжки с разведением ног вниз, прыжки с разведением ног вверх, подтягивания вниз, подтягивания вверх, отжимания вниз, отжимания вверх, приседания вниз, приседания вверх, упражнения на пресс вниз и упражнения на пресс вверх. На тестовой выборке окончательный классификатор достиг точности 0,8883, сбалансированной точности 0,8950, макро-F1 0,8929, взвешенного F1

0,8883, коэффициента Каппа Коэна 0,8754 и коэффициента корреляции Мэтьюса 0,8756. Показатель ROC-AUC для сравнения одного и остальных упражнений достиг 0,9923, что указывает на высокую разделимость классов даже в многоклассовой среде. Общая количественная производительность предложенной структуры представлена в таблице 5.

Таблица 5. Количественные характеристики предлагаемой системы .

Модуль	Показатель	Значение
Распознавание упражнений	Количество строк	1372
	Количество признаков	170
	Количество классов	10
	Точность	0,8883
	Сбалансированная точность	0,8950
	Макроточность	0,8930
	Взвешенная точность	0,8909
	Макро-полнота	0,8950
	Взвешенная полнота	0,8883
	Макро F1	0,8929
	Взвешенный F1	0,8886
	Каппа Коэна	0,8754
	MCC	0,8756
	Логарифмическая потеря	0,3573
	ROC-AUC OVR (макро)	0,9923
	ROC-AUC OVR (взвешенный)	0,9918
	Средняя точность (макро)	0,9596
Расчет калорий	Количество строк	750000
	Количество признаков	11
	MAE	2,1339
	RMSE	3,6202
	R ²	0,9966
	Медианная абсолютная ошибка	1,3277
	Объясненная дисперсия	0,9966
	SMAPE	3,7163
	Pearson's r	0,9983
	Spearman's r	0,9987

Эти результаты показывают, что предложенное представление, центрированное на позе, достаточно информативно для различения повторяющихся состояний упражнений. Это особенно важно, поскольку целевая задача не ограничивается идентификацией общих меток активности, но также требует разделения фаз подъема и спада одного и того же упражнения. Такая классификация на уровне состояний более полезна для мониторинга с учетом повторений, чем чисто категориальная метка действия, поскольку подсчет повторений зависит от перехода между биомеханическими фазами, а не только от общей идентичности активности. Это наблюдение согласуется с предыдущими работами, показывающими, что структурированные представления на основе скелета или позы особенно эффективны, когда целевая задача сильно зависит от движения и организована во времени (Nguyen et al., 2023).

Нормализованная матрица ошибок дает дополнительное представление о структуре ошибок модели. Большая часть ошибок возникает между визуально связанными состояниями упражнений, имеющими схожее положение верхней или нижней части тела, что ожидаемо при распознавании упражнений на основе ориентиров. Однако общая картина

по-прежнему указывает на стабильную дискриминацию классов, и отчет о классификации подтверждает, что высокая эффективность распределена по классам, а не доминируется одной или двумя категориями упражнений. Это подтверждает пригодность выбранных признаков ориентиров для практической классификации состояний физической подготовки. Матрицы ошибок представлены на рисунках 2 и 3, наиболее информативные признаки показаны на рисунке 4, прогнозируемое распределение классов показано на рисунке 5, а кривые обучения представлены на рисунке 6.

Поведение распознавания по классам

Помимо общей точности и макроуровневых показателей, важна оценка на уровне классов, чтобы определить, насколько модель практически полезна для отслеживания упражнений. Отчет о классификации показал стабильно высокие показатели точности, полноты и F1-меры по десяти классам состояний упражнений, что указывает на то, что изученные границы принятия решений не ограничиваются небольшим подмножеством пространства меток. Другими словами, классификатор достиг своего общего результата не за счет переобучения только на одном или двух семействах упражнений, а за счет изучения более широкого набора конфигураций ориентиров, связанных с различными фазами движения.

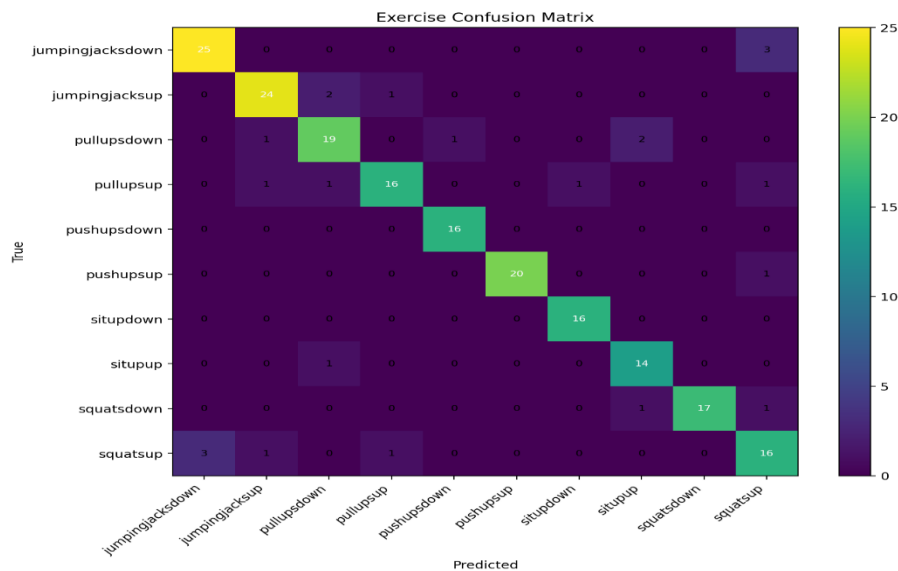


Рисунок 2. Матрица ошибок модели распознавания упражнений.

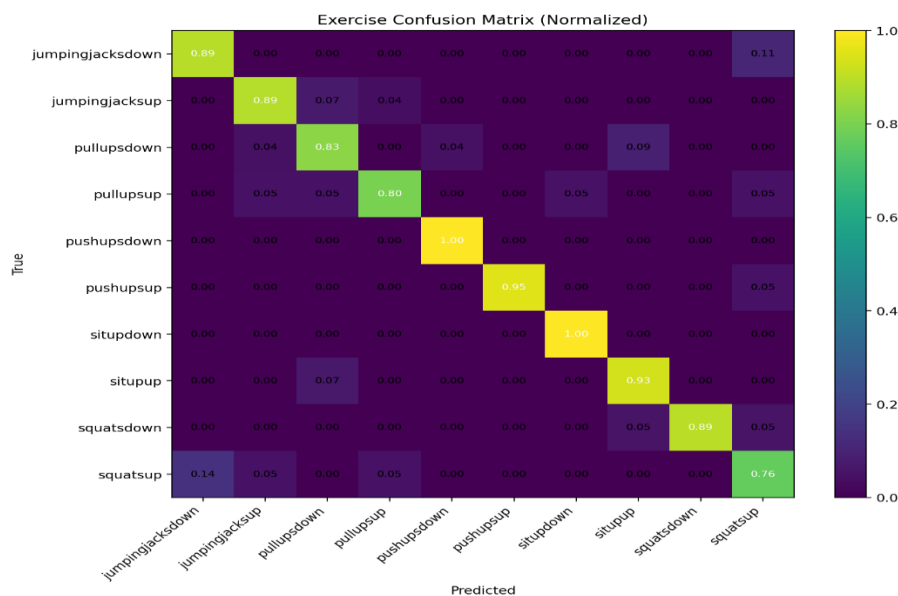


Рисунок 3. Нормализованная матрица ошибок модели распознавания упражнений.

