МИНОБРНАУКИ РОССИИ

Федеральное государственное автономное образовательное   
учреждение высшего образования   
«Южный федеральный университет»

Институт математики, механики   
и компьютерных наук им. И. И. Воровича

Кафедра прикладной математики и программирования

Дурманенко Всеволод Сергеевич

Оценка положения тела на видео

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА  
по направлению подготовки  
02.03.02 – Фундаментальная информатика и информационные технологии

**Научный руководитель** –   
доц. Демяненко Яна Михайловна

Допущено к защите:  
заведующий кафедрой \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Сидоров С. С.

Ростов-на-Дону – 2024

Оглавление

[Введение 3](#_Toc164363905)

[1. Постановка задачи 4](#_Toc164363906)

[2. Обзор подходов 5](#_Toc164363907)

[2.1. Сверточная нейросеть 5](#_Toc164363908)

[2.2. Методы обнаружения человека 5](#_Toc164363909)

[2.2.1. OpenPose 5](#_Toc164363910)

[2.2.2. AlphaPose (RMPE) 6](#_Toc164363911)

[2.2.3. DeepCut 8](#_Toc164363912)

[2.2.4. Mask R-CNN 9](#_Toc164363913)

[2.2.5. MediaPipe 10](#_Toc164363914)

[2.3. Представление анатомии человека для компьютера 11](#_Toc164363915)

[2.3.1. Моделирование анатомии человека 11](#_Toc164363916)

[3. Основная часть 13](#_Toc164363917)

[3.1. Идея решения 13](#_Toc164363918)

[3.2. Разработка алгоритма 13](#_Toc164363919)

[3.3. Статистика 15](#_Toc164363920)

[3.4. Графический интерфейс 15](#_Toc164363921)

[3.5. Результаты 16](#_Toc164363922)

[Приложения 17](#_Toc164363923)

[Заключение 18](#_Toc164363924)

[Литература 19](#_Toc164363925)

# Введение

В современном мире, где активный образ жизни и забота о здоровье становятся все более важными, фитнес-индустрия находится в центре внимания, предлагая различные инновационные способы поддержания физической формы. В рамках этого контекста разработка приложений становится неотъемлемой частью современной фитнес-культуры, обеспечивая пользователей инструментами для эффективного тренировочного процесса и отслеживания своих достижений.

Цель данной дипломной работы заключается в разработке и реализации фитнес-приложения, предназначенного для отслеживания выполнения упражнений. Данное приложение будет ориентировано на широкий круг пользователей, начиная от начинающих заниматься фитнесом до опытных спортсменов, стремящихся оптимизировать свои тренировочные процессы.

Целью данного проекта является не только создание удобного и эффективного инструмента для тренировок, но и способствование привлечению большего числа людей к здоровому образу жизни, путем предоставления им доступных и мотивирующих средств для занятий спортом.

# Постановка задачи

Целью данной работы является разработка приложения для правильного выполнения упражнений, связанных с разминкой и растяжкой. Для достижения данной цели можно выделить следующие задачи:

* Исследовать существующие методы анализа видеоданных и оценки положения объектов на видео.
* Разработать алгоритм для определения правильности выполнения упражнений, используя данные из изученных методов.
* Разработать метод ведения статистики для отслеживания прогресса.
* Разработать графический интерфейс.

# Обзор подходов

## Сверточная нейросеть

Сверточная нейронная сеть (CNN) представляет собой специализированный тип нейронной сети, применяемый в основном для обработки изображений и других данных с пространственной структурой, таких как аудио и видео. Основное отличие CNN от обычной нейронной сети заключается в способе обработки входных данных.

Основные концепции сверточных нейронных сетей включают в себя сверточные слои, которые обнаруживают локальные шаблоны, такие как грани и текстуры, путем скольжения окна фильтра по входным данным. Затем следует операция подвыборки, которая уменьшает размерность данных, сохраняя наиболее важные признаки.

После каждого сверточного или подвыборочного слоя обычно применяется нелинейная функция активации, такая как ReLU, для введения нелинейности в модель. И, наконец, после серии сверточных и подвыборочных слоев следует один или несколько полносвязных слоев, которые объединяют глобальные признаки для принятия окончательного решения.

