

Министерство образования и науки Российской Федерации  
Московский физико-технический институт (национальный  
исследовательский университет)

Физтех-школа радиотехники и компьютерных технологий  
Кафедра интеллектуальных информационных систем и технологий

Выпускная квалификационная работа магистра

Исследование методов доменной адаптации для  
улучшения распознавания ключевых точек на теле  
человека

**Автор:**

Студент M01-205a группы  
Токарев Андрей Сергеевич

**Научный руководитель:**

Доктор технических наук  
Назаров Алексей Николаевич

**Научный консультант:**

Ст. Преподаватель  
Воронков Илья Михайлович



Москва 2024

**Аннотация**

Исследование методов доменной адаптации для  
улучшения распознавания ключевых точек на теле  
человека

*Токарев Андрей Сергеевич*

Краткое описание задачи и основных результатов,  
мотивирующее прочитать весь текст

# Содержание

<b>1</b>	<b>Введение</b>	<b>5</b>
<b>2</b>	<b>Сверточные нейронные сети</b>	<b>8</b>
<b>3</b>	<b>Распознавание ключевых точек</b>	<b>9</b>
3.1	Ключевые точки . . . . .	9
3.2	Распознавание ключевых точек на теле человека. Скелет человека . . . . .	11
<b>4</b>	<b>Domain Adaptation</b>	<b>15</b>
4.1	Перенос обучения . . . . .	15
4.2	Доменная адаптация . . . . .	17
4.2.1	Классификация методов доменной адаптации . . . .	17
<b>5</b>	<b>Обзор существующих решений</b>	<b>18</b>
5.1	Обзор моделей для распознавания КТ . . . . .	18
5.1.1	DeepPose . . . . .	18
5.1.2	AlphaPose . . . . .	18
5.1.3	OpenPose . . . . .	18
5.1.4	BlazePose . . . . .	18
5.1.5	HRNet . . . . .	18
5.1.6	ViTPose . . . . .	19
5.1.7	RTMPose . . . . .	19
5.1.8	Swin . . . . .	19
5.1.9	Simcc + Resnet . . . . .	19
5.1.10	Dekr + HRNet . . . . .	19
5.1.11	YoloPose . . . . .	20
5.2	Анализ и выбор методов доменной адаптации . . . . .	20
5.2.1	PUL . . . . .	20

5.2.2	RegDA . . . . .	20
5.2.3	UDA PoseEstimation . . . . .	20
5.2.4	POST . . . . .	20
5.2.5	SFDA . . . . .	21
<b>6</b>	<b>Эксперимент</b>	<b>22</b>
6.1	Описание деталей эксперимента . . . . .	22
6.2	Данные. Сбор и разметка . . . . .	22
6.3	Результаты эксперимента . . . . .	23
<b>7</b>	<b>Заключение</b>	<b>24</b>
	<b>Список литературы</b>	<b>25</b>

# 1 Введение

Здесь необходимо описать описание проблемы. Ее актуальность. Где это можно использовать и что делать. Заворожить читателя для прочтения твоей работы. ОПИСАТЬ научную НОВИЗНУ работы.

Современные технологии машинного обучения и компьютерного зрения продолжают активно развиваться, находя применение в самых разнообразных областях. Одно из направлений, активно развивающихся в последние годы, является решение задачи распознавания ключевых точек на теле человека (Keypoint Detection) или оценка позы человека (Human Pose Estimation). На сегодняшний день решения данной задачи могут иметь множество практических применений. (, включая системы наблюдения, анимацию, медицинскую диагностику и интерактивные интерфейсы.)

Одной из возможностей использовать распознавание позы человека является виртуальная реальность. Оцифровка позы человека с помощью неросетей позволяет экономить на закупке дорогостоящих костюмов. Можно установить несколько камер, которые будут восстанавливать позу человека и переносить ее компьютерное пространство. При добавлении генеративных алгоритмов можно создавать всевозможные аватары и погрузиться в "Оазис" из фильма Стивена Спилберга "Первому игроку приготовиться".

