**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ**
**ІМЕНІ ТАРАСА ШЕВЧЕНКА**
Факультет комп'ютерних наук та кібернетики
Кафедра математичної інформатики

«До захисту допущено»
Завідувач кафедри
В. М. Терещенко _____
(підпис)

«___» _____20__ р.

Кваліфікаційна робота
на здобуття ступеня бакалавра
за освітньо-професійною програмою "Інформатика"
спеціальності 122 "Комп'ютерні науки"
на тему:

# Розробка системи комп'ютерного моделювання та аналізу рухів людини у центрах реабілітації

Виконав студент 4 курсу
Щербина Павло Андрійович _____
(підпис)

Науковий керівник:
доктор фіз.-мат. наук, професор
Терещенко Василь Миколайович _____
(підпис)

Засвідчую, що в цій дипломній роботі немає запозичень з праць інших авторів без відповідних посилань.

Студент _____
(підпис)

Київ – 2022

# РЕФЕРАТ

Робота складається зі вступу, 3 розділів, висновку, списку використаних джерел (14 найменування). Робота містить 12 рисунків. Загальний обсяг становить 38 сторінок.

**Ключові слова:** КОМП'ЮТЕРНИЙ ЗІР, ФІЗИЧНА РЕАБІЛІТАЦІЯ, РОЗПІЗНАВАННЯ ОБРАЗІВ, ПОБУДОВА СКЕЛЕТОНУ, КЛЮЧОВІ ТОЧКИ, АНАЛІЗ РУХІВ ЛЮДИНИ, ШТУЧНИЙ ІНТЕЛЕКТ, ОЦІНКА ПОЗИ, ОБРОБКА ДАНИХ.

**Об'єкт роботи:** система комп'ютерного моделювання та аналізу рухів людини для реабілітації, яка корегує виконання фізичних вправ, виконуваних пацієнтами.

**Мета роботи:** ознайомлення аудиторії з особливостями побудови систем реабілітації, які стають можливими з швидким розвитком технологій комп'ютерного зору та штучного інтелекту, реалізація різних методів аналізу коректності виконання фізичних вправ.

**Методи дослідження:** аналіз літератури та існуючих підходів до вирішення подібних задач; тестування та визначення найбільш ефективних алгоритмів, створення власних програм, що їх реалізують;

**Інструменти**: безкоштовне середовище розробки та мова програмування Python, безкоштовні бібліотеки обробки даних та застосування алгоритмів машинного навчання NumPy, Pandas, MatPlotLib, scikit-learn, безкоштовна бібліотека для оцінки пози та побудови скелетона людини OpenPose, інструменти для роботи з ОС.

**Результати роботи**: ми проаналізували сучасні підходи до побудови систем реабілітації та обрали свій стек технологій.

Вдалося розробити свою систему аналіза виконання вправ пацієнта з генеруванням відгуку про коректність його рухів для можливого покращення

виконання чи уникнення травм. Аналіз точності показав, що система працює добре, тому ми запропонували можливі шляхи розвитку цієї системи.

# ЗМІСТ

**РЕФЕРАТ** **2**

**ЗМІСТ** **4**

**СКОРОЧЕННЯ ТА УМОВНІ ПОЗНАЧЕННЯ** **6**

**ВСТУП** **7**

**РОЗДІЛ 1. ОГЛЯД ІСНУЮЧИХ РІШЕНЬ** **9**

1.1 Реабілітація 9

1.1.1 Дані про фізичні порушення 9

1.1.2 Приватність пацієнтів 10

1.1.3 Оцінка 10

1.2 Загальний логічний процес для систем оцінки пацієнтів на основі комп’ютерного зору 11

1.2.1 Сенсори та камери 11

1.2.2 Системи побудови скелетонів 14

1.2.3 Математичне представлення скелетонів 15

1.2.4 Порівняння ознак 15

**РОЗДІЛ 2. ПРАКТИЧНА РЕАЛІЗАЦІЯ** **17**

2.1 Постановка задачі 17

2.1.1 Опис системи 17

2.1.2 Підготовка даних 18

2.2 Технічний підхід 19

2.2.1 Огляд робочого процесу системи 19

2.2.2 Запис відео 19

2.2.3 Оцінка пози 20

2.2.4 Нормалізація опорних точок 21

2.2.5 Визначення перспективи 22

2.2.6 Геометричні оцінки 22

2.2.7 Оцінка за допомогою методів машинного навчання 24

**РОЗДІЛ 3. РЕЗУЛЬТАТИ** **26**

3.1 Вправи 26

3.1.1 Скручування на біцепс 26

3.1.2 Піднімання гантелей спереду 28

3.1.3 Знизування плечами з гантелями 29

3.1.4 Жим плечей стоячі 31

3.2 Точність 33

3.3 Наступні етапи розвитку системи 34

**ВИСНОВКИ 35**

**ДОДАТОК А 36**

**СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ 37**

# СКОРОЧЕННЯ ТА УМОВНІ ПОЗНАЧЕННЯ

CNN — скорочення convolutional neural network;

CV — скорочення computer vision;

NUI — скорочення natural user interface;

DTW — скорочення dynamic time warping;

ANOVA — скорочення analysis of variance;

FMA — скорочення Fugl-Meyer Assessment;

DL — скорочення deep learning;

# ВСТУП

**Оцінка сучасного стану об’єкта розробки.**

З кожним днем кількість відео та фото даних стрімко зростає, а цифрові камери уже стали звичними речами для кожного з нас. Для обробки цих зображень та отримання корисної інформації розробляються технології комп’ютерного зору (CV), які знаходять застосування у різних сферах нашого життя. Однією з таких сфер є аналіз людських рухів.

Моделювання та аналіз людських рухів за допомогою технологій комп’ютерного зору широко досліджується спільнотою. Все ж більшість робіт присвячені таким темам, як оцінка пози [1], взаємодія людини з об’єктом, розпізнавання активності/жестів [2] або взаємодію людини з людиною [3]. Проте порівняльний аналіз рухів людини отримував відносно менше уваги з боку спільноти через широкий діапазон аномалій рухів людини та труднощі з автоматичною оцінкою цих відхилень. Варто відмітити, що покращення алгоритмів оцінки пози відкривають нові можливості в порівняльному аналізі.

**Актуальність роботи та підстави для її виконання.**

Порівняльний аналіз рухів [4] людини необхідний в таких сферах застосування, як перевірка правильності виконання фізичних вправ, автоматизована реабілітація, оцінка інсульту, травми спинного мозку, хвороби Паркінсона та інші фізичні вади. Пацієнтам, які відновлюються після порушень, проходять широку фізичну реабілітацію та оцінюються клініцистами (лікарями, фізіотерапевтами або ерготерапевтами), також важливо весь час тренувань проводити під наглядом. Неправильне виконання фізичних вправ, яке може бути помічено тільки наглядачем або комп’ютером зі сторони, може призвести до регресу в реабілітації або до ще більш серйозних травм. Тож процес реабілітації є дорогий, трудомісткий, вимагає багато часу і схильний до людської помилки. Статистика показує що неформальна допомога може зберегти до 27% ціни всього лікування.

Також оцінки клініцистів можуть страждати від неточностей, оскільки візуальна схема звітування про процес реабілітації схильна до непослідовного сприйняття. Неточності також можуть виникнути через суб'єктивність клінічних оцінок. Крім того, інтеграція оцінки на основі кінематичних параметрів може бути більш надійною та точною в порівнянні з візуальною оцінкою лише клініцистами.

Існують різні способи зчитування рухів людини. Датчики, що носятся на тілі, або системи на основі маркерів є дорогими і можуть бути дуже нав'язливими для повсякденної діяльності пацієнта. Моделювання руху людини на основі комп'ютерного зору без маркерів і подальше порівняння може забезпечити навіть домашній, недорогий і ненав'язливий моніторинг. Він також має потенційне застосування в спорті.

Найважче в системах для реабілітації це не розпізнавання виконання фізичної вправи, а  реалізація логіки коригування нашої пози у випадку неправильного виконання.