В целом, сверточная нейронная сеть отличается от обычной тем, что она эффективно обрабатывает данные с пространственной структурой, сохраняя локальные шаблоны и признаки, что делает ее более подходящей для задач компьютерного зрения и обработки изображений.

## Методы обнаружения человека

### OpenPose

OpenPose – это восходящий подход, при котором сеть сначала обнаруживает части тела или ключевые точки на изображении, а затем сопоставляет соответствующие ключевые точки для формирования пар.

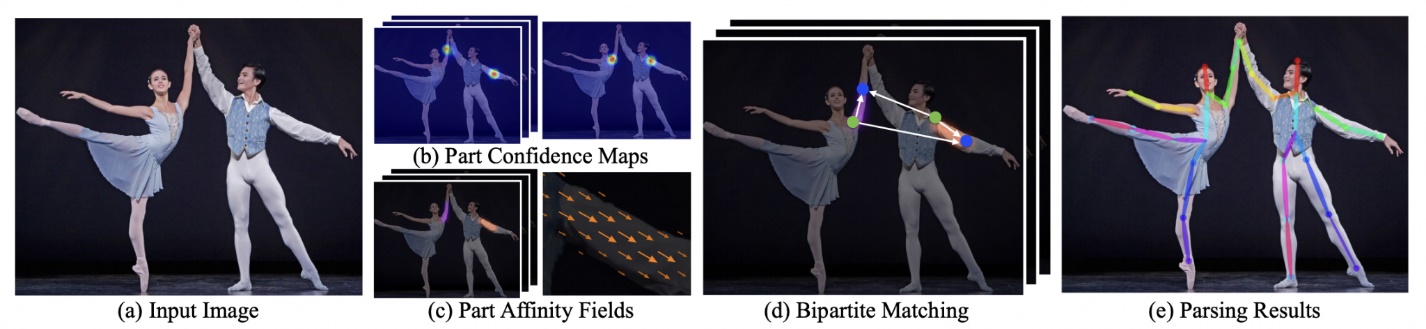
OpenPose также использует CNN в качестве основной архитектуры. Он состоит из сверточной сети VGG-19, которая используется для извлечения шаблонов и представлений из заданных входных данных. Выходные данные VGG-19 попадают в две ветви сверточных сетей.

Первая сеть прогнозирует набор карт достоверности для каждой части тела, а вторая ветвь прогнозирует поля сходства частей (PAF), которые создают определенную степень связи между частями. Также полезно удалить более слабые звенья в двудольных графах.

(фото)

На изображении выше показана архитектура OpenPose, которая представляет собой многоступенчатую CNN.