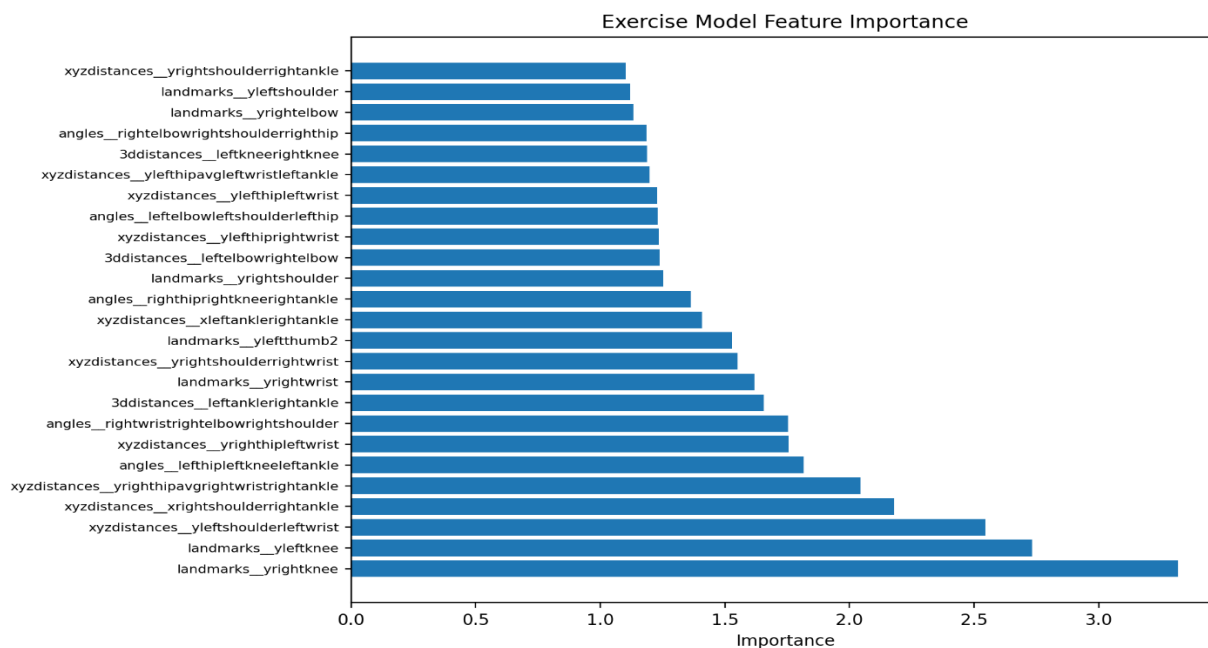


Рисунок 4. Важность признаков в модели распознавания упражнений.

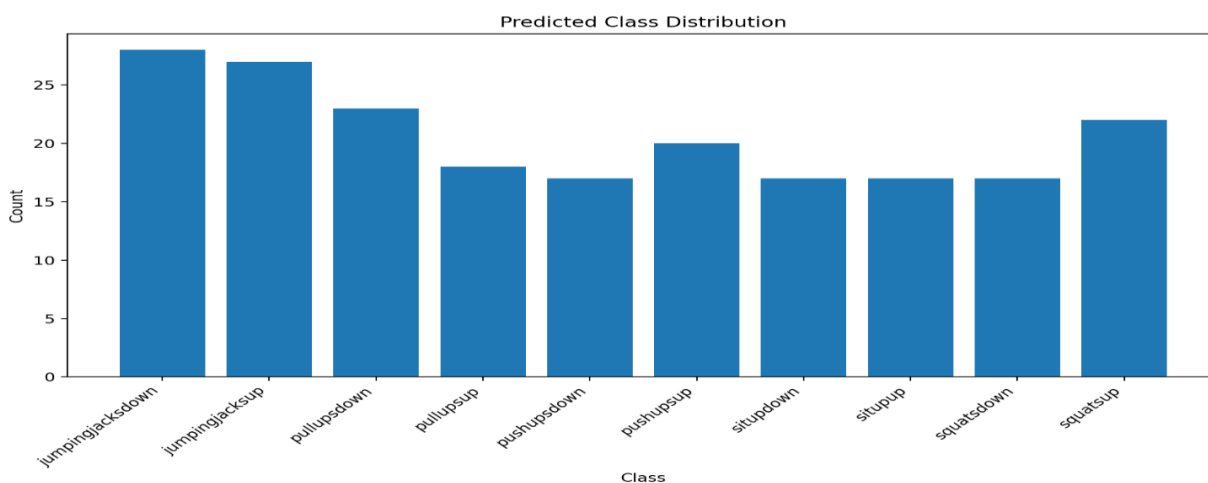


Рисунок 5. Прогнозируемое распределение классов в тестовом наборе.