# РОЗДІЛ 1. ОГЛЯД ІСНУЮЧИХ РІШЕНЬ

## 1.1 Реабілітація

У реабілітаційних системах головною метою є забезпечення автоматизованого будинку або клініки, щоб пацієнти проходили фізичну терапію, жестотерапію або інші реабілітаційні вправи. Така система спрямовує пацієнтів виконувати свої реабілітаційні завдання. Реабілітація може бути повністю автоматизована або опосередкована лікарем. Дослідження в цій категорії зазвичай спрямовані на поліпшення фізичного стану пацієнта. Більшість досліджень відноситься до типу віртуальної реабілітації. У віртуальній реабілітації оцінюється ефективність пацієнта у віртуальному світі, а не безпосередньо оцінка фізичної працездатності пацієнта. Це і аватар, що виконує завдання у віртуальному світі, і використання ігор для реабілітації. Тут суб'єкти повинні виконувати дії у віртуальному світі через рухи в реальному. У системах Direct Rehabilitation користувачі керуються веб-інтерфейсом для виконання реабілітаційних вправ, а їхні рухи безпосередньо відстежуються за допомогою датчика зору. У цьому випадку вимірюється фізична працездатність пацієнта, а не його аватар чи здатність виконувати завдання у віртуальному світі. Оцінка пацієнта може бути вбудованою або може вимагати клініцистів.

### 1.1.1 Дані про фізичні порушення

Рухи людини багатовимірні, як і їх аномалії. Порушення опорно-рухового апарату проявляються по-різному у різних пацієнтів протягом певного періоду часу. Безліч факторів, таких як порушення, ступінь ушкодження, зона ураження, фізіологічні характеристики та наданий догляд, призводять до надзвичайно різноманітних проявів у пацієнтів. Це на додаток до широкого діапазону рухових можливостей людини. Клініцисти мають спеціальні тести та вправи, призначені для реабілітації та оцінки різних типів рухових аномалій. Тому дослідники також повинні провести специфічний експеримент, щоб отримати дані для оцінки конкретних м'язово-скелетних порушень. Через труднощі з

доступом до пацієнтів, етичні проблеми та інші подібні проблеми важко отримати дані, а набори даних часто невеликі.

### 1.1.2 Приватність пацієнтів

Слід зазначити, що якщо розробляється система для реабілітації у медичних закладах, розробники мають потурбуватися про приватність пацієнтів та медичну таємницю. Системи з використанням комп'ютерного бачення є особливо небезпечні зі сторони конфіденційності даних, так як вони безпосередньо записують та обробляють відеодані. Якщо наша система не повинна працювати з аналізом емоційного стану пацієнта шляхом обробки зображень обличчя, то вихідом буде додавання алгоритму розпізнавання обличчя та наступного його розмиття перед збереженням відеозапису до бази даних.

### 1.1.3 Оцінка

У додатках для оцінки, мета полягає в тому, щоб на момент часу оцінити якість рухів пацієнта, пов'язану з однією або кількома частинами тіла. Оцінка може проводитися в клінічних або доклінічних умовах. Додаток для оцінки можна розділити на три типи залежно від того, як користувач отримує кінцевий результат. Перший тип – це порівняння , де дані пацієнта (наприклад, кінематичні параметри) добуваються для порівняння, але не існує вирішальної автоматизованої системи оцінки. У таких додатках може бути статистичне порівняння, як-от дисперсійний аналіз (ANOVA [5]) або просте графічне порівняння кінематики, представленої траєкторіями ідеалу та кута або положення суглоба пацієнта в часі. По-друге, маємо системи типу категоризації , де головною метою є класифікація. Рухи можуть бути класифіковані як правильні-неправильні або можуть бути класифіковані на кілька типів відхилень. Це включає в себе розпізнавання жестів/пози та діяльності. У третьому застосунку типу підрахунку балів, вирішальний бал прикріплюється до рухів пацієнта для оцінки якості їх руху.

## 1.2 Загальний логічний процес для систем оцінки пацієнтів на основі комп'ютерного зору

Існує багато досліджень на основі комп'ютерного бачення, включаючи вихідні дані, вилучення ознак, представлення ознак, порівняння ознак, статистичне та стохастичне моделювання (DL). Однак загальний потік досліджень у сфері реабілітації та моніторингу на основі зору можна широко проілюструвати на рис.1.

**Рисунок 1.** логіка застосунків для реабілітації

Ілюстрація включає в себе датчик зору, такий як монокуляр RGB або камера глибини для визначення даних, функцію низького рівня, наприклад положення суглобів людини, метод кодування та представлення ознак, наприклад група позицій суглобів або комбінація кінематичних параметрів людини. Потім закодовані ознаки порівнюються за допомогою простих графічних і статистичних методів або за допомогою інтелектуальних алгоритмів. Нарешті, оцінка проводиться у формі порівняння кінематичних параметрів, розпізнавання пози, автоматизованого клінічного оцінювання, класифікації порушень тощо. Системи реабілітації зазвичай мають програму вправ і забезпечують зворотний зв'язок. Ці характеристики можна широко описати у трьох основних частинах: первинні дані, виділення та представлення ознак, а також порівняння ознак.

### 1.2.1 Сенсори та камери

Існують різні датчики зору (рис. 2) для роботи з комп'ютерним зором. Доступні прилади зору, які підтримують 2D або 3D скелетного пошуку може

мати три типи: система Mocap, глибина камера та камера RGB. Такі системи пропонують компанії OptiTrack, Qualisys та VICON. Для таких сфер, як біомеханіка, спорт, інженерія та розваги. Ці Mocap системи можуть надавати дуже точні скелетні дані, але мають високу ціну та низьку гнучкість з метою комерціалізації.

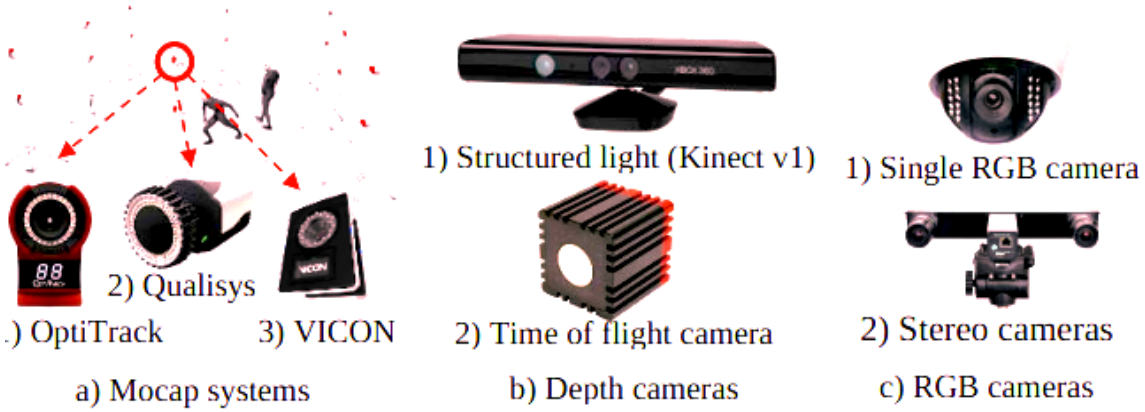


**Рисунок 2.** Різни типи сенсорів та камер

Камери глибини базуються на таких технологіях, як датчик структурованого світла. Деякі стандартні комерційні камери глибини, як-от Kinect і Intel RealSense можуть отримувати скелетні дані значно доступнішим способом.

Kinect — це надбудова Microsoft для датчиків руху для ігрової консолі Xbox 360. Пристрій забезпечує природний інтерфейс користувача (NUI), який дозволяє користувачам взаємодіяти інтуїтивно та без будь-яких проміжних пристроїв, таких як контролер.

Система Kinect ідентифікує окремих гравців за допомогою розпізнавання облич і голосу. Камера глибини, яка «бачить» у 3-D, створює зображення скелета гравця, а датчик руху фіксує його рухи. Програмне забезпечення для розпізнавання мовлення дозволяє системі розуміти голосові команди, а розпізнавання жестів дозволяє відстежувати рухи гравця.