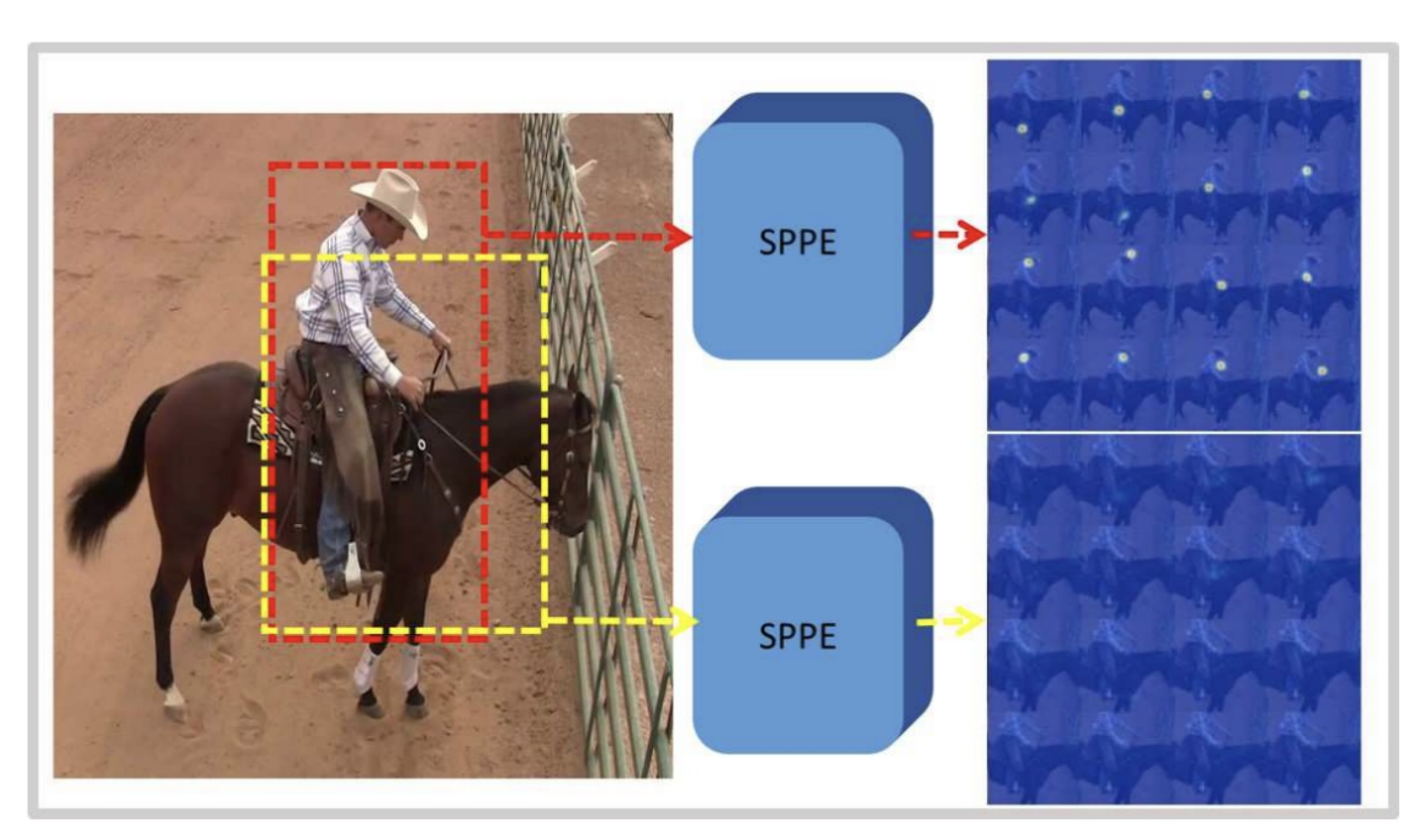
По сути, предсказания двух ветвей вместе с признаками объединяются на следующем этапе, чтобы сформировать человеческий скелет в зависимости от количества людей, присутствующих во входных данных. Последовательные этапы CNN используются для уточнения прогноза.



### AlphaPose (RMPE)

Региональная оценка положения нескольких человек (RMPE) или AlphaPose реализует нисходящий подход к HPE.

Нисходящий подход к HPE приводит к большому количеству ошибок при локализации и неточностям во время прогнозирования и, следовательно, является довольно сложной задачей.



Например, на изображении выше показаны две ограничивающие рамки: красная рамка представляет достоверное выделение, а желтая рамка представляет прогнозируемую ограничивающую рамку.

Хотя, когда дело доходит до классификации, желтая ограничивающая рамка будет считаться «правильной» ограничивающей рамкой для классификации человека. Однако позу человека невозможно оценить даже с помощью «правильной» ограничивающей рамки.

Авторы AlphaPose решили проблему несовершенного обнаружения людей с помощью двухэтапной схемы. В рамках этой структуры они представили две сети:

Симметричная пространственная трансформаторная сеть (SSTN): помогает вырезать соответствующую область на входе, что впоследствии упрощает задачу классификации, что приводит к повышению производительности.

Оценщик позы одного человека (SPPE): он используется для извлечения и оценки позы человека.

Цель AlphaPose — извлечь высококачественную область с одним человеком из неточной ограничивающей рамки путем присоединения SSTN к SPPE. Этот метод повышает эффективность классификации, обеспечивая инвариантность и обеспечивая при этом стабильную основу для оценки позы человека.



### DeepCut

DeepCut был предложен Леонидом Пищулиным и др. ал. в 2016 году с целью совместного решения задач обнаружения и оценки позы одновременно.

Это восходящий подход к оценке позы человека.

Идея заключалась в том, чтобы обнаружить все возможные части тела на данном изображении, затем обозначить их, например, голову, руки, ноги и т. д., с последующим процессом разделения частей тела, принадлежащих каждому человеку.