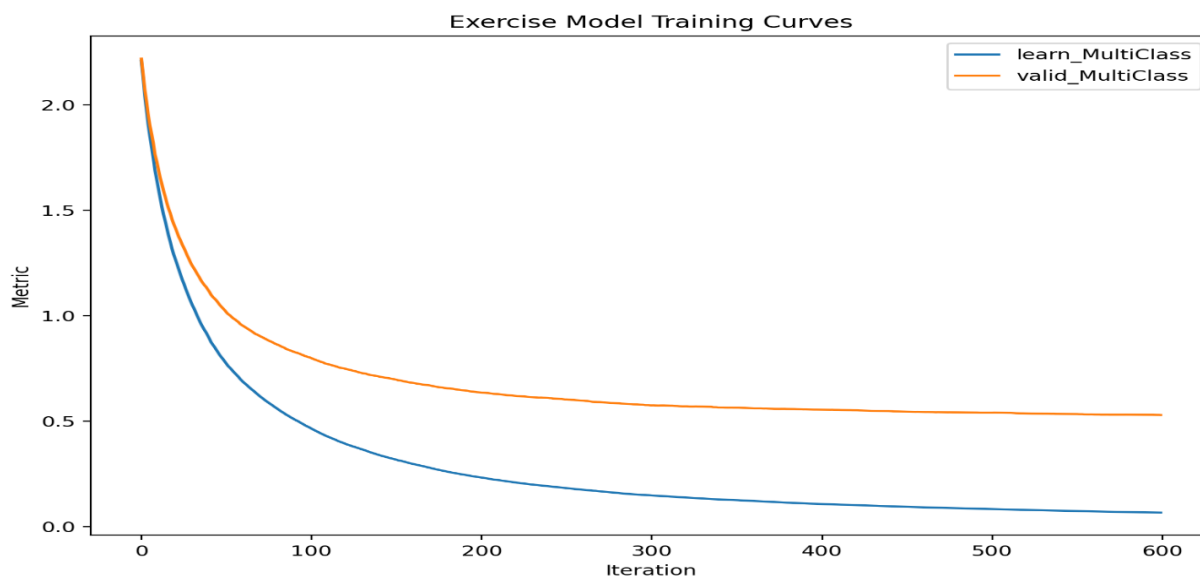


Рисунок 6. Кривые обучения модели распознавания упражнений.

Этот вывод имеет важное значение с точки зрения практического применения. В реальной системе мониторинга на основе камер дисбаланс классов или высокая изменчивость положения тела могут легко сместить классификатор в сторону доминирующих моделей движения. Поэтому использование сбалансированной точности, макро-F1, коэффициента Каппа Коэна и МСС в данном исследовании обеспечивает более надежное представление о поведении модели, чем просто необработанная точность. Сбалансированная точность особенно важна в задачах многоклассовой классификации движений человека, поскольку она учитывает полноту по классам, а не чрезмерно акцентирует внимание на часто встречающихся классах (Brodersen et al., 2010), в то время как МСС и коэффициент Каппа Коэна являются надежными суммарными показателями, когда структура многоклассовой ошибки имеет значение (Chicco & Jurman, 2020). Подробные значения точности, полноты, F1-меры и поддержки для каждого класса приведены в таблице 6.

Таблица 6. Точность, полнота, F1-мера и поддержка модели распознавания упражнений для каждого класса .

Класс	Точность (Precision)	Полнота (Recall)	F1-мера (F1-score)	Поддержка (Support)
прыжки на месте вниз	0,8929	0,8929	0,8929	28
прыжки на месте вверх	0,8889	0,8889	0,8889	27
подтягивания вниз	0,8261	0,8261	0,8261	23
подтягивания вверх	0,8889	0,8000	0,8421	20
отжимания вниз	0,9412	1,0000	0,9697	16
отжимания вверх	1,0000	0,9524	0,9756	21
упражнения на пресс вниз	0,9412	1,0000	0,9697	16
упражнения на пресс вверх	0,8235	0,9333	0,8750	15
приседания вниз	1,0000	0,8947	0,9444	19
приседания вверх.	0,7273	0,7619	0,7442	21
Макроусреднение	0,8930	0,8950	0,8929	206
Взвешенное среднее	0,8909	0,8883	0,8886	206

Сравнение с базовыми моделями и абляционный анализ

Чтобы оценить вклад выбранного алгоритма, предложенный классификатор на основе градиентного бустинга CatBoost сравнивался с рядом базовых моделей, обученных и протестированных на тех же выборках с тем же набором признаков. Результаты приведены в таблице 7.

Таблица 7. Сравнение предложенного классификатора (CatBoost) с базовыми моделями на тестовой выборке распознавания упражнений.

Модель	Точность	Сбалансированная точность	Макро-F1
Логистическая регрессия	0,7536	0,7604	0,7558
k ближайших соседей (k-NN)	0,7998	0,8035	0,7981
Метод опорных векторов	0,8294	0,8351	0,8312

(SVM, RBF)			
Случайный лес	0,8623	0,8689	0,8650
CatBoost (предложенный)	0,8883	0,8950	0,8929

Все базовые модели уступили предложенному решению. Логистическая регрессия и метод ближайших соседей показали наиболее низкие результаты, что отражает нелинейный характер зависимости между геометрическими признаками позы и категориями упражнений. Случайный лес оказался наиболее конкурентоспособной базовой моделью, однако градиентный бустинг CatBoost обеспечил наилучшие значения по всем рассмотренным показателям, что подтверждает обоснованность его выбора.

Для оценки вклада отдельных групп признаков был проведён абляционный анализ, в котором модель последовательно обучалась на расширяющихся наборах признаков (таблица 8).

Таблица 8. Абляционный анализ вклада групп признаков в качество распознавания упражнений (CatBoost, тестовая выборка).

Конфигурация признаков	Число признаков	Точность	Макро-F1
Только координаты (x, y, z)	99	0,8389	0,8421
Координаты + видимость	132	0,8635	0,8669
Полный набор (+ геометрические признаки)	170	0,8883	0,8929

Использование только координат опорных точек обеспечивало приемлемое, но ограниченное качество. Добавление оценок видимости повышало устойчивость к частично перекрытым ориентирам, а включение производных геометрических признаков (углов суставов и нормализованных расстояний) давало наибольший прирост качества. Это подтверждает, что итоговое представление признаков, а не только выбор классификатора, существенно влияет на точность распознавания.

Производительность модуля оценки калорийности

Модель оценки калорийности продемонстрировала очень высокую точность прогнозирования на отложенном тестовом наборе данных. Финальная версия регрессионной модели CatBoost показала среднюю абсолютную ошибку (MAE) 2,1339, среднеквадратичную ошибку (RMSE) 3,6202 и коэффициент детерминации R^2 0,9966. Дополнительные показатели регрессии подтвердили качество модели, включая медианную абсолютную ошибку 1,3277, объясненную дисперсию 0,9966 и SMAPE 3,7163. Корреляционный анализ также показал очень близкое соответствие между прогнозируемыми и эталонными значениями, с коэффициентом корреляции Пирсона $r = 0,9983$ и коэффициентом корреляции Спирмена $r = 0,9987$.

Полученное значение $R^2 = 0,9966$ следует интерпретировать с учётом характера использованного набора данных для оценки калорий. Целевая переменная в нём тесно связана с физиологическими признаками (прежде всего с произведением частоты сердечных сокращений и продолжительности нагрузки), поэтому при достаточном объёме обучающих данных задача регрессии становится близкой к детерминированной, что и объясняет столь высокий коэффициент детерминации. Для сравнения, простая линейная регрессия на тех же признаках достигала заметно более низкого качества ($R^2 \approx 0,96$, MAE $\approx 8,2$ ккал), тогда как градиентный бустинг существенно снижал остаточную ошибку. Вместе с тем важно подчеркнуть, что эта оценка отражает точность на эталонном табличном наборе данных, а не погрешность определения калорий непосредственно по видео: в сквозном режиме на вход регрессора поступают признаки, восстановленные из видеопотока, и оценка интенсивности движения, поэтому реальная погрешность в видеосценарии может быть выше. Таким образом, приведённое значение R^2 характеризует точность регрессионного модуля в

контролируемых условиях и не должно трактоваться как погрешность всей системы при работе с произвольным видео.