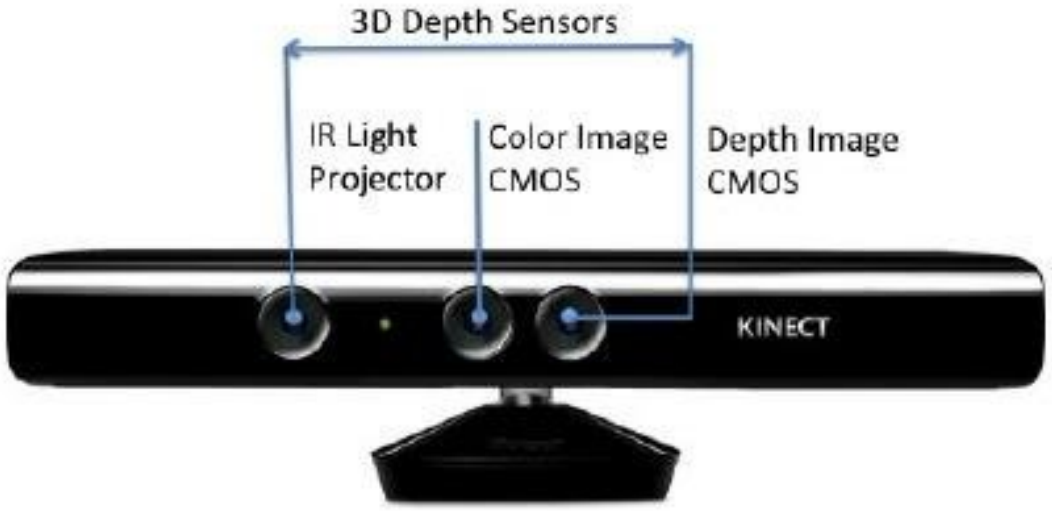


**Рисунок 3.** Microsoft Kinect камера

Незважаючи на те, що Kinect було розроблено для ігор, ця технологія була застосована до таких різноманітних реальних додатків, як цифрові вивіски, віртуальні покупки, освіта, надання послуг дистанційного медичного обслуговування та інші сфери медичних IT.

Камери RGB включають монокулярну камеру. Найдешевші RGB камери також можуть застосовуватися для побудови 2D або 3D скелетону, що вимагає сильніших обчислювальних потужностей і важко працювати з ними в системах реального часу.

### 1.2.2 Системи побудови скелетонів

OpenPose [6] використовує VGGNet-19 як екстрактор функцій під час навчання моделі мережі. Однією з найбільших особливостей VGGNet є те, що він може обробляти вхідні дані зображення різної роздільної здатності. Правильна обробка даних зображення, що вводяться в мережу, може покращити плавність роботи системи. OpenPose це система реального часу, яка здатна фіксувати 2D положення декількох людей на фото чи відео . Саме виявлення положення тіла, реалізовано підходом знизу вгору: спочатку алгоритм знаходить окремі частини тіла, які потім можна асоціювати з певними індивідами на фото. Сама система дуже гнучка і може бути налаштована під вашу задачу. Поточні популярні бібліотеки розпізнавання поз людини включають OpenPose, Mask R-CNN [7] і Alpha-Pose [8]. Час розпізнавання людського тіла порівнюється серед трьох, і було виявлено, що зі збільшенням кількості людей на зображенні час роботи Mask R-CNN і Alpha-Pose демонструє лінійне зростання, а час роботи OpenPose залишається незмінним. , тому система, розроблена за допомогою OpenPose, буде більш стабільною.

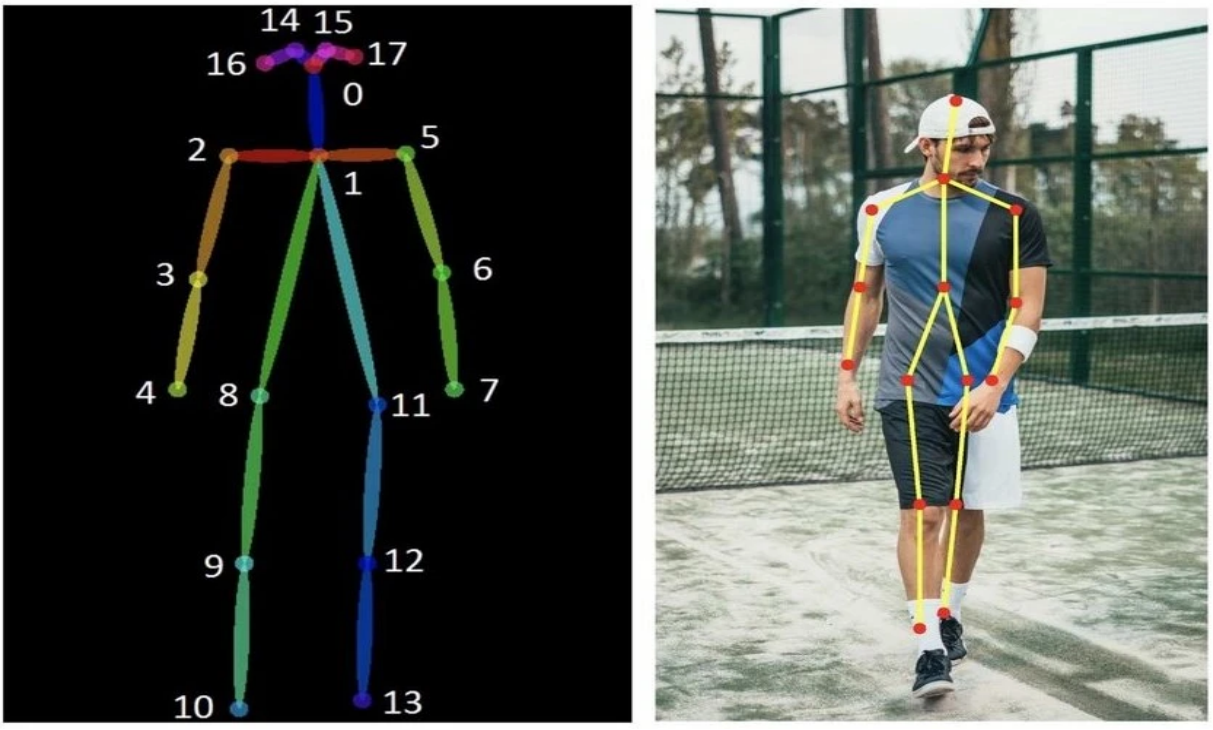


**Рисунок 4.** Результат роботи OpenPose

### 1.2.3 Математичне представлення скелетонів

Кожен скелетон може складатися з 25 опорних точок, які можемо позначити як голова, шия, ліве коліно тощо. Для набору суглобів у каркасі тіла, якій спостерігається в момент часу t можемо його представити як набір $j^t = (j_1^t, \dots, j_i^t, \dots, j_{25}^t)$ де $j_i^t$ має 8 атрибутів, які відповідають за позицію і орієнтацію опорної точки в просторі. Позиція опорної точки $i$ має 4 атрибути: 3D координати та його висота від підлоги, тоді $j_{i-pos}^t = (j_x^t, j_y^t, j_z^t, j_h^t)$. Орієнтація опорної точки $i$ в просторі визначається як кватерніон з набором значень $X, Y, Z$ і $W$. [ 9 ]

### 1.2.4 Порівняння ознак

Однією з головних цілей досліджень фізичної реабілітації та моніторингу на основі зору є забезпечення автоматизованої клінічної оцінки фізичного стану м’язово-скелетного пацієнта. Для багатьох задач CV, таких як виявлення об’єктів або розпізнавання діяльності, ціль чітко визначена (наприклад, класифікація). Однак для оцінки пацієнтів ціль сильно змінюється і часто залежить від клінічних вимог. Вимоги варіюються від статистичного аналізу до методів автоматичного встановлення клінічних балів, таких як оцінка Фугла-Мейєра (FMA)[10], уніфікована шкала оцінки та інші. Для деяких випадків достатньо простого представлення та порівняння траєкторій кута суглоба, але для інших випадків, таких як автоматизована клінічна оцінка, часто потрібні розширені алгоритми порівняння.

Важливо розробити методику порівняльного аналізу ознак, засновану лише на позиціях суглобів. Такі порівняння можна виконувати багатьма способами, включаючи простий графічний аналіз, статистичний аналіз, порівняння послідовності, класифікацію та регресію. Такі методи, як графічне порівняння, часто прості і можуть не вимагати великих наборів даних. З іншого боку, для створення автоматизованої клінічної оцінки потрібні розширені алгоритми та великі набори даних для надійної роботи. Як пояснювалося

раніше, отримати великомасштабний набір даних для кожного типу аномального руху досить важко. Тому основним завданням у цій сфері є максимізувати застосовність передових алгоритмів з обмеженими даними.

