

УДК 004.4

ПРИМЕНЕНИЕ НЕЙРОННЫХ СЕТЕЙ ДЛЯ РАСПОЗНАВАНИЯ И КОНТРОЛЯ ПОЗЫ ТЕЛА ЧЕЛОВЕКА НА ВИДЕО ПОТОКЕ

Разуминин В.В. (Университет ИТМО), **Анохин И.И.** (Университет ИТМО)

Научный руководитель – Гаврилов А. В., Преподаватель факультета ПИиКТ

Университет ИТМО

Введение

Благодаря стремительному развитию технологий, возникают новые возможности для автоматизации процессов общества и бизнеса. Так, машинное видение уже сейчас может быть эффективным решением для осуществления контроля качества и корректности выполнения различных физических действий человеком на видео дорожке. В том числе в режиме реального времени.

Использование подобных технологий будет полезным для различных областей применения, таких например как: спорт, развлечения, здравоохранение, безопасность. Обучение различным спортивным навыкам, оценка исполнения технических элементов спортивными судьями и жюри, контроль поведения сотрудников, детей, животных - всё это примеры потенциального использования разрабатываемого программного обеспечения.

Спектр применения рассматриваемых технологий велик, как и спектр возможных технических реализаций. Данная работа направлена на исследование современных подходов и методов к осуществлению треккинга виртуального скелета человека для дальнейшей обработки и преобразования полученных данных.

Цель работы

Целью работы является разработка прототипа мобильного приложения на операционных системах IOS и Android, которое сможет распознавать выполнение человеком определенных физических действий. Примером такого физического действия будет являться комплекс физических упражнений общеоздоровительной подготовки: сгибание-разгибание рук и ног, наклоны тела.

На основе выбранного примера появится возможность сравнить виды реализации алгоритмов распознавания и контроля позы человека в кадре, определив актуальные для рынка технические решения.

Базовые положения исследования

Поскольку задача распознавания положения тела человека на видео потоке актуальна для разных профессиональных областей применения, основной идеей работы является разбор существующих подходов к реализации широкого спектра типовых бизнес-процессов, связанных с физической активностью человека.

Для осуществления наиболее эффективного распознавания позы человека в кадре, были использованы методы машинного зрения. Основными требованиями к подходу являлись: скорость выполнения программы и способность к быстрому масштабированию без резкого возрастания нагрузки на вычислительные мощности.

Промежуточные результаты

Для легкой разработки мобильного приложения был использован язык программирования Dart, фреймворк Flutter. С помощью него удалось достичь 100% совместимости прототипа как со смартфонами на базе операционной системы iOS, так и на базе Android.

Немаловажно, что Flutter предоставляет возможность писать и использовать низкоуровневые операции различных систем, путём написания нативного кода, что позволяет использовать для каждой системы различные нативные оптимизации. Также, сам Flutter хорошо оптимизирован, что позволяет ему поддерживать частоту обновления мобильного приложения в 60 кадров в секунду.

Для реализации работы приложения было использовано 2 подхода: облачные вычисления положения человека в кадре и вычисления на конечном устройстве пользователя.

Для облачных вычислений использовался следующий стек технологий: Python, Flask, PyTest. Также сама система была развернута на сервере с видеокартой, что позволило использовать GPU для ускорения процесса видеобработки. Распознавание движений было реализовано с помощью PyTorch и модели OpenPose.

Для реализации вычислений на телефоне использовалась библиотека tflite, которая даёт доступ к использованию TensorFlow Lite. Для определения человека в кадре использовалась модель PoseNet, оптимизированная для мобильных устройств.

Эффективность использования библиотеки TensorFlow Lite подтверждают исследования, в которых она успешно применяется для решения похожих задач.

В ходе исследования было замечено, что решение для мобильных устройств имеет более высокий приоритет, так как позволяет обрабатывать видеопоток в режиме реального времени, без необходимости использования интернета на телефоне пользователя. Также при таком подходе не нужно содержать серверные мощности.

Основной результат

В результате работы был разработан прототип кросс-платформенных приложений для операционных платформ IOS и Android с использованием передовой технологии распознавания позы тела человека в кадре. Проект был пилотно протестирован на фокус группе, проведены проблемное и решенческое интервью ключевых сегментов целевой аудитории.

На основе обратной связи пользователей и подтвержденных гипотезах ценности, проект был одобрен к дальнейшему развитию: созданию полноценной бизнес-модели и виральной концепции для осуществления маркетинговых кампаний.

Разуминин В.В. (автор)

Подпись

Анохин И.И. (соавтор)

Подпись

Гаврилов А. В. (научный руководитель)

Подпись