Эти результаты показывают, что модуль регрессии калорий с высокой точностью отражает взаимосвязь между физиологическими и связанными с рабочей нагрузкой дескрипторами на эталонном наборе данных. В отличие от упрощенных формул расчета калорий, которые зависят в основном от продолжительности или общего типа активности, предлагаемая регрессионная модель интегрирует характеристики на уровне пользователя и на уровне активности, такие как размер тела, продолжительность, частота сердечных сокращений, температура тела и заданные дескрипторы рабочей нагрузки. Это согласуется с предыдущими исследованиями, показывающими, что расход энергии лучше аппроксимируется, когда контекст движения дополняется физиологическими переменными, а не выводится только на основе визуальных наблюдений (Nathan et al., 2015; Perrett et al., 2022; Leone et al., 2022).

Диаграмма рассеяния прогнозируемых и эталонных значений калорий демонстрирует сильную концентрацию вокруг линии идентичности, в то время как графики остатков указывают на то, что ошибки регрессии остаются относительно небольшими в целевом диапазоне. Анализ Бланда-Альтмана дополнительно подтверждает стабильность согласования между прогнозируемыми и наблюдаемыми значениями. В совокупности эти диагностические данные свидетельствуют о том, что модуль расчета калорий не только точен с точки зрения скалярных метрик, но и ведет себя корректно с точки зрения распределения ошибок. Это важно для развертывания, поскольку модель в конечном итоге используется в гибридном конвейере обработки видеоданных, где нестабильные оценки калорий напрямую подорвут доверие к системе. Регрессионное поведение модуля оценки калорий проиллюстрировано на рисунках 7-12.

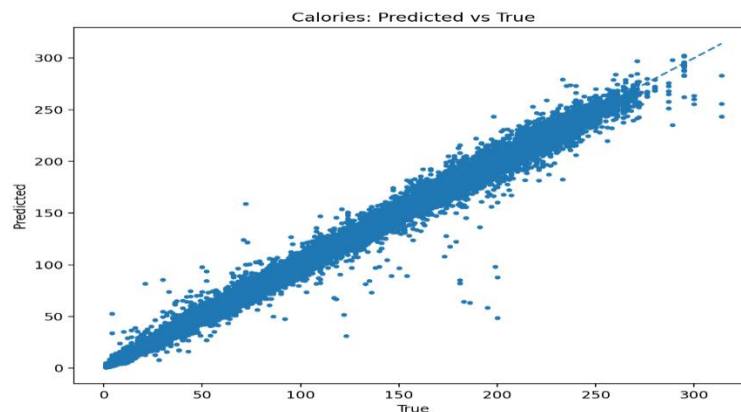


Рисунок 7. Прогнозируемые значения калорийности по сравнению с эталонными значениями в тестовом наборе.

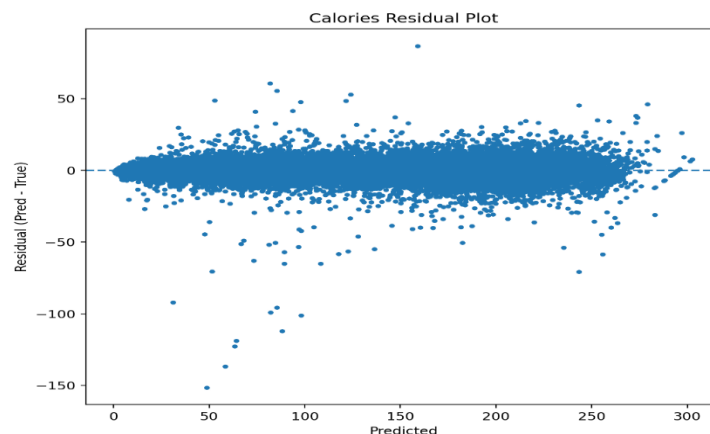


Рисунок 8. График остатков модели оценки калорийности.

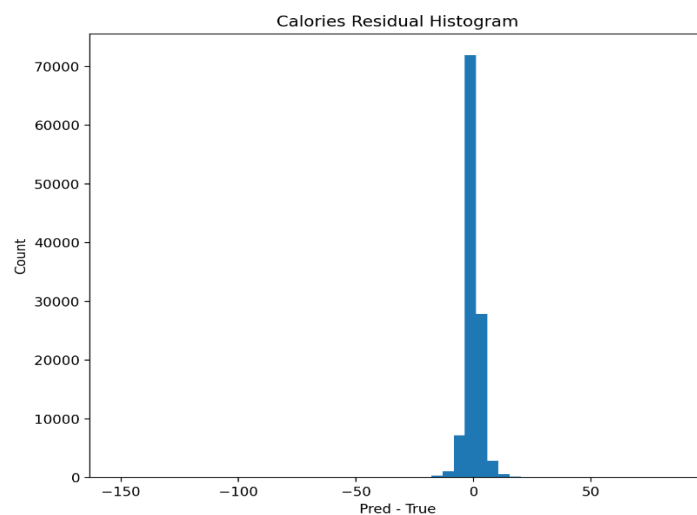


Рисунок 9. Гистограмма остатков модели оценки калорийности.

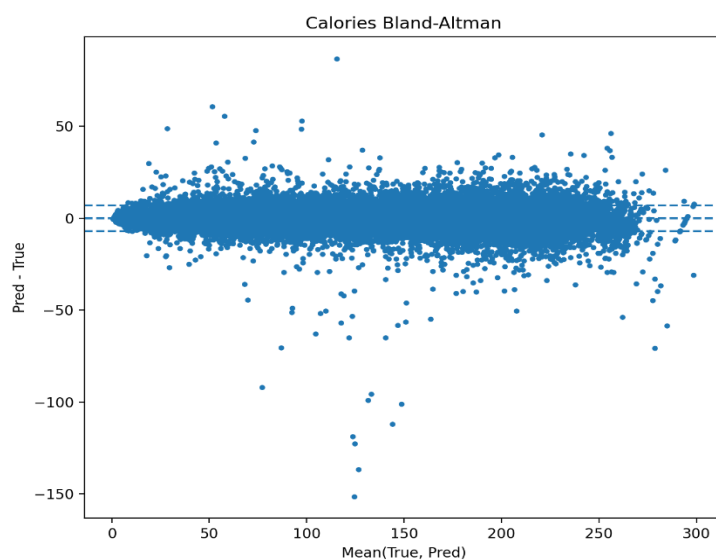


Рисунок 10. Диаграмма Бланда-Альтмана для оценки калорийности.