# РОЗДІЛ 2. ПРАКТИЧНА РЕАЛІЗАЦІЯ

## 2.1 Постановка задачі

Як ми знаємо з попереднього розділу фізична реабілітація це досить обширна тема з різними складними задачами. Комп'ютерний зір можна застосувати для різних систем. Деякі підраховують кількість повторень рухів разом зі швидкістю виконання, інші ж аналізують саме виконання вправи і дають коригувальну оцінку в реальному часі.

Зробити систему яка б мала в своєму інструментарії усі технології для аналізу, збирала інформацію про реабілітаційну динаміку пацієнта, а також працювала з ним в реальному часі для коригування виконання вправ є досить важка задача. Вона потребує багато часу на розробку, та досить сучасних і потужних комп'ютерних потужностей.

В практичній частині нашої роботи ми розробили систему, яка визначає позу пацієнта і генерує відгук про правильність виконання тієї чи іншої вправи, використовуючи останні досягнення в технологіях оцінки пози та машинного навчання. Мета – запобігти травмам та покращити якість тренувань пацієнтів лише за допомогою комп'ютера та вебкамери.

### 2.1.1 Опис системи

Перша частина нашої системи оцінює позу людини. Pose estimation складна, але дуже важлива задача комп'ютерного зору. З отриманих візуальних даних, які можуть бути RGB зображення, навчена модель прогнозує скелетон опорних точок пацієнта. Оцінка пози є одним з найважливіших етапів для задач, пов'язаних з розпізнаванням вправ та рухів людини. Ми використовуємо сучасну глибоку нейронну мережу для оцінки пози OpenPose, про яку було згадано в попередньому розділі.

Друга частина програми аналізує наскільки правильна є поза користувача, при виконанні тієї чи іншої вправи. Для цього ми застосували евристичний

підхід аналізу рухів та модель машинного навчання, використовуючи пози та інструкції персональних тренерів та інших фахівців як еталонні виконання вправ.

### 2.1.2 Підготовка даних

Для оцінки пози ми використовуємо попередньо навчену систему реального часу OpenPose, яка може виявляти опорні точки тіла людини на відео.

Ця система працює без додаткових налаштувань, її досить легко встановити. Використання OpenPose дозволяє скористатися перевагами сучасних алгоритмів оцінки пози для нашого завдання, і зосередитися на фактичній оцінці виконання вправ.

Для частини яка аналізує виконання вправи, ми записали деякі відео, як ми тренуємося, а також частину взяли з вільного доступу. Ми намагалися виконати вправи коректно, а також навмисно створили неправильні приклади. Оцінка ідентифікатора пози відбувається різними способами в залежності від алгоритму: для евристичних алгоритмів ми надаємо всі відео для оцінки, а для алгоритмів машинного навчання, ми оцінюємо, розділяючи наш відео датасет на тренувальні та тестувальні набори.

## 2.2 Технічний підхід

### 2.2.1 Огляд робочого процесу системи

Зараз ми розглянемо з технічної точки зору робочий процес системи, якій складається з декількох етапів (Рис. 4).

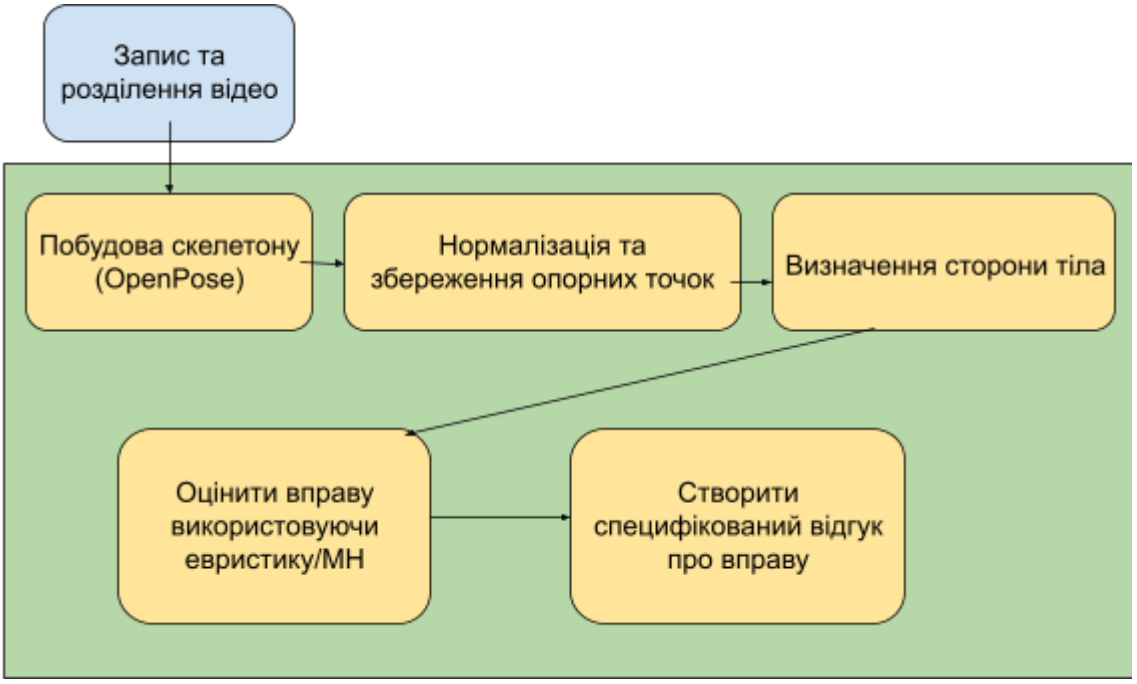


**Рисунок 5.** Процес роботи програми

Наша програма починається з запису відео виконання вправи, а закінчується, специфікованим відгуком користувачу про його виконання.

### 2.2.2 Запис відео

Спочатку, пацієнт чи користувач записує відео, як він виконує обрану вправу. Це відео записано з певної перспективи (обличчям до камери, боком до камери, тощо) це дозволяє побачити виконання вправи. Немає рекомендацій до типу камери або ж дистанції до неї, користувач тільки має впевнитись, що він повністю попадає в кадр. Потім він обрізає відео, щоб на ньому залишилося тільки виконання вправи.

Ми залишаємо запис та обрізання відео для користувача, усі сучасні смартфони мають досить зручні та швидкі вбудовані відеоредактори. Наша система працює з усіма найбільш поширеними форматами відеоданих.

### 2.2.3 Оцінка пози

Для оцінки пози ми використовуємо глибокі згорткові нейронні мережі (CNN) для розпізнавання скелетонів людини на RGB зображеннях. Після експериментів із кількома найсучаснішими системами для оцінки пози людини, ми обрали попередньо навчену модель OpenPose. Ця система представляє новий підхід для оцінки пози з використанням полів спорідненості частини, які є векторами положення та орієнтацію кінцівок у просторі. OpenPose є точним і ефективним, а також масштабується до кількох людей без зростання часу на виконання.

Ще одним важливим фактором у виборі OpenPose є його простота у встановленні та використанні для наших кінцевих користувачів. Більшість схожих систем зараз існують як Tensorflow або Caffe вихідний код, що містить архітектуру моделі, і вимагає складної установки від користувача. Крім того, бібліотеки GPU, такі як CUDA і CuDNN також потрібно встановити, щоб ці оцінювачі працювали коректно. Це далеко за межами досяжності більшості пересічних користувачів. До того ж, OpenPose може працювати як на Windows так і на Linux. Крім того, немає зовнішньої установки бібліотек GPU. OpenPose автоматично використовуватиме NVIDIA, GPU всередині комп’ютера, якщо є.

Вихідні дані OpenPose складаються зі списків, що містять прогнози координат усіх місцезнаходження опорних точок людини та відповідну впевненість прогнозу. Ми розглядаємо прогнози 18 ключових точок пози, до яких відноситься ніс, шия, плечі, лікті, зап'ястя, стегна, коліна та щиколотки.