Сеть использует моделирование интегрального линейного программирования (ILP ) для неявной группировки всех обнаруженных ключевых точек в заданных входных данных так, чтобы конечный результат напоминал скелетное представление человека.

### Mask R-CNN

Модель Mask R-CNN имеет возможность одновременно локализовать и классифицировать объекты, создавая ограничивающую рамку вокруг объекта, а также создавая маску сегментации.

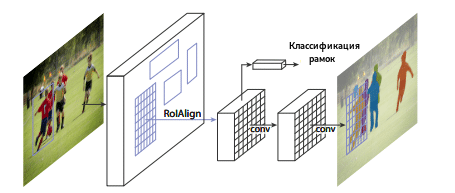


Fast R-CNN использует CNN для извлечения функций и представлений из заданных входных данных.

Извлеченные функции затем используются, чтобы предложить, где объект может присутствовать через сеть предложений регионов (RPN).

Поскольку ограничивающая рамка может иметь разные размеры, как на изображении выше, слой под названием RoIAlign используется для нормализации извлеченных объектов, чтобы все они имели одинаковые размеры.

Извлеченные признаки передаются в параллельные ветви сети для уточнения предлагаемой области интереса (RoI) для создания ограничивающих рамок и масок сегментации.



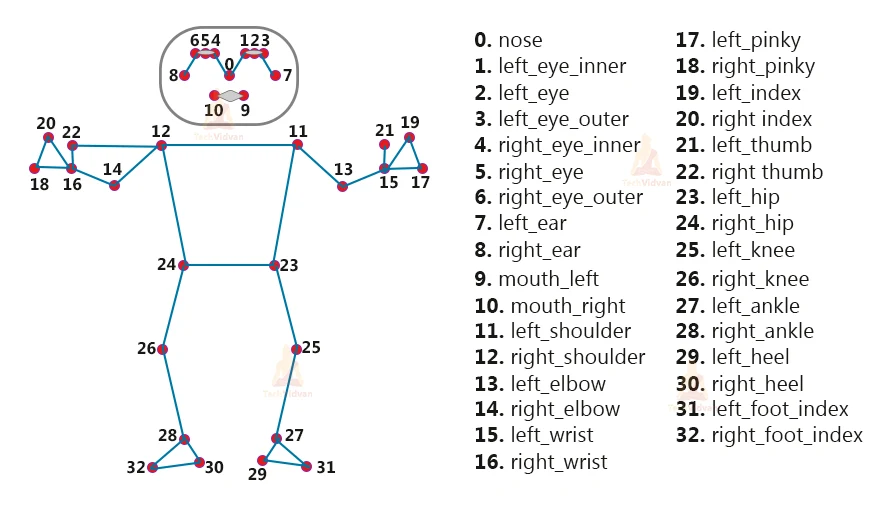
Когда дело доходит до оценки позы человека, выходные данные сегментации маски, полученные сетью, можно использовать для обнаружения людей на заданных входных данных. Поскольку сегментация по маске очень точна при обнаружении объектов, в данном случае при обнаружении человека, позу человека можно оценить довольно легко.

Этот метод напоминает подход «сверху вниз», при котором этап обнаружения человека выполняется параллельно этапу обнаружения деталей.

Другими словами, этап обнаружения ключевой точки и этап обнаружения человека независимы друг от друга.

### MediaPipe

MediaPipe Pose — это система оценки позы одного человека. Она использует топологию ориентиров BlazePose 33. BlazePose — это расширенный набор ключевых точек COCO, топологии Blaze Palm и Blaze Face. В свою очередь COCO – это набор данных, широко используемый для оценки алгоритмов компьютерного зрения, особенно для задачи обнаружения объектов и сегментации. Набор данных COCO содержит большой набор изображений, размеченных для различных задач компьютерного зрения. Алгоритм работает в два этапа – обнаружение и отслеживание. Поскольку обнаружение не выполняется в каждом кадре, MediaPipe может выполнять вывод быстро.