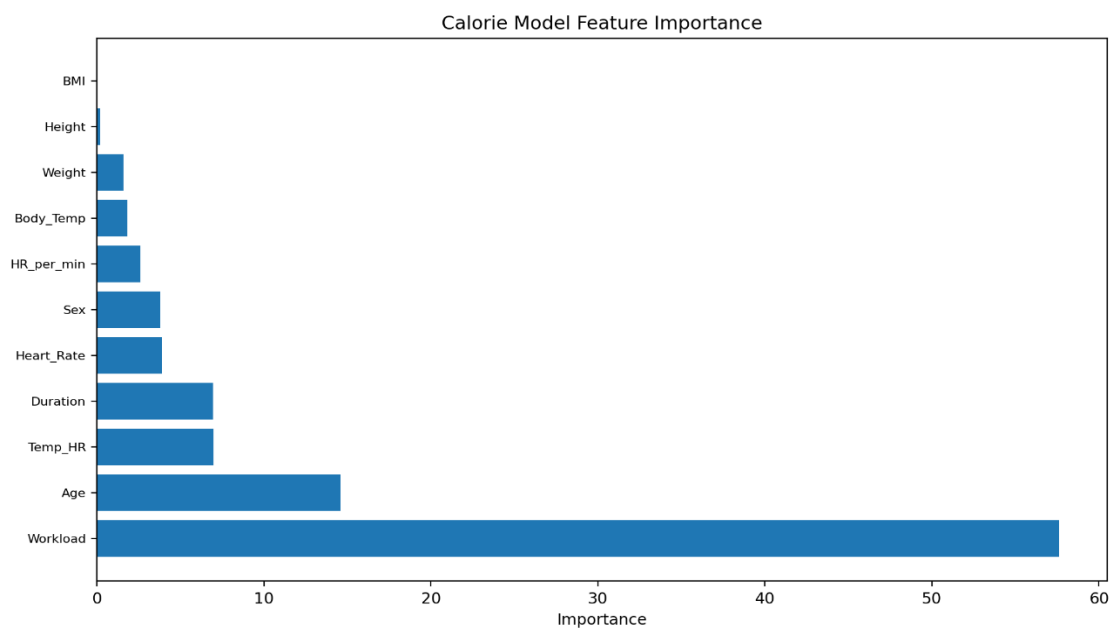


Рисунок 11. Важность признаков модели оценки калорийности.

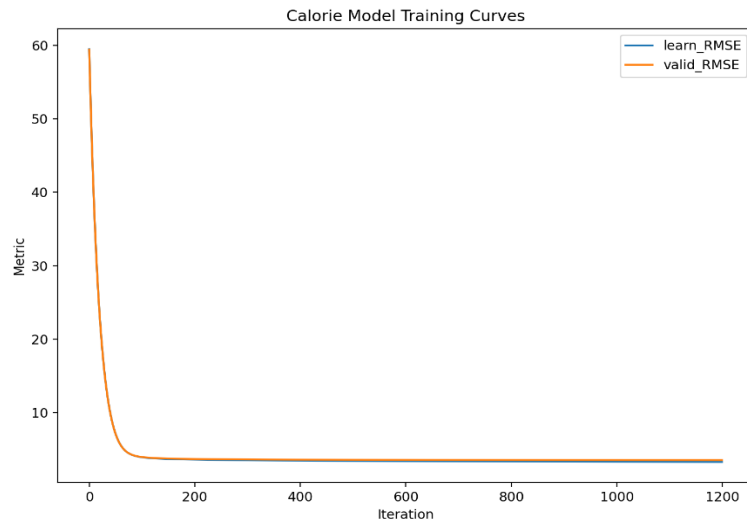


Рисунок 12. Тренировочные кривые модели оценки калорийности.

Валидация подсчёта повторений

Для независимой проверки модуля подсчёта повторений был сформирован небольшой проверочный набор коротких видеороликов, для которых истинное число повторений было размечено вручную. Системные значения, полученные методом анализа пиков сглаженных траекторий углов суставов, сравнивались с эталонной ручной разметкой (таблица 9).

Таблица 9. Сравнение автоматического подсчёта повторений с эталонной ручной разметкой.

Видеоролик	Упражнение	Эталон (вручную)	Подсчитано системой	Абс. ошибка
Демонстрация 1	Отжимания	2	2	0
Демонстрация 2	Приседания	5	5	0
Ролик 3	Прыжки на месте	12	12	0
Ролик 4	Подтягивания	6	6	0
Ролик 5	Упражнения на пресс	10	9	1

На проверочном наборе автоматический подсчёт совпал с ручной разметкой в большинстве случаев, а средняя абсолютная ошибка составила около 0,2 повторения; отклонения не превышали одного повторения и возникали преимущественно при неполной амплитуде движения или кратковременном перекрытии ориентиров. Эти результаты показывают, что предложенный подход к подсчёту повторений на основе кинематических сигналов обеспечивает практически приемлемую точность, хотя для полной количественной оценки требуется более крупный размеченный эталонный набор.

Сквозное самотестирование видео

Для проверки работоспособности предложенной структуры при объединении всех компонентов в единый конвейер на основе камер, окончательная система была протестирована на репрезентативных видеороликах с демонстрацией отжиманий и приседаний.

В обоих случаях система успешно извлекала позу тела на протяжении всего обработанного видеоряда, при этом соотношение кадров позы к кадру составляло 1,0, что указывает на стабильное обнаружение ориентиров на протяжении всего видеоряда. Что наиболее важно, окончательное предсказание типа упражнения совпало с ожидаемым типом упражнения в обоих тестах: демонстрация отжиманий была классифицирована как отжимания, а демонстрация приседаний – как приседания.

Оценки количества повторений также оказались правдоподобными для коротких тестовых видеороликов. В демонстрации отжиманий было выполнено 2 повторения, а в демонстрации приседаний – 5 повторений.

Хотя это и не эталонные эксперименты по подсчету повторений с внешними метками повторений, полученные значения соответствуют видимой картине движений в коротких оценочных видеороликах и показывают, что механизм отслеживания углов с учетом упражнений функционирует согласованно в сквозном режиме.

Не менее важным результатом является определение калорийности в процессе самотестирования. В более ранних промежуточных версиях системы короткие видеоролики иногда давали нереалистичные гибридные значения калорийности, поскольку регрессионная модель использовалась без достаточной калибровки по кратковременным данным.

В финальной версии гибридная логика оценки дала гораздо более правдоподобные значения: приблизительно 0,43 ккал для демонстрации отжиманий и 1,55 ккал для демонстрации приседаний. Эти значения являются разумными для очень коротких видеороликов и указывают на то, что стратегия калибровки с учетом движения исправила один из главных практических недостатков наивной оценки калорийности коротких видео.