Другим уже реальным применением данной технологии является рефери спортивных соревнований. Уже сейчас система полуавтоматического определения оффсайда помогает судьям футбольных матчей по всему миру. А работает она на распознавании ключевых точек чисте тела, которыми футболист может сыграть в мяч и определяет были ли нарушены правила или гол был забит чисто. (ССЫЛКА)

Недалеко отходя от спорта, можно применять оценку позы для со-

здания личных тренеров прямо в телефоне. Проект MediaPipe (ССЫЛКА) предлагает уже сейчас возможности для использования его моделей на смартфонах для подсчета количества выполнений таких упражнений, как отжимания, приседания или подтягивания, а также дает возможность оценивать правильность различных асан в йоге.

## ЕЩЕ ПРИМЕРОВ ПРО ИСПОЛЬЗОВАНИЕ ЗАДАЧИ ОЦЕНКИ ПОЗЫ

Обучение модели и разработка алгоритма её работы представляют собой чрезвычайно сложный и трудоемкий процесс. Этот процесс требует значительных ресурсов, как со стороны специалистов, так и в плане вычислительной мощности. Сначала необходимо собрать и подготовить данные, затем обучить модель, настроить её параметры и протестировать на различных наборах данных, чтобы убедиться в её точности и надёжности. Это часто занимает много времени и требует значительных финансовых вложений.

Данные, на которых модели обучаются имеют общий характер и могут не подходить под определенную узко-специализированную задачу. Например, в футболе не особо важны ключевые точки рук и лица, но важны ключевые точки тела, ног и общая точка головы. Для адаптации модели под данную задачу необходимо проделать гигантский объем работы по сбору данных, их фильтрации и разметке. Особенно выделяется последняя часть работы, так как она требует работы нескольких экспертов, которые расставят точки на каждой отобранной фотографии или кадре видео.

Для избежания дополнительных работ по адаптации широкоспециализированной модели к узкой задаче ученые задумались над вопросом приспособливания модели к новым запросам. Таким образом появилась такая область науки о нейросетях, как адаптация модели к новым доме-

нам данных (англ. Domain Adaptation). ЧТО ТАКОЕ ДОМЕНЫ. КАК  
МОЖНО АДАПТИРОВАТЬ. ЧТО МОЖНО ИЗ ЭТОГО ПОЛУЧАТЬ.

ДУМАЮ ТУТ ЕЩЁ СТОИТ СКАЗАТЬ ПРО ПЕРЕНОС ОБУЧЕ-  
НИЯ.

В конце стоит описать план по главам и что там будет описано. Кратко  
и по делу.

В данной работе будет рассмотрено применение алгоритма доменной  
адаптации Progressive Unsupervised Learning для нескольких решений  
задачи распознавания ключевых точек на теле человека. Также будут  
представлены качественные и количественные результаты проведенного  
эксперимента.

## 2 Сверточные нейронные сети

На анный момент раздел находится под вопросом. Надо узнать насколько необходимо о СНС описывать тут. Если не будет хватать объема, то точно придется дополнять.



## 3 Распознавание ключевых точек

С развитием технологий человечество начало ставить все более разнообразные задачи, для решения которых применялись сверточные нейронные сети (convolutional neural network, CNN). Одной из таких задач оказалось распознавание ключевых точек (Keypoint Detection). При этом распознавание ключевых точек на теле человека выделилось в отдельный раздел, известный как оценка позы (Pose Estimation). Далее рассмотрим формулировку этих задач( , а также проведем обзор различных методов, которые на сегодняшний день применяются для их решения).