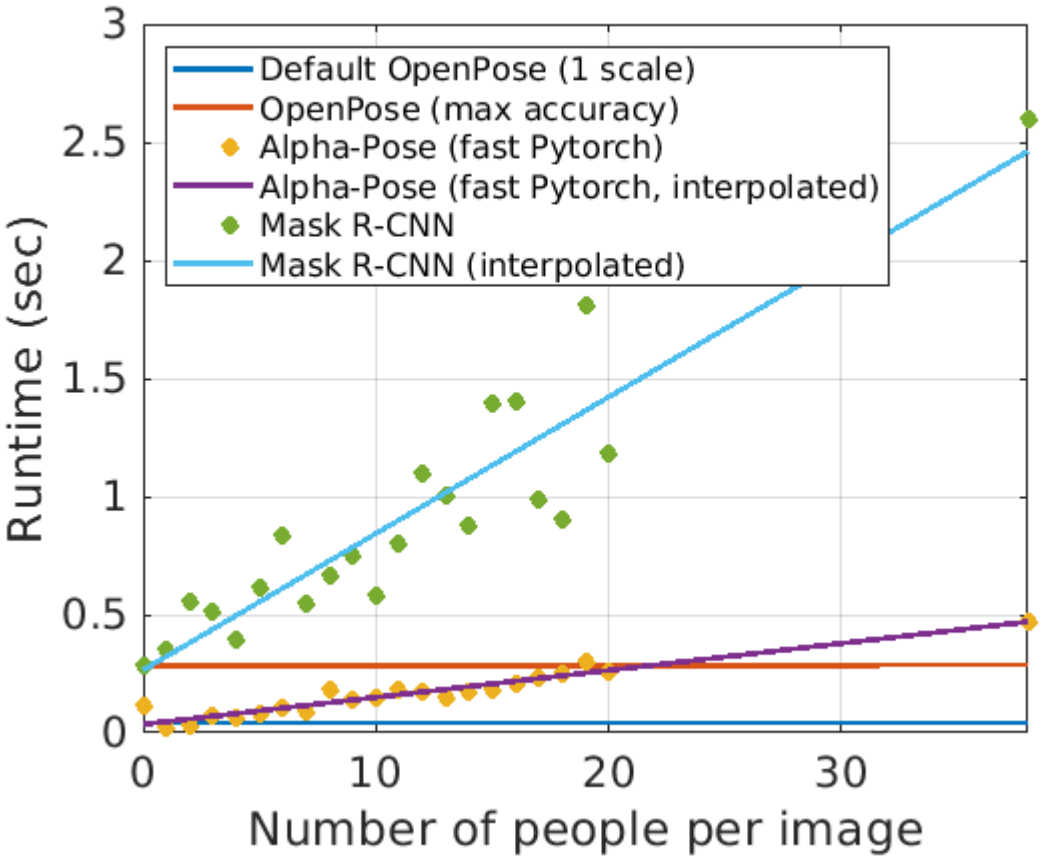


**Рисунок 6.** Демонстрація залежності швидкості роботи та кількості людей на фото для популярних бібліотек визначення опорних точок

### 2.2.4 Нормалізація опорних точок

Після виконання оцінки пози ми використовуємо власний парсер, щоб використовувати вихідні дані опорних точки. Спочатку зчитуємо списки опорних точок для кожного відео та сегментуємо їх на об’єкти Part(частина тіла), які зберігають x, y та впевненість прогнозу. Крім того, ми відстежуємо, чи є прихована опорна точка у кадрі (впевненість прогнозу 0). Для кожного кадру у відео ми складаємо кожну частину тіла, щоб побудувати об’єкт Pose, який представляє повний скелетон людини. Для повного відео ми складаємо скелетони у об'єкт PoseSequence, який зв'язує оцінки Pose для кожного кадру у відео.

Ми відзначаємо необхідність узагальнення нашої програми для обліку для користувачів з різними вимірами довжини тіла, відстань від камери, а також інші відносні фактори. Щоб врахувати ці відмінності, ми нормалізуємо позу на

основі довжина тулуба в пікселях. Розраховується довжина тулуба за середнім значенням відстані від ключової точки шиї до ключові точки правого та лівого стегна. Ця нормалізація працює добре: ми спостерігаємо, що довжина тулуба залишається постійною на всіх кадрах вхідних відео. Відстані таким чином представлені як співвідношення довжини тулуба: наприклад, довжина передпліччя 0,6 означає, що передпліччя дорівнює 0,6 довжина тулуба.

### 2.2.5 Визначення перспективи

Для з деякими фізичними вправами, ми визначаємо неоднозначність у перспективі з якої знімає камера. Наприклад, вправа згинання руки на біцепс записується з боку тіла і може бути виконана лівою або правою рукою. Для визначення, яка саме рука виконує вправу, ми обираємо опорні точки які є найбільш помітні на кадрах (ліворуч або праворуч). Це точно визначає перспективу на всіх вхідних даних.

### 2.2.6 Геометричні оцінки

Наш евристичний підхід полягає в тому, що ми оцінюємо векторні дані опорних точок тіла, які є ключовими для виконання тієї чи іншої вправи. Використовуємо персональні вказівки до тренування та записані відео для проектування геометричної евристики.

До прикладу розглянемо вправу скручування на біцепс. (Рисунок 6)

По-перше, передпліччя слід тримати нерухомим. Ми можемо це визначити порахувавши кут між вектором плеча та вектором тулуба, плече має бути паралельним тулубу, або з мінімальними відхиленнями. Значний рух призведе до зміни кути між двома векторами.

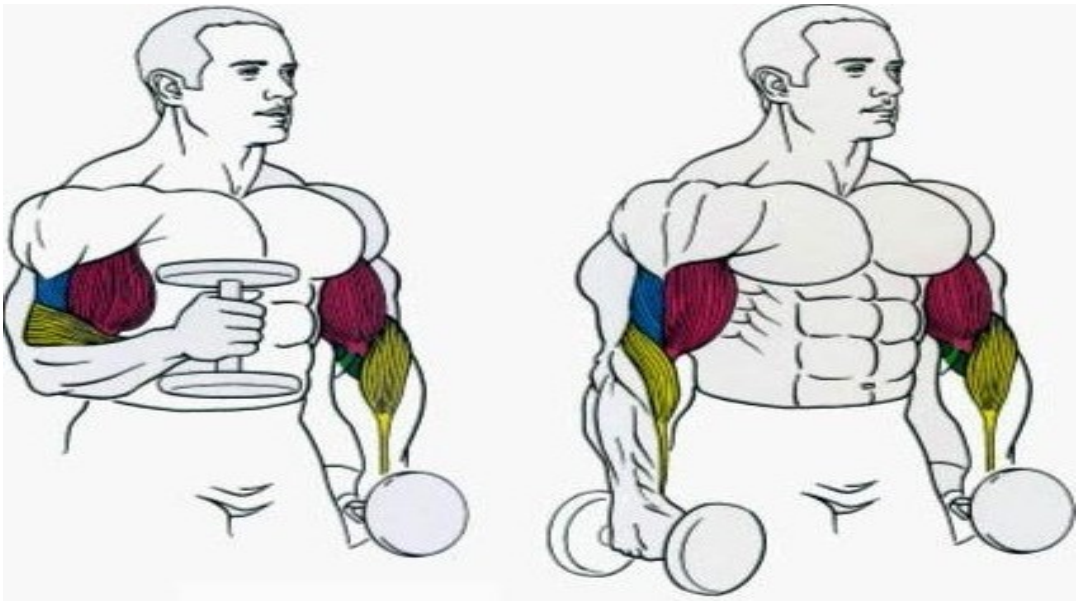

**Рисунок 7.**. Скручування на біцепс

По-друге, ця вправа вимагає виконання з додатковою вагою. Вправа виконується до моменту поки наш м’яз не скоротиться. Серединна точка є 90 між плечем та передпліччям. Якщо користувач виконує вправу до цієї точки це означає, що гантелі для нього є занадто важкими, про це йому повідомить програма. Іншими словами ми оцінюємо виконання вправи розміром мінімального кута, досягнутого між передпліччям та плечем. У вихідному положенні з опущеною вагою кут повинен бути близько 180 . З підняттям ваги кут повинен зменшуватися до моменту, коли користувач зупиниться. Таким чином, пошук мінімального кута протягом усього відеоряду покаже, наскільки користувач підняв вагу.

Проаналізувавши відеодані, ми знаходимо що якщо кут між векторами плеча та тулуба зростають до 35 , це означає що користувач надмірну рухає плече під час виконання вправи. Якщо мінімальний кут між плечем та передпліччям не нижче 70 , це говорить про те що користувач не піднімає вагу до кінця.

Використовуючи кількісні показники та порогові значення, ми попереджуємо користувача, про неправильне виконання вправи, і генеруємо

пропозиції для покращення виконання. Детальна інформація про геометричні оцінки інших вправ наведена в розділі Результати.