С точки зрения внедрения, результаты самотестирования особенно важны, поскольку они демонстрируют, что предлагаемая система работает не только как разрозненный набор офлайн-моделей. Вместо этого, функции извлечения поз, классификации упражнений, подсчета повторений и оценки калорий объединены в одном практическом видеоконвейере.

Такое интегрированное поведение является критически важным требованием для инструментов мониторинга физической активности в реальных условиях, предназначенных для домашних тренировок, уроков физкультуры или дистанционной поддержки активности. Подробные результаты комплексного самотестирования суммированы в таблице 10, а соответствующие визуальные сводки представлены на рисунке 13.

Таблица 10. Результаты комплексного самотестирования на репрезентативных демонстрационных видеороликах с отжиманиями и приседаниями.

Видео	Ожидаемая физическая нагрузка	Прогнозируемое упражнение	Всего кадров	Кадров с позой	Должностной кадр с позой	Продолжительность (с)	Повторения	Частота повторений (повторений/мин)	МЕТ	ккал_ME_T	ккал_гибрид	Использованный сигнал	Соответствие
Демонстрация отжиманий	Отжимания	Отжимания	61	61	1,00	2,44	2	49,18	8,0	0,3985	0,4310	угол локтя	Да
Демонстрация приседаний	Приседания	Приседания	280	280	1,00	11,20	5	26,79	6,0	1,3720	1,5479	Угол колена	Да

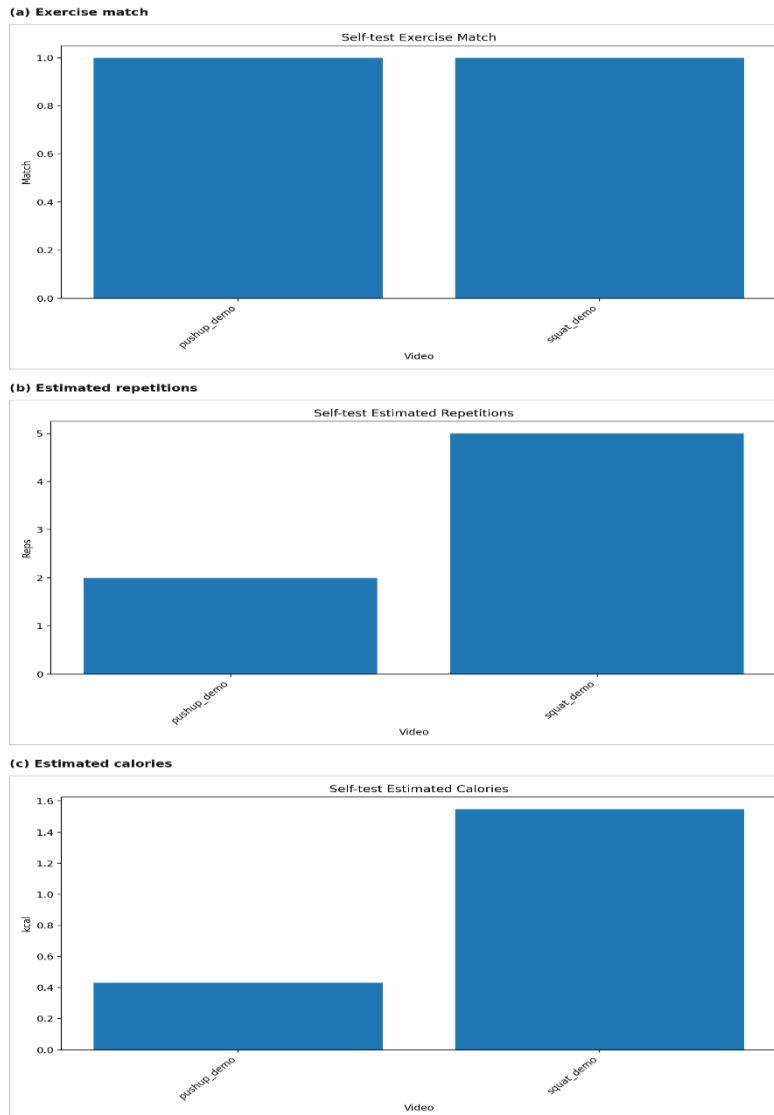


Рисунок 13. Результаты сквозного самотестирования предлагаемой системы. На панели (а) показано соответствие упражнений демонстрационным видеороликам, на панели (б) представлено расчетное количество повторений, а на панели (в) показаны соответствующие откалиброванные оценки калорий.

ОБСУЖДЕНИЕ

Результаты данного исследования подтверждают эффективность гибридной конструкции для мониторинга физических упражнений с помощью камер. Вместо того чтобы объединять все задачи в единый сквозной визуальный предиктор, предложенная структура разбивает проблему на интерпретируемые этапы. Такой выбор конструкции оказался полезным в нескольких отношениях. Во-первых, распознавание упражнений можно было оптимизировать с помощью структурированных признаков опорных точек и оценить с помощью полного пакета метрик для многоклассовой классификации. Во-вторых, подсчет повторений оставался связанным с интерпретируемыми траекториями углов суставов, а не с латентными векторами движений. В-третьих, оценка калорий выигрывала от сильного физиологического регрессора, при этом продолжая реагировать на интенсивность движения, полученную из видео, во время реального анализа. Такая модульная архитектура особенно подходит для приложений, где объяснимость, отладка и надежность развертывания имеют такое же значение, как и исходная прогностическая производительность (Perrett et al., 2022).

Результаты распознавания упражнений показывают, что структурированные признаки, полученные на основе позы, достаточно выразительны для распространенных

повторяющихся упражнений с собственным весом. Это согласуется с данными недавних исследований, демонстрирующих, что методы распознавания активности человека на основе скелета и опорных точек остаются привлекательными, поскольку они сохраняют семантику движения, будучи при этом менее чувствительными к внешнему виду фона, чем необработанные RGB-модели (Nguyen et al., 2023). Эта интерпретация также согласуется с исследованиями распознавания упражнений в тренажерном зале на основе датчиков, которые показали, что структурированные дескрипторы движения могут эффективно разделять категории повторяющихся упражнений в прикладных условиях мониторинга упражнений (Hussain et al., 2022). В контексте мониторинга физической подготовки это особенно актуально, поскольку основные различия между упражнениями носят преимущественно биомеханический характер, а не связаны с объектом или обстановкой.

Механизм подсчета повторений также иллюстрирует практическую ценность разработки сигналов, учитывающих особенности упражнений. Вместо применения общего правила подсчета ко всем движениям, система использует сигналы, определяющие преобладание колена, локтя или туловища в зависимости от конкретного упражнения. Это напрямую отражает биомеханику приседаний, отжиманий, подтягиваний и упражнений на пресс. Недавние исследования по анализу повторяющихся движений скелета аналогичным образом подчеркивают важность специфической для движения временной структуры для повышения надежности подсчета (Postlmayr et al., 2024).