### 3.1 Ключевые точки

Задача распознавания ключевых точек заключается в том, чтобы обнаружить и точно локализовать определенные точки или места внутри изображения или кадра видео. Эти точки, называемые ключевыми, могут быть определены для различных объектов, таких как лица, тела человека или других структур. Например, в случае распознавания лица ключевые точки могут включать углы глаз, кончик носа, уголки рта и другие характерные особенности лица. В контексте человеческого тела ключевыми точками могут быть суставы, такие как локти, колени, плечи и так далее. Для других структур это могут быть уникальные элементы, которые помогают идентифицировать или анализировать объект.

Другими словами, ключевые точки являются опорными объектами, которые используются для определения положения и/или местонахождения объекта в пространстве. Они играют важную роль в задачах компьютерного зрения, таких как отслеживание движений, 3D-моделирование, анимация, медицинская визуализация и другие. Они могут использоваться для создания каркасов объектов, анализа их формы, измерения расстояний между различными частями и выполнения других аналитиче-

ских задач.

Более точно эти объекты можно определить следующим образом: *ключевые точки (КТ)* - это специфические, заранее определенные части объекта, которые имеют особое значение для дальнейшего анализа местоположения объекта на изображении. Каждая ключевая точка обычно соответствует определенной анатомической или структурной особенности, которая легко распознается и может служить ориентиром для алгоритмов обработки изображений. Иными словами, КТ необходимо обладать следующими характеристиками, чтобы можно было использовать их в качестве референсных для заданного объекта:

1. *Уникальность*

Точки должны быть уникальными и отличаться от других точек на изображении

2. *Инвариантность*

Точки должны сохранять свою идентичность при общих преобразованиях изображения, таких как вращение, масштабирование и изменения условий освещения

3. *Повторяемость*

Точки должны быть обнаруживаемы в разных экземплярах одного и того же объекта или сцены

Для анализа структуры объекта и взаимосвязи между его различными КТ часто применяется схематичное описание, которое обеспечивает более наглядное визуальное представление. Этот метод помогает лучше понять анатомическую и функциональную структуру распознаваемого предмета. При этом саму схему, представляющую собой своего рода «скелет», принято называть *топологией*. Примеры различных топологий представлены на рис. 1.



(a) Animal Keypoints



(b) Car keypoints

Рис. 1: Примеры топологий объектов от OpenPifPaf [1]

ВОЗМОЖНО ТУТ СТОИТ СФОРМУЛИРОВАТЬ ЗАДАЧУ РАСПОЗНАВАНИЯ КЛЮЧЕВЫХ ТОЧЕК. ОНА СФОРМУЛИРОВАНА, НО МОЖЕТ СМОГУ ДОБАВИТЬ МАТЕМАТИКИ.

### 3.2 Распознавание ключевых точек на теле человека. Скелет человека

На теле человека тоже можно выделить несколько ключевых точек, информация о которых дает возможность цифровизовать позу человека и использовать ее для аналитики с помощью методов машинного обучения. Именно для этого и была разработана задача распознавания ключевых точек на теле человека или, как ее часто называют в англоязычной литературе, задача оценки позы человека (Human Pose Estimation).

Основной вопросом для НРЕ стал выбор набора ключевых точек и топологии, по которой они будут соединяться. В первых работах были представлены различные примеры топологий, некоторые примеры представлены на рис. 2. Если точки туловища имеют большое количество пересечений, то точки головы сильно отличались: где-то учитывалось только положение головы (то есть добавлена верхняя точка головы), где-то