### 2.2.7 Оцінка за допомогою методів машинного навчання

Наш другій підхід оцінки рухів на опорних точок базується на аналізі даних та технологіях машинного навчання. Оскільки записані відео можуть мати довільну довжину, це призводить до зміни довжини векторів опорних точок в кожному прикладі. Класифікація з різною довжиною вектора ознак є складною задачею для машинного навчання. Ми розв'язуємо цю проблему за допомогою DTW [6] з класифікатором найближчого сусіда [7].

DTW — це показник, який використовується для вимірювання нелінійної подібності між двома часовими рядами (Рисунок 7). Коли дві подібні послідовності зміщені по фазі (тобто зміщені у часовому вимірі), такі показники, як Евклідова відстань[8] не можуть бути обчисленими, тому що це пряме порівняння двох точок в один момент часу. З DTW ми намагаємося динамічно ідентифікувати опорно точку другої послідовності рухів, що відповідає заданій послідовності у першій.

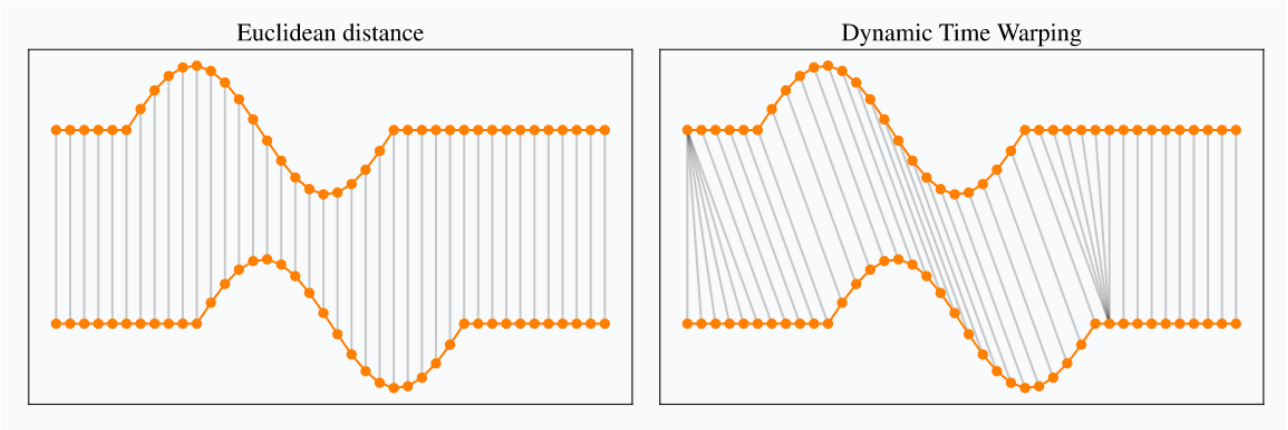


**Рисунок 8.** Dynamic Time Warping

Маємо дві послідовності опорних точок $P \in \mathbb{R}^m$ та $Q \in \mathbb{R}^n$, будуємо матрицю відстаней $D \in \mathbb{R}^{m \times n}$, де $D_{i,j}$ це Евклідова відстань між точками $p_i$ та

$q_j$. Перебираємо усі точки матриці за допомогою динамічного програмування, знаходимо оптимальний збіг точок у кожній послідовності і обчислюємо відстань між відповідними точками. Недоліком DTW є те, що він не стійкий до шуму. Коли OpenPose генерує шумні опорні точки, це впливає на продуктивність DTW. Щоб покращити результати ми двічі застосовуємо медіанний фільтр[9] на послідовність опорних точок перед обчисленням мір. Потім беремо послідовність опорних точок і обчислюємо DTW відстань на усіх навчальних послідовностях. Врешті решт створюємо двійковий класифікатор найближчого сусіда, який прогнозує «правильна» чи «неправильна» форма на основі відстані DTW.

# РОЗДІЛ 3. РЕЗУЛЬТАТИ

## 3.1 Вправи

Демонструємо кількісні та якісні результати роботи нашої системи із чотирьох різних вправ з гантелями: скручування на біцепс, підйом спереду, знизування плечами та жим плечей стоячи. Для кожної вправи ми використали як геометричний/евристичний підхід, так і підхід машинного навчання з використанням DWT.

### 3.1.1 Скручування на біцепс

Скручування на біцепс однією рукою — це вправа, яка ізолює біцепс, великий м’яз у верхній частині руки, що відповідає за згинання та скручування в лікті. У скручування на біцепс гантель піднімається вгору, обертаючись навколо ліктя, з положення спокою, при цьому інші частини тіла залишаються нерухомими. Ця вправа спрямована на укріплення двоголового м’язу плеча або біцепсу. Поширені помилки під час виконання скручування на біцепс включають використання плеча, щоб допомогти підняти вагу вгору, а також піднімати вагу не повністю.

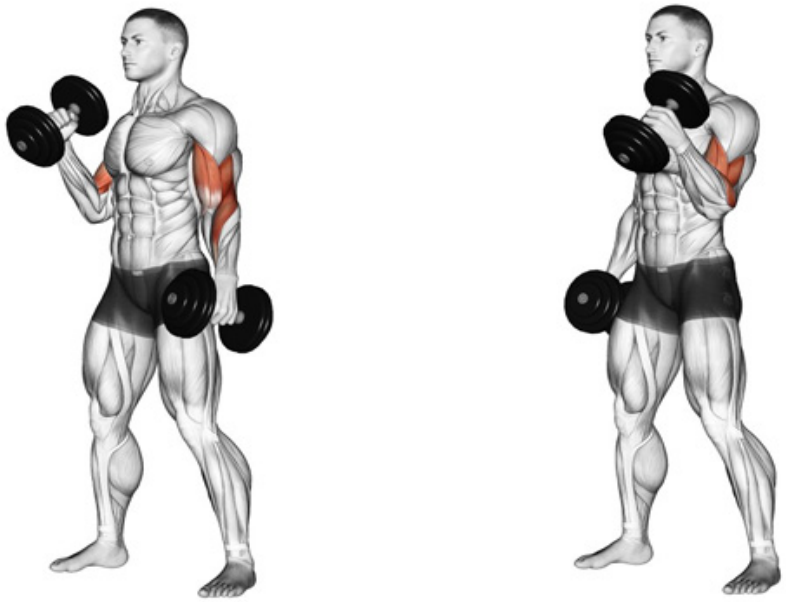

**Рисунок 9.** Скручування на біцепс

У нашому евристичному алгоритмі ми враховуємо діапазон кута між верхньою частиною плеча і тулубом (вимірювання відбувається коли користувач крутить плечем при виконанні вправи), а мінімальний кут між плечем та передпліччям (означає висоту на яку користувач піднімає вагу). Якщо діапазон між верхньою плечем і тулубом виходить за 35°, ми визначаємо це як значне обертання плеча.

Якщо мінімальний кут між плечем та передпліччям більше за 70°, це означає не виконання вправи до кінця.

Наш геометричний алгоритм класифікував усі відеозаписи на яких вправа виконана коректно як «правильні», 80% записів де вправа виконана завідомо некоректно він визначив як «неправильні». Після візуального аналізу помилкових результатів стало зрозуміло що виконання вправ на них знаходиться на межі між належною та неналежною формою виконання. Як і очікувалося, налаштування алгоритмів коригує межі рішення, а отже, і суворість під час оцінки вправи користувача.

Для нашого алгоритму машинного навчання ми розділили наш датасет з 16-відеозаписів на 9 тренувальних та 7 тестувальних. DTW класифікатор досягнув точності у 0.85. Точність на всіх розглянутих нами вправах зібрана у Таблиці 1.

Нижче наведений приклад виводу результату програми для коректного виконання вправи:

*обробка відео файлу ...*
*Рука яка виконує вправу: права*
*Кут між плечем та тулубом:*
*→21.138462278650042*
*Мінімальний кут між плечем та передпліччям:*
*→40.84669523416520*
*Вправу виконано відмінно!*

*Вправу виконано відмінно! Вага була піднята:*

*→до верху, і плече не рухалося*

Нижче наведений приклад виводу результату програми для некоректного виконання вправи:

*обробка відео файлу ...*

*Рука яка виконує вправу: права*

*Кут між плечем та тулубом:*

*→36.18554253665129*

*Мінімальний кут між плечем та передпліччям:*

*→31.27681963442689*

*Виконання вправи може бути покращене:*

*Ваше плече здійснює значний рух*

*→намагайтеся тримати плече нерухомим, паралельно до тулуба*

*→зконцентруйте рух тільки на вашому лікті*

### 3.1.2 Піднімання гантелей спереду

Підйом гантелей спереду — це вправа з вільною вагою, спрямована на плечі, зокрема на передню частину дельтоподібних м’язів. Користувач спочатку тримає гантелі на боці а потім піднімає їх перед собою, переважно з прямими руками. Виконавець повинен тримати тіло нерухомим, щоб ізолювати навантаження на плечі, також піднімати гантелі слід трішки вище плечей щоб виконати повний цикл вправи. Поширені помилки включають черезмірне напруження та використання рухів тулуба для полегшення навантаження, а також не піднімання вище плечей.