## Представление анатомии человека для компьютера

Определение положения и ориентации различных частей человеческого тела на изображениях или видео - это ключевая задача в области компьютерного зрения и искусственного интеллекта. Эта задача имеет широкий спектр применений, включая разработку систем для управления роботами, анализа видеонаблюдения, виртуальной реальности, медицинской диагностики и даже в области спортивных технологий. Оценка положения тела позволяет компьютерным системам понимать и интерпретировать человеческое движение, что является важным шагом к созданию более умных и автономных устройств.

### Моделирование анатомии человека

В большинстве методов используется жесткая кинематическая модель N-суставов, в которой человеческое тело представлено как объект с суставами и конечностями, содержащими кинематическую структуру тела и информацию о форме тела.

Существуют три типа моделирования анатомии человека:

* **Kinematic model** (Скелетная модель) - это модель также известная как скелетно-ориентированная модель, и она применяется для анализа позы в форматах 2D и 3D. Эта удобная и интуитивно понятная модель человеческого тела включает набор суставов и ориентацию конечностей для описания структуры тела. Таким образом, модели, оценивающие положение скелета, используются для улавливания взаимосвязей между различными частями тела. Однако кинематические модели имеют ограничения в передаче информации о текстуре или форме.
* **Planar model** (Плоская модель) – это модель, известная также как модель на основе контура, применяется для оценки 2D позы. Плоские модели используются для представления внешнего вида и формы человеческого тела. Обычно различные части тела представлены несколькими прямоугольниками, которые приближены к контуру человеческого тела. Хорошим примером такой модели является модель активных форм, которая используется для охвата полного контура человеческого тела и его деформации с помощью анализа главных компонентов.
* **Volumetric model** (Объемная модель) – это объемная модель, применяемая для трехмерной оценки положения, представляет собой важный инструмент в анализе поз. Они предназначены для моделирования статистических и артикулированных форм человеческого тела и поз. Эти модели могут быть использованы для вывода различных характеристик и параметров позы.



# Основная часть

## Идея решения

В первую очередь необходимо определить тело человека и ключевые точки на видео. Под ключевыми точками подразумеваются части тела: локти, стопы, нос и т.д. Для отслеживания правильности выполнения упражнения необходимо рассчитать угол между ключевыми точками, соответствующими текущему упражнению.

## Разработка алгоритма

Для решения поставленной задачи был выбран язык программирования Python с библиотеками Mediapipe, OpenCV и PyQt.

Алгоритму на вход передается изображение, полученное с камеры. MediaPipe обрабатывает данное изображение. В результате получаются координаты ключевых точек и коэффициенты их видимости. Формат, в котором Mediapipe предоставляет данные, не очень удобен для дальнейшей работы. В связи с этим был создан класс Workout, в котором используется словарь в формате string:Point. Ключом здесь является название части тела, а значением – точка, содержащая координаты расположения этой части тела и коэффициенты видимости. Данный класс также содержит методы для определения выполнения упражнений и вспомогательный метод, отвечающий за то, что человек находится в кадре целиком. Для удобства понимания стадии выполнения упражнения был создан Enum класс или класс с перечислением. В нем содержатся такие фазы выполнения упражнения как начальная, в процессе выполнения и законченная.

Основная логика алгоритма находится в функции display\_frame класса VideoWindow. Данная функция отвечает за отображение изображения с камеры. В качестве параметра данному методу передается словарь упражнений. В первую очередь необходимо проверить, остались ли невыполненные упражнения. Если таковых нет, то выводится сообщение о том, что комплекс выполнен, и окно можно закрывать. Также в конце происходит подсчет выполненных повторений для каждого упражнения, и эта информация записывается в файл data.csv для статистики. Если упражнения еще остались, то изображение с камеры преобразуется и происходит поиск ключевых точек. На данном этапе необходимо убедиться, что тело полностью находится в кадре, иначе возникает исключение и выводится сообщение о том, что необходимо сместиться. Если все в порядке, то на экран выводится список оставшихся к выполнению упражнений. Происходит проверка фазы выполнения, и если она имеет значение законченной, то из словаря с упражнениями вычитается одно повторение для текущего упражнения. В случае если количество оставшихся повторений равно нулю, то данное упражнение из словаря убирается. В конце изображение появляется и видно выделенные ключевые точки.