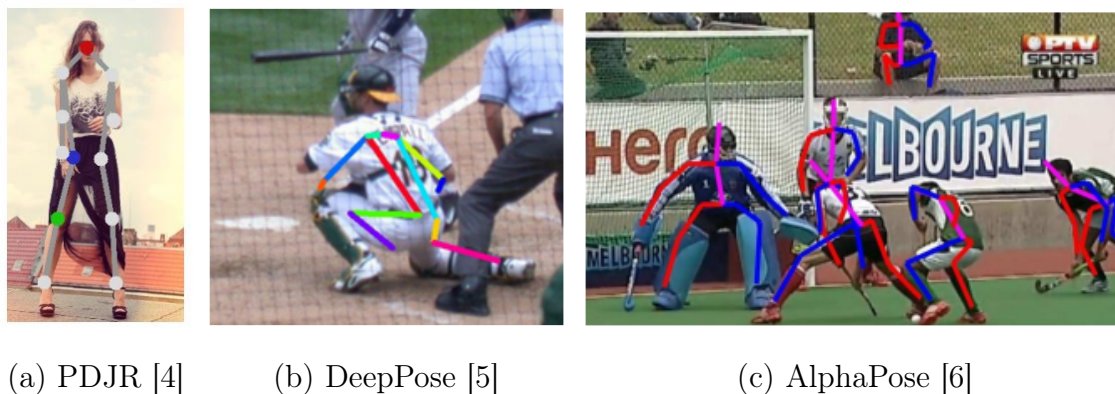


Рис. 2: Примеры различных топологий у первых решений задачи распознавания ключевых точек на теле человека

рассматривались некоторые точки на лице, где-то голова вообще не учитывалась. Все зависело от задачи и возможностей исследователей. Позже в 2015 году Microsoft выпустило набор данных с детальным описанием 17 точек на теле человека и запустила соревнование по распознаванию этих точек [2, 3]. Исследователей это заинтересовало и они начали адаптировать свои модели под топологию, описанную в датасете COCO [2]. Хотя данные в наборах перестали обновлять после 2017 года, многие новые модели до сих пор оценивают свои модели по набору данных COCO. Отсюда и получилось, что данная топология стала основной для задачи оценки позы.

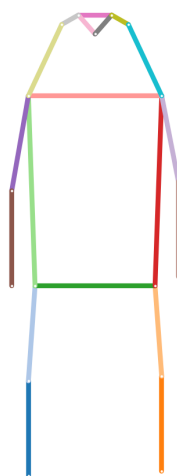


Рис. 3: Топология COCO

Другим вопросом для описываемой задачи стал подход к распознаванию. Делать детекцию человека и уже на кропнутом изображении производить поиск или искать все возможные точки, а потом собирать их в скелет. Эти две идеи и сформировали два направления развития методов НРЕ. Немного о них:

1. *Подход сверху-вниз (англ, top-down)*

Для данного этапа вам понадобится дополнительная модель детекции, которая локализует человека на изображении, выделяя его в прямоугольную область. Затем этот прямоугольник передается в модель распознавания ключевых точек. Этот подход обеспечивает высокую точность предсказания КТ, но может быть чувствителен к ошибкам на этапе обнаружения и ориентации человека в кадре, что требует дополнительной предобработки изображений.

2. *Подход внизу-вверх (англ, bottom-up)*

Для данного подхода вам не нужны помощники в виде детектора, но все равно имеет два этапа в своей работе. Первоначально модель распознает все ключевые точки на полученном на вход изображении, получая таким образом карту распределения КТ по фото. А вторым шагом модель собирает все полученные точки единый скелет. При детекции нескольких людей необходимо верно сопоставить их части тела. Для этого есть несколько способов, один из которых является построение полей сходства частей тела (part affinity fields), используемый в проекте OpenPose [7].

Но в результате обоих подходов к решению задачи на выход модели получаем массив данных  $[N \times K \times 3]$  в двумерном пространстве, где третьей размерностью идет предсказание видимости точки на изображении.

СТОИТ ДОБАВИТЬ ФОРМУЛ И МАТЕМАТИЧЕСКИЙ ЗАДАЧ.  
ПЕРЕДЕЛАТЬ ПОСЛЕДНИЙ АБЗАЦ

## 4 Domain Adaptation

В современном мире машинного обучения и искусственного интеллекта способность моделей адаптироваться к новым условиям и данным стала одной из ключевых задач. Традиционные методы обучения моделей предполагают, что данные, используемые для обучения и тестирования, имеют схожие характеристики и распределения. Однако в реальных приложениях часто возникает необходимость применять модели на данных, которые существенно отличаются от тех, на которых они были изначально обучены. Это приводит к снижению точности и эффективности моделей, что ставит под угрозу их практическое применение.