Геометричний алгоритм фронтального підйому вимірює горизонтальний діапазон руху спини, а також максимальний кут між тулубом і рукою. Рух спини вимірюється шляхом обчислення найбільшої зміни вектора опорних точок спини протягом усього відеозапису, щоб визначити чи полегшує користувач піднімання ваги шляхом розхитування тулуба. Кут нахилу рук відносно тулуба, для того щоб визначити чи повністю користувач виконує вправу.

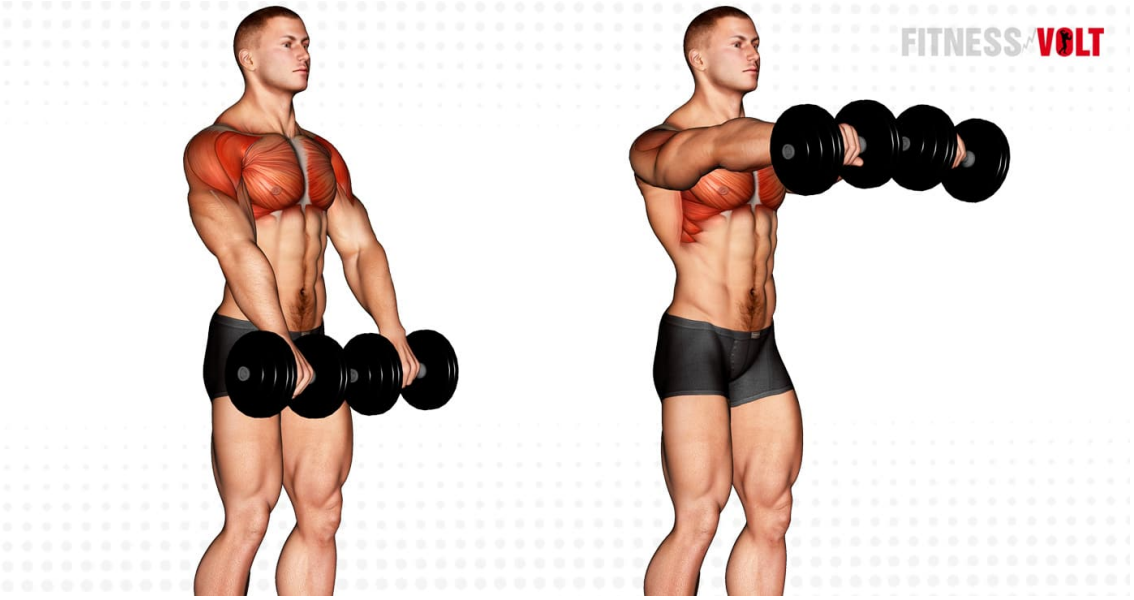


**Рисунок 10.** Піднімання гантелей спереду

Для алгоритму машинного навчання ми розділили датасет з 24 відео на 13 навчальних та 11 тестових прикладів. Класифікатор DTW досягає ідеального результату.

Система може генерувати такий відгук для користувача:

“*Ваша спина демонструє значний рух. Спробуйте тримати спину прямо та нерухомо, коли піднімаєте вагу.*”

“*Подумайте про використання легшої ваги*”

*“Ви не піднімаєте вагу до кінця. Намагайтеся підняти ваги до моменту коли зап’ястя будуть на рівні плечей або трохи вище”*

### 3.1.3 Знизування плечами з гантелями

Знизування плечима з гантелями — це вправа, яка зосереджена на плечі, зокрема на верхню трапецію. Користувач тримає гантелі з бокім і піднімає плечі настільки наскільки це можливо, зберігаючи тулуб в статичному положенні. Поширені помилки включають не виконання вправи на весь діапазон руху в плечах, а також згинання ліктів і використання рук при підйомі.

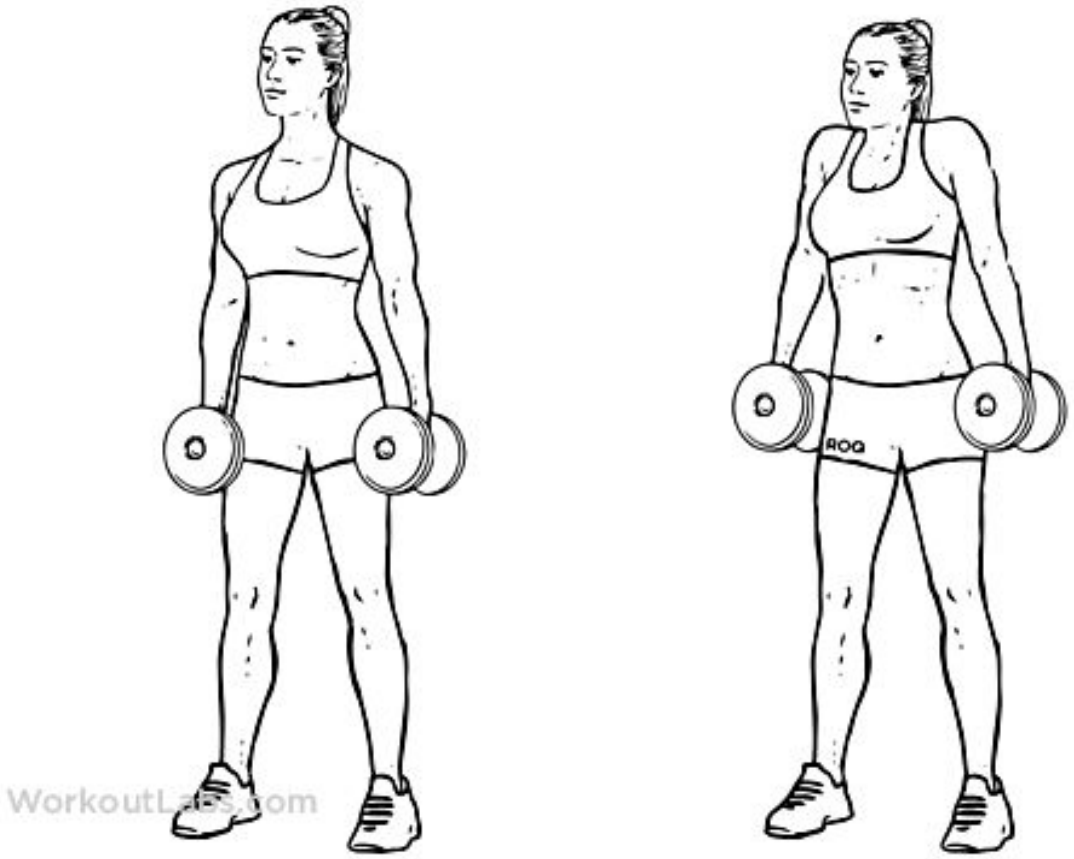


**Рисунок 11.** Знизування плечей з гантелями

Геометричний алгоритм для вправи знизування плечами обчислює діапазон руху плечей та кут між плечем і передпліччям. Якщо рух плечей дуже незначний, система визначає це як некоректне виконання вправи. Якщо кут між плечем та передпліччям малий, це означає що користувач занадто сильно згинає лікоть.

Для алгоритму машинного навчання ми розділили датасет з 30 відео на 17 навчальних та 13 тестових прикладів. На цій вправі DTW класифікатор досягає точності у 0.85.

Система може видати наступний відгук користувачу:

*“Ваші плечі недостатньо рухаються. Під час виконання вправи сильніше стискайте та підіймайте плечі”*

*“Ваші руки згинаються під час підйому. Тримайте руки рівно та нерухомо, зосередьтесь на руху плечей.”*

### 3.1.4 Жим плечей стоячі

Жим гантелей від плечей — це тренувальна вправа, де гантелі витягуються прямо вгору. Вправа опрацьовує багато груп м’язів плечей, грудей та частини спини. У цій вправі можна зробити багато помилок: підняти гантелі занадто вперед або назад, сильно рухати тіло під час виконання, недостатньо високо піднімати ваги.