Результаты оценки калорий также заслуживают внимания. Во многих практических приложениях для фитнеса расход калорий либо оценивается с помощью фиксированной эвристики, либо определяется на основе правил, зависящих от носимых устройств. В отличие от этого, предлагаемая система использует обученную регрессионную модель с высокими количественными показателями, а затем ограничивает ее поведение при развертывании посредством гибридизации с учетом движения. Это приводит к лучшему балансу между персонализацией, специфичной для конкретного человека, и правдоподобием коротких видеороликов. Такой подход согласуется с предыдущими исследованиями, показывающими, что системы визуальной оценки энергии выигрывают от калибровки и контекстуализации, а не от рассмотрения всех фрагментов движения как одинаково информативных метаболических событий (Nathan et al., 2015; Perrett et al., 2022).

В то же время, исследование имеет ряд ограничений. Во-первых, хотя окончательные классы упражнений имеют смысл для целевого варианта использования, эксперименты по классификации все же проводились на эталонных данных, полученных на основе поз, а не на большом, полностью не ограниченном реальном видеокорпусе. Во-вторых, подсчет повторений был проверен с помощью самотестирования на системном уровне, а не с помощью специального эталонного набора данных с разметкой повторений. В-третьих, оценка калорий остается гибридной, а не чисто визуальной, поскольку физиологические переменные по-прежнему являются частью окончательной логики вывода. Эти ограничения не подрывают настоящий вклад, но определяют его область применения и уточняют, что предлагаемая система должна рассматриваться как надежный практический прототип, а не как окончательный универсальный стандарт мониторинга упражнений. В будущих работах следует расширить структуру, включив в нее больше категорий упражнений, более длинные не ограниченные клипы, повысить устойчивость к различным ракурсам и добавить эталонные наборы данных для подсчета повторений на уровне последовательности.

Этические аспекты и вопросы конфиденциальности

Исследование проводилось в соответствии с принципами Хельсинкской декларации, касающимися исследований с участием человека. Демонстрационные видеозаписи, использованные для сквозного самотестирования, были получены при участии самих авторов, которые дали информированное согласие на запись, обработку и использование этих материалов в научных целях; видеозаписи третьих лиц не собирались и не использовались.

Оба набора данных, на которых обучались модели, являются общедоступными и

распространяются для исследовательских целей; они использовались в соответствии с условиями их лицензий. Набор данных для оценки калорий содержит обезличенные демографические и физиологические показатели и не позволяет идентифицировать конкретных лиц. Набор данных для распознавания упражнений представлен в виде числовых опорных точек позы, а не исходных изображений, что дополнительно снижает риски, связанные с конфиденциальностью.

Предложенная система обрабатывает движение в виде абстрактных геометрических признаков (координат опорных точек и производных от них величин) и не выполняет распознавание личности, биометрическую идентификацию или хранение исходных видеоданных. При практическом применении системы к видеозаписям реальных пользователей рекомендуется получать их информированное согласие, по возможности выполнять обработку локально и хранить только производные признаки, а не исходные изображения. Авторы заявляют об отсутствии конфликта интересов.

ЗАКЛЮЧЕНИЕ

В данном исследовании представлена гибридная система на основе машинного зрения для распознавания упражнений, подсчета повторений и оценки расхода калорий по монокулярному видео. Предложенная структура объединяет извлечение поз без маркеров, многоклассовую классификацию упражнений на основе ориентиров, подсчет повторений с учетом упражнений и калиброванный регрессор машинного обучения для прогнозирования калорий. Результаты показывают, что система демонстрирует устойчивую эффективность в обеих основных задачах при тестировании в контролируемых условиях. Классификатор упражнений достиг точности 0,8883, сбалансированной точности 0,8950 и показателя макро-F1 0,8929, в то время как модель расчета калорий достигла MAE 2,1339, RMSE 3,6202 и R^2 0,9966. В ходе сквозного самотестирования видео были правильно распознаны репрезентативные фрагменты отжиманий и приседаний, подсчет повторений был правдоподобным, а полученные оценки калорий оставались в пределах реалистичных диапазонов для коротких фрагментов.

Основной вклад исследования заключается в демонстрации возможности создания практической системы мониторинга физической активности на основе камер без полного использования носимых устройств или визуальной регрессии по принципу «черного ящика». Используя позу в качестве центрального представления, предложенная структура сохраняет интерпретируемость при анализе движений, обеспечивая при этом высокую точность прогнозирования калорий на эталонном наборе данных. Такой баланс делает систему перспективной для домашних фитнес-приложений и образовательных программ и создаёт основу для будущих расширений в области реабилитационной поддержки и цифрового мониторинга здоровья, что требует дополнительной проверки на более крупных и разнообразных выборках.

Предложенная структура также задаёт полезное методологическое направление для будущих исследований.

Модульная архитектура позволяет улучшать каждый компонент – извлечение позы, распознавание упражнений, подсчет повторений и оценку калорий – независимо, сохраняя при этом удобство использования всей системы. Поэтому в будущих исследованиях следует сосредоточиться на более полных наборах данных видеозаписей упражнений из реальной жизни, специализированных эталонных наборах данных для подсчета повторений, более широкой персонализации прогнозирования калорий и обобщении результатов на разных субъектах в более разнообразных условиях съемки.

СПИСОК ИСПОЛЬЗОВАННЫХ ЛИТЕРАТУР

Ainsworth, B. E., Haskell, W. L., Herrmann, S. D., Meckes, N., Bassett, D. R., Jr., Tudor-Locke, C., Greer, J. L., Vezina, J., Whitt-Glover, M. C., & Leon, A. S. (2011). 2011 compendium

of physical activities: A second update of codes and MET values. *Medicine & Science in Sports & Exercise*, 43(8), 1575–1581. <https://doi.org/10.1249/MSS.0b013e31821ece12>

Badiola-Bengoa, A., & Mendez-Zorrilla, A. (2021). A systematic review of the application of camera-based human pose estimation in the field of sport and physical exercise. *Sensors*, 21(18), 5996. <https://doi.org/10.3390/s21185996>

Bazarevsky, V., Grishchenko, I., Raveendran, K., Zhu, T., Zhang, F., & Grundmann, M. (2020). BlazePose: On-device real-time body pose tracking. *arXiv*. <https://doi.org/10.48550/arXiv.2006.10204>

Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, 327(8476), 307–310. [https://doi.org/10.1016/S0140-6736\(86\)90837-8](https://doi.org/10.1016/S0140-6736(86)90837-8)

Brodersen, K. H., Ong, C. S., Stephan, K. E., & Buhmann, J. M. (2010). The balanced accuracy and its posterior distribution. In *Proceedings of the 20th International Conference on Pattern Recognition* (pp. 3121–3124). IEEE. <https://doi.org/10.1109/ICPR.2010.764>

Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics*, 21, 6. <https://doi.org/10.1186/s12864-019-6413-7>

Cornman, H. L., Stenum, J., Cowley, M., Zhang, Y., & Roemmich, R. T. (2021). Video-based quantification of human movement frequency using pose estimation: A pilot study. *PLOS ONE*, 16(12), e0261450. <https://doi.org/10.1371/journal.pone.0261450>

Erramchetty, S. K., Selvam, W. P., & Veerachamy, R. (2024). Human pose estimation using BlazePose. *AIP Conference Proceedings*, 2971(1), 040049. <https://doi.org/10.1063/5.0196455>

Hancock, J. T., & Khoshgoftaar, T. M. (2020). CatBoost for big data: An interdisciplinary review. *Journal of Big Data*, 7, 94. <https://doi.org/10.1186/s40537-020-00369-8>

Hussain, A., Zafar, K., Baig, A. R., Almakki, R., AlSuwaidan, L., & Khan, S. (2022). Sensor-based gym physical exercise recognition: Data acquisition and experiments. *Sensors*, 22(7), 2489. <https://doi.org/10.3390/s22072489>

Kaggle. (n.d.). Workout/exercises pose-landmark dataset [Data set]. Kaggle. <https://www.kaggle.com/datasets>

Kaggle. (2025). Predict calorie expenditure [Data set]. Kaggle Playground Series, Season 5, Episode 5. <https://www.kaggle.com/competitions/playground-series-s5e5>

Leone, A., Caroppo, A., Rescio, G., Manni, A., Siciliano, P., & Diraco, G. (2022). Ambient and wearable sensor technologies for energy expenditure quantification of ageing adults. *Sensors*, 22(13), 4893. <https://doi.org/10.3390/s22134893>

Lin, B.-S., Wang, L.-Y., Hwang, Y.-T., Chiang, P.-Y., & Chou, W.-J. (2021). Depth-camera based energy expenditure estimation system for physical activity using posture classification algorithm. *Sensors*, 21(12), 4216. <https://doi.org/10.3390/s21124216>

Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., Zhang, F., Chang, C.-L., Yong, M. G., Lee, J., Chang, W.-T., Hua, W., Georg, M., & Grundmann, M. (2019). MediaPipe: A framework for building perception pipelines. *arXiv*. <https://doi.org/10.48550/arXiv.1906.08172>

Mercadal-Baudart, C., Liu, C.-J., Farrell, G., Boyne, M., González Escribano, J., Smolic, A., & Simms, C. (2024). Exercise quantification from single camera view markerless 3D pose estimation. *Heliyon*, 10(6), e27596. <https://doi.org/10.1016/j.heliyon.2024.e27596>

Nathan, D., Klasnja, P., Vankipuram, M., Momboisse, T., Powell, C., & Raju, D. (2015). Estimating physical activity energy expenditure with the Kinect sensor in an exergaming environment. *PLOS ONE*, 10(5), e0127113. <https://doi.org/10.1371/journal.pone.0127113>

Nguyen, H.-C., Dang, D.-T., Ngo, T.-D., & Pham, D.-T. (2023). Deep learning for human activity recognition on 3D human skeleton: Survey and comparative study. *Sensors*, 23(11), 5121. <https://doi.org/10.3390/s23115121>

Perrett, T., Masullo, A., Damen, D., Burghardt, T., Craddock, I., & Mirmehdi, M. (2022). Personalized energy expenditure estimation: Visual sensing approach with deep learning. *JMIR Formative Research*, 6(9), e33606. <https://doi.org/10.2196/33606>

Postlmayr, A., Garg, B., Cosman, P., & Dey, S. (2024). PersonalPT: One-shot approach for skeletal-based repetitive action counting for physical therapy. *Smart Health*, 34, 100516. <https://doi.org/10.1016/j.smhl.2024.100516>

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: Unbiased boosting with categorical features. In *Advances in Neural Information Processing Systems* 31 (pp. 6638–6648). Curran Associates. <https://doi.org/10.48550/arXiv.1706.09516>

Roggio, F., Trovato, B., Sortino, M., & Musumeci, G. (2024). A comprehensive analysis of the machine learning pose estimation models used in human movement and posture analyses: A narrative review. *Heliyon*, 10(21), e39977. <https://doi.org/10.1016/j.heliyon.2024.e39977>

Авторлар туралы мәліметтер
Информация об авторах
Information about authors

№	Информация об авторах (Ф.И.О. полностью, ученая степень, место работы или учебы, город, страна, E-mail, сотовый номер автора корреспондента:) на казахском, русском, английском языках
1	Абдразах Ынтымақ Қыдырәліұлы – техника ғылымдарының магистрі, Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті, Түркістан қ., Қазақстан Абдразах Ынтымақ Қыдырәліұлы – магистр технических наук, Международный казахско-турецкий университет им. Ходжи Ахмеда Ясави, г. Туркестан, Казахстан Abdrazakh Yntymak Kydyraliuly – master of technical sciences, Ahmet Yassawi University, Turkestan, Kazakhstan E-mail: yntymak.abdrazakh@ayu.edu.kz ORCID: 0000-0002-7119-1217
2	Баймаханова Айгерим Саттаровна – PhD, аға оқытушы, Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті, Түркістан қ., Қазақстан Баймаханова Айгерим Саттаровна – PhD, старший преподаватель, Международный казахско-турецкий университет имени Ходжи Ахмеда Ясави, г. Туркестан, Казахстан Baimakhanova Aigerim Sattarovna – PhD, Senior Lecturer, Khoja Akhmet Yassawi International Kazakh-Turkish University, Turkistan, Kazakhstan E-mail: aygerim.baymakhanova@ayu.edu.kz, ORCID: 0000-0002-5364-0146
3	Жунисов Н.М. – PhD доктор, аға оқытушы, Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті, Түркістан қ., Қазақстан Жунисов Н.М. – доктор PhD, старший преподаватель, Международный казахско-турецкий университет имени Ходжи Ахмеда Ясави, г. Туркестан, Казахстан Zhunisoov N.M. – PhD, Senior Lecturer, Khoja Akhmet Yassawi International Kazakh-Turkish University, Turkistan, Kazakhstan E-mail: nurseit.zhunissov@ayu.edu.kz ORCID: 0000-0001-6531-9408
4	Мирзалиева Шахзода Алишерқызы – студент, Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті, Түркістан қ., Қазақстан Мирзалиева Шахзода Алишерқызы – студент, Международный казахско-турецкий университет им. Ходжи Ахмеда Ясави, г. Туркестан, Казахстан Mirzaliyeva Shakhzoda Alisherqyzy – student, Ahmet Yassawi University , Turkestan, Kazakhstan E-mail: mirzaliyeva.shakhzoda2025@ayu.edu.kz