Для решения данной проблемы была разработана задача доменной адаптации, целью которой является создание методов для акклиматизации модели к целевым данным, отличающимся от исходных. В рамках этой задачи были разработаны подходы и техники, направленные на уменьшение расхождений между исходным и целевым доменами, что позволяет сохранять точность и производительность моделей в новых условиях. Эти методы способствуют переносу знаний, накопленных в исходном домене, на целевой, что относит данную задачу к области методов transfer learning.

### 4.1 Перенос обучения

ЗДЕСЬ ТОЧНО НАДО БУДЕТ ИСПОЛЬЗОВАТЬ КАРТИНКУ ДЛЯ  
ОПИСАНИЯ ЧТО ТАКОЕ ПЕРЕНОС ОБУЧЕНИЯ

СЛЕДУЮЩИЕ АБЗАЦЫ ДО МЕТКИ НЕОБХОДИМО ПРОВЕ-  
РИТЬ И ВОЗМОЖНО ПЕРЕНЕСТИ ВО ВВЕДЕНИЕ

Обучение модели и разработка алгоритма её работы представляют собой чрезвычайно сложный и трудоемкий процесс. Этот процесс требует значительных ресурсов, как со стороны специалистов, так и в плане

вычислительной мощности. Сначала необходимо собрать и подготовить данные, затем обучить модель, настроить её параметры и протестировать на различных наборах данных, чтобы убедиться в её точности и надёжности. Это часто занимает много времени и требует значительных финансовых вложений.

Столкнувшись с этими трудностями, исследователи начали искать способы повысить эффективность и упростить процесс обучения моделей. Они пришли к идее использования знаний, полученных при решении одной задачи, для решения другой, сходной задачи. Такой подход получил название *переноса обучения* (англ. *transfer learning*). Он позволяет существенно сократить время и ресурсы, необходимые для разработки новых моделей, путём повторного использования уже существующих знаний и наработок.

ВЫГЛЯДИТ КАК АБЗАЦ ДЛЯ ВВЕДЕНИЯ :)

Суть переноса обучения заключается в том, что модель, предварительно обученная на одном наборе данных, может быть адаптирована для работы на другом наборе данных, даже если они различаются по своему содержанию или структуре. Например, модель, обученная на огромном корпусе текстов, может быть успешно применена для анализа тональности текстов или для распознавания именованных сущностей в другом наборе данных. Точно так же, модели, обученные на больших наборах изображений, могут быть использованы для решения задач распознавания объектов или классификации изображений в других доменах.

Поэтому исследователи задались вопросом применения знаний, полученных в результате решения одной задачи, для решения другой задачи. Такой процесс получил название *переноса обучения* (англ. *transfer learning*).

МЕТКА ДО КОТОРОЙ НАДО ПРОВЕРЯТЬ АБЗАЦЫ



## 4.2 Доменная адаптация

В данной части необходимо рассказать как из второго появилось первое.

### 4.2.1 Классификация методов доменной адаптации

Есть ли деление в доменной адаптации. Какое, какие бывают виды. Провести обзор.

ЗДЕСЬ ТОЧНО НАДО БУДЕТ ИСПОЛЬЗОВАТЬ КАРТИНКУ  
КАК В ПРЕЗЕНТАЦИИ ДЛЯ КЛАССИФИКАЦИИ МЕТОДОВ ДА

## 5 Обзор существующих решений

В данной главе необходимо провести анализ существующих моделей. Аргументировать выбор тех или иных моделей, методов. Подготовить теоретическую базу перед экспериментов.