## Упражнения

## Статистика

Как было сказано ранее, статистика является неотъемлемой частью тренировок. В данном случае для ее отображения используется библиотека Matplotlib и Pandas. Исходные данные хранятся в файле data.csv, куда информация поступает при завершении комплекса упражнений. Для отображения гистограммы данные из этого файла считываются при помощи функции read\_csv из Pandas. Полученную информацию уже преобразуется Matplotlib. В результате можно увидеть наглядный результат тренировок.

## Графический интерфейс

Для создания графического интерфейса была использована библиотека PyQt5. PyQt5 – это набор Python-обёрток над Qt, фреймворком для создания графических интерфейсов. Выбор был сделан в пользу данной библиотеки из-за простоты её использования и кроссплатформенности.

Для создания основного графического окна был создан класс MainWindow, унаследованный от QMainWindow. Внутри вызывается конструктор, в котором определяются элементы интерфейса и их расположение. Для вертикального расположения элементов используется QVBoxLayout, обернутый в QWidget. Для удобства представления списка упражнений и заданных повторений используется таблица (QTableWidget). Чтобы перемещать элементы внутри таблицы используются две кнопки (QPushButton). Это нужно для того, чтобы задать порядок выполнения упражнений. PyQt5 содержит некоторое количество изображений, которые можно использовать для создания понятного графического интерфейса. В данном случае оттуда были взяты изображения стрелок вверх и вниз. Кроме того, пользователь может устанавливать количество повторений не только вручную, но и использовать уже заготовленные наборы упражнений. Для этого необходимо открыть выпадающий список (QHBoxLayout). И выбрать желаемый комплекс упражнений на утро, день или вечер. Одним из условий прогресса является статистика, которая отображает ваши достижения. Чтобы увидеть свой прогресс на протяжении определенного периода, нужно воспользоваться кнопкой “Показать статистику”. При нажатии будет отображена гистограмма по датам и количеству выполненных упражнений. В случае отсутствия файла со статистикой вы получите сообщение об этом. В самом конце графического окна расположена кнопка для того, чтобы начать тренировку. При нажатии на данную кнопку произойдет парсинг данных из таблицы и будет запущено другое окно, в котором присутствует изображение с камеры и отображаются выбранные упражнения.

Основная логика приложения заложена именно в классе VideoWindow. Для получения изображения с камеры используется библиотека OpenCV. Видеопоток, полученный оттуда передается в QLabel. И каждые 30мс обновляется при помощи QTimer. Такая обертка нужна для того, чтобы окна могли работать друг с другом и реагировать на определенные события. Например, при попытке открыть еще одно окно с выводом изображения будет выведено сообщение о том, что необходимо закрыть уже открытое окно.

Сам список упражнений отображается в левом верхнем углу по порядку, заданному пользователем. При завершении комплекса упражнений можно увидеть сообщение об этом. Также для начала выполнения упражнений необходимо, чтобы тело полностью помещалось в камеру. В противном случае можно увидеть сообщение о том, что необходимо изменить свое положение.

## Результаты

…

# Приложения

# Заключение

В заключении следует отразить ключевые аспекты и достижения приложения, а также его потенциальное влияние на пользователей.

Приложение для отслеживания выполнения упражнений представляет собой мощный инструмент для улучшения физической активности и заботы о здоровье. За время разработки данного приложения были решены ряд ключевых задач:

* Разработан алгоритм, позволяющий распознавать различные упражнения и оценивать их правильность выполнения.
* Создан пользовательский интерфейс, интуитивно понятный и удобный в использовании, что способствует привлечению широкой аудитории пользователей.
* Реализована система ведения статистики выполнения упражнений, которая позволяет пользователям отслеживать свой прогресс, оценивать эффективность тренировок и планировать дальнейшие занятия.

Эти достижения делают приложение не только эффективным инструментом для тренировок, но и мощным мотиватором для пользователей. Предоставление обратной связи о правильности выполнения упражнений и возможность наблюдать свой прогресс стимулируют регулярные тренировки и помогают пользователям достигать своих фитнес-целей.

В целом, разработанное приложение открывает новые перспективы для здорового образа жизни и физического развития, подтверждая свою значимость как для отдельных пользователей, так и для общества в целом.

# Литература

…