### 5.1 Обзор моделей для распознавания КТ

#### 5.1.1 DeepPose

Чисто для исторической справки. Если нужен будет объем.

#### 5.1.2 AlphaPose

Чисто для исторической справки. Если нужен будет объем.

#### 5.1.3 OpenPose

??? год. Проект от лаборатории ??? от института ???.

Показала хорошие результаты при предыдущих сравнениях. Использует

#### 5.1.4 BlazePose

Проект MediaPipe от гугл использует сеть архитектуры BlazePose.

Показывает хорошие результаты. Имеет хорошие

#### 5.1.5 HRNet

2019 год. Основная модель от проекта MMPose. Использует интересную архитектуру и хорошо показала себя при предыдущих сравнениях.

### 5.1.6 ViTPose

Трансформер. ??? год. Интересная архитектура. Рассказать про трансформер, рассказать про кодеки/докедеки. Рассказать про способ работы. Можно расписать неплохо.

### 5.1.7 RTMPose

2023 год. На момент исследования наиболее актуальная модель от проекта MMPose.

Работает неплохо. Обучается тоже неплохо. Попробуем заюзать в экспериментах.

ОБЗОР ДАЛЬНЕЙШИХ РЕШЕНИЙ ПОЗЫ ПОД БОЛЬШИМ ВОПРОСОМ.

### 5.1.8 Swin

Трансформер. 2021 год. Все )

### 5.1.9 Simcc + Resnet

Вообще хз что это, но если смогу, то сделаю эксперимент )

### 5.1.10 Dekr + HRNet

Подход снизу вверх. ??? год. Модель, обученная лабораторией показывает хорошие результаты, но очень прожорлива в плане ресурсов для обучения.

### **5.1.11 YoloPose**

Несколько различных версий архитектуры Yolo дают несколько различных версий модели для распознавания позы. Использует подход снизу-вверх и хорошо

## **5.2 Анализ и выбор методов доменной адаптации**

Опираясь на модели из предыдущей главы надо сделать анализ и выбрать только несколько методов, которые применимы к нашим требованиям и которые мы будем сравнивать между собой.

### **5.2.1 PUL**

Будем это использовать в работе. Надо будет хорошо расписать...

### **5.2.2 RegDA**

Интересный алгоритм. На нем базируются все остальные

ПРИ НЕХВАТКЕ ВРЕМЕНИ МОЖНО БУДЕТ ОСТАВИТЬ ТОЛЬКО ОДИН ИЗ СЛЕДУЮЩИХ МЕТОДОВ.

### **5.2.3 UDA PoseEstimation**

Адаптация от синтетических данных к реальным. Должно бтыть интересно описать. Не использовали, так как до сих пор мы не перешли в 3х мерное пространство.

### **5.2.4 POST**

Похожее на предыдущее

## **5.2.5 SFDA**

Похожее на предыдущее

## 6 Эксперимент

В данной главе будет описан эксперимент, как таковой.

### 6.1 Описание деталей эксперимента

Постановка эксперимента. Что планируется сделать и какие результаты хочется получить. Какие метрики будем использовать и по каким метрикам будем сравнивать.

### 6.2 Данные. Сбор и разметка

Описание собранных данных. Численный объем датасета. Возможно количество локаций и распределение по ним.

Необходимо предоставить данные по распределению данных между людьми, по количеству данных для обучения/тестирования. Предоставить данные по локациям. Предоставить картинки с примерами данных, которые были собраны.

Описание системы полуавтоматической разметки данных. Что там использовалось и как проходит разметка.

Для разметки собранных данных была создана система полуавтоматической разметки данных pose-markup (ССЫЛКА). Система состоит из двух частей: автоматическая разметка ключевых точек с помощью нейросети от MediaPipe и корректировка полученных данных экспертом.

На данный момент проект ориентирован на разметку позы только на видео файлах, так как это было необходимо реализовать в рамках эксперимента. В будущем планируется добавить возможность размечать точки и на одиночных изображениях.

Добавить коллаж примера работы системы (3 фото: исходное, после автоматической разметки, после корректировки. Необходимо явно ука-

зять точки, которые были изменены.)

Численные данные о разметке данных. Сколько пришлось разметить, сколько пришлось скорректировать. Предоставить некоторые картинки по распределению корректировки данных.

В рамках разметки данных можель имела неточности, поэтому роль эксперта была существенна при валидации работы автоматической части системы. Процент фотографий на которых были проведены изменения: собрать статистику и предоставить результаты корректировки разметки: сколько суммарно точек было изменено, сколько точек имели изменения более 5% от роста человека (метрика в бакалаврском дипломе есть), тоже самое для 15%, три случая для среднего количества точек на фотографии, три случая для фотографий (может большинство изменений были на трети фотографий, а остальные не меняли), представить сводную таблицу изменений для точек по топографии КОКО, визуализировать сводную таблицу на картинке.

### 6.3 Результаты эксперимента

Предоставить относительно сухо результаты эксперимента. Можно дать базовый анализ ситуации и того, что мы видим.

Необходимо предоставить результаты по времени обучения нейросетей, времени дообучения нейросетей.

Собрать данные по количеству ошибок до-после обучения. Собрать данные по количеству ошибок при изменении домена.

Данные по ресурсам, на которых обучались нейросетки.

## 7 Заключение

Необходимо рассказать о результатах исследования и о том, что мы получили. Привести сравнения, если таковые будут иметь место. Закинуть удочку для будущих исследований.



## Список литературы

- [1] *Kreiss, Sven*. OpenPifPaf: Composite Fields for Semantic Keypoint Detection and Spatio-Temporal Association / Sven Kreiss, Lorenzo Bertoni, Alexandre Alahi // *IEEE Transactions on Intelligent Transportation Systems*. — 2022. — Vol. 23, no. 8. — Pp. 13498–13511.
- [2] *Lin, Tsung-Yi*. Microsoft COCO: Common Objects in Context. — 2014. <https://arxiv.org/abs/1405.0312>.
- [3] *Tsung-Yi Lin Matteo Ruggero Ronchi, Alexander Kirillov*. COCO 2020 Keypoint Detection Task. — <https://cocodataset.org/#keypoints-2020>.
- [4] Human Pose Estimation Using Body Parts Dependent Joint Regressors / Matthias Dantone, Juergen Gall, Christian Leistner, Luc Van Gool // 2013 IEEE Conference on Computer Vision and Pattern Recognition. — 2013. — Pp. 3041–3048.
- [5] *Toshev, Alexander*. DeepPose: Human Pose Estimation via Deep Neural Networks / Alexander Toshev, Christian Szegedy // 2014 IEEE Conference on Computer Vision and Pattern Recognition. — 2014. — Pp. 1653–1660.
- [6] RMPE: Regional Multi-person Pose Estimation / Hao-Shu Fang, Shuqin Xie, Yu-Wing Tai, Cewu Lu // ICCV. — 2017.
- [7] OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields / Z. Cao, G. Hidalgo Martinez, T. Simon et al. // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. — 2019.
- [8] *Contributors, MMPose*. OpenMMLab Pose Estimation Toolbox and Benchmark. — <https://github.com/open-mmlab/mmpose>. — 2020.

- [9] Detailed, Accurate, Human Shape Estimation From Clothed 3D Scan Sequences / Chao Zhang, Sergi Pujades, Michael J. Black, Gerard Pons-Moll // The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). — 2017. — July.
- [10] Learning from Synthetic Humans / Gül Varol, Javier Romero, Xavier Martin et al. // CVPR. — 2017.
- [11] Deep 3D human pose estimation: A review / Jinbao Wang, Shujie Tan, Xiantong Zhen et al. // *Computer Vision and Image Understanding*. — 2021. — Vol. 210. — P. 103225. <https://www.sciencedirect.com/science/article/pii/S1077314221000692>.
- [12] *Tome, Denis*. Lifting from the Deep: Convolutional 3D Pose Estimation from a Single Image / Denis Tome, Chris Russell, Lourdes Agapito // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). — 2017. — Pp. 5689–5698.
- [13] *Whittle, Michael W.* Clinical gait analysis: A review / Michael W. Whittle // *Human Movement Science*. — 1996. — Vol. 15, no. 3. — Pp. 369–387. <https://www.sciencedirect.com/science/article/pii/0167945796000061>.
- [14] *google.github.io*. Pose Classification. — [https://google.github.io/mediapipe/solutions/pose\\_classification.html](https://google.github.io/mediapipe/solutions/pose_classification.html).
- [15] OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields / Z. Cao, G. Hidalgo Martinez, T. Simon et al. // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. — 2019.
- [16] *Kocabas, Muhammed*. MultiPoseNet: Fast Multi-Person Pose Estimation using Pose Residual Network. — 2018. <https://arxiv.org/abs/1807.04067>.

- [17] Human3.6M: Large Scale Datasets and Predictive Methods for 3D Human Sensing in Natural Environments / Catalin Ionescu, Dragos Papava, Vlad Olaru, Cristian Sminchisescu // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. — 2014. — jul. — Vol. 36, no. 7. — Pp. 1325–1339.
- [18] Prediction of Human Activities Based on a New Structure of Skeleton Features and Deep Learning Model / Neziha Jaouedi, Francisco J. Perales, José Maria Buades et al. // *Sensors*. — 2020. — Vol. 20, no. 17. <https://www.mdpi.com/1424-8220/20/17/4944>.
- [19] PaStaNet: Toward Human Activity Knowledge Engine / Yong-Lu Li, Liang Xu, Xinpeng Liu et al. // *CVPR*. — 2020.
- [20] *Bazarevsky, Valentin*. BlazePose: On-device Real-time Body Pose tracking. — 2020. <https://arxiv.org/abs/2006.10204>.
- [21] *Bazarevsky, Valentin*. BlazeFace: Sub-millisecond Neural Face Detection on Mobile GPUs. — 2019. <https://arxiv.org/abs/1907.05047>.
- [22] *Zhang, Fan*. MediaPipe Hands: On-device Real-time Hand Tracking. — 2020. <https://arxiv.org/abs/2006.10214>.
- [23] *Ivan Grishchenko, Valentin Bazarevsky*. MediaPipe Holistic — Simultaneous Face, Hand and Pose Prediction, on Device. — <https://ai.googleblog.com/2020/12/mediapipe-holistic-simultaneous-face.html>.
- [24] *google.github.io*. MediaPipe.Home. — <https://google.github.io/mediapipe/>.
- [25] *Sandler, Mark*. MobileNetV2: Inverted Residuals and Linear Bottlenecks. — 2019.

- [26] 2D Human Pose Estimation: New Benchmark and State of the Art Analysis / Mykhaylo Andriluka, Leonid Pishchulin, Peter Gehler, Bernt Schiele // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). — 2014. — June.
- [27] *Johnson, Sam*. Clustered Pose and Nonlinear Appearance Models for Human Pose Estimation / Sam Johnson, Mark Everingham // Proceedings of the British Machine Vision Conference. — 2010. — doi:10.5244/C.24.12.
- [28] Hand Keypoint Detection in Single Images using Multiview Bootstrapping / Tomas Simon, Hanbyul Joo, Iain Matthews, Yaser Sheikh // CVPR. — 2017.
- [29] Convolutional pose machines / Shih-En Wei, Varun Ramakrishna, Takeo Kanade, Yaser Sheikh // CVPR. — 2016.