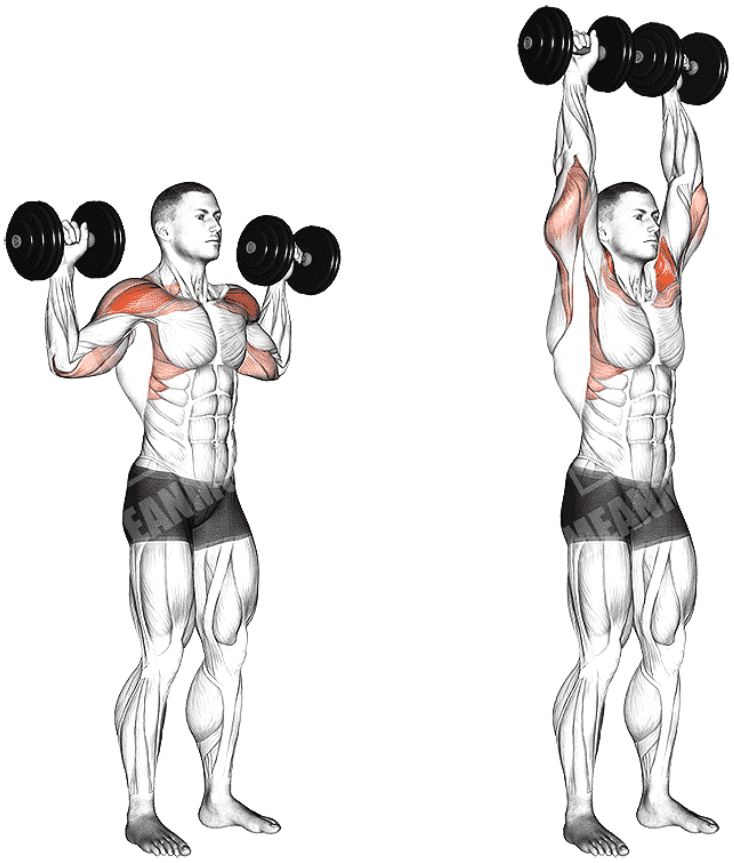

**Рисунок 12.** Жим плечей стоячі

Геометричний алгоритм для жима плечей вимірює діапазон руху спини за допомогою опорних точок шиї та стегна, діапазон руху руки за допомогою ліктя та шиї, максимальний кут між плечем та передпліччям. Якщо рух спини занадто великий, ми попереджаємо користувача тримати спину рівно.

Якщо опорна точка ліктя розташовується за шиєю, ми просимо користувача не крутити плечами під час відтискання вагів. Якщо кут між плечем та

передпліччям занадто малий, ми попередимо користувача, щоб він підняв вагу до кінця.

Для нашого алгоритму машинного навчання ми розділили наш датасет з 35 відео на 20 навчальних та 15 тестувальних прикладів. DTW класифікатор досягає точності у 0.73.

Система може згенерувати наступний відгук користувачеві:

*“Ваша спина демонструє значний рух під час відтискання вагів. Намагайтеся тримати спину рівною і нерухомою, коли піднімаєте вагу”*

*“Ви крутите плечі, коли підіймаєте гантелі. Намагайтеся тримати плечі рівно”*

*“Ви не підіймаєте вагу до кінця. Намагайтеся виконувати вправу повністю або спробуйте зменшити вагу*

## 3.2 Точність

Точність DTW класифікатора представлена в Додатку А. У таблиці представлені 3 метрики: Precision, Recall та F1.

Precision – рахує яка частина наших передбачень “правильно” дійсно збулася.

Recall – коли в прикладах було “правильно”, рахуємо яку частку із цих випадків модель передбачила правильно.

F1 – це комбінація Recall і Precision: $F1 = 2 * \frac{recall * precision}{recall + precision}$.

## 3.3 Наступні етапи розвитку системи

Звичайно наша програма поки що не є повністю готова для застосування в реабілітаційних центрах, але слід зазначити що основну задачу вона вирішує. Залишилося ще багато роботи по вдосконаленню інструментарія системи для полегшення взаємодії з пацієнтом.

Почнемо з того, що потрібно розробити алгоритм, якій би розпізнавав початок виконання тієї чи іншої вправи, аби позбавити користувача таких незручностей, як запис та редагування відео вручну. До того ж, ця функція зробить можливим роботу програми в реальному часі, але на досить потужних комп'ютерах. Потім з'явиться необхідність у голосовому помічнику, щоб відгук який згенерувала програма одразу ж був озвучений і пацієнт скорегував своє виконання вправи.

Також є необхідною умовою є створення бази даних з відеозаписами, і додавання методів аналізу даних динаміки реабілітації пацієнтів. Щоб лікарі могли побачити чи вдається пацієнту краще виконувати вправи, робити більше повторень тощо.

Якщо в нашій системі буде збереження даних пацієнтів, нам потрібно буде подумати про конфіденційність, про це було згадано у Розділі 1. Рішенням можуть стати алгоритми комп'ютерного зору, які розмивають обличчя для конфіденційності.

Кожен пацієнт є унікальним і програма фізичних вправ для всіх різна. Евристичні алгоритми аналізу рухів показали себе досить добре, і їхня реалізація в програмному коді може бути перенесена в програму з користувацьким інтерфейсом для створення таких алгоритмів, щоб научений клініцист без знань програмування міг створювати унікальні вправи для своїх пацієнтів.

# ВИСНОВКИ

Ми розглянули сучасні підходи до побудови систем реабілітації. Розібрали основні пункти розробки таких систем та технології які в них застосовуються.

У цій роботі ми представили додаток для реабілітації, якій використовує технології комп’ютерного зору, оцінює позу, аналізує дані опорних точок скелетону евристичними алгоритмами та технологіями машинного навчання для надання персоналізованого відгуку про коректність виконання фізичних вправ.

На вхід оціночних алгоритмів подаються дані скелетону людини отримані за допомогою OpenPose. Програма може аналізувати чотири різні вправи: скручування на біцепс, підіймання гантелей спереду, знизування плечима з гантелями та жим гантелей стоячи. Для оцінки коректності виконання вправ ми використовуємо два підходи: евристичний та алгоритм машинного навчання. Також програма генерує відгук, як покращити виконання вправи.

Тестування нашої програми показало гарні результати, які також зазначені в підрозділі 3.2.

Також сформували список етапів розвитку системи в майбутньому, які інструменти та функції потрібно додати, щоб створити повноцінну систему реабілітації.

# ДОДАТОК А

| | **Precision** | **Recall** | **F1** | **Набір даних** |
|---|---|---|---|---|
| Скручування на біцепс | | | | |
| Коректно | 0.80 | 1.00 | 0.89 | 4 |
| Некоректно | 1.00 | 0.67 | 0.80 | 3 |
| **Середнє** | **0.89** | **0.86** | **0.85** | **7** |
| Піднімання гантелей спереду | | | | |
| Коректно | 1.00 | 1.00 | 1.00 | 6 |
| Некоректно | 1.00 | 1.00 | 1.00 | 5 |
| **Середнє** | **1.00** | **1.00** | **1.00** | **11** |
| Знизування плечами з гантелями | | | | |
| Коректно | 1.00 | 0.75 | 0.86 | 8 |
| Некоректно | 0.71 | 1.00 | 0.83 | 5 |
| **Середнє** | **0.89** | **0.85** | **0.85** | **13** |
| Жим плечей стоячі | | | | |
| Коректно | 0.67 | 0.86 | 0.75 | 7 |
| Некоректно | 0.83 | 0.62 | 0.71 | 8 |
| **Середнє** | **0.76** | **0.73** | **0.73** | **15** |

**Таблиця 1.** точність DTW класифікара

# СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1.Zhang, H.-B., Lei, Q., Zhong, B.-N., Ji-Xiang, D., Peng, J.L.: A survey on human pose estimation. Intelligent Automation & Soft Computing 22(3), 483–489 (2016)

2.Herath, S., Harandi, M., Porikli, F.: Going deeper into action recognition: A survey. Image Vis. Comput. 60, 4–21 (2017)

3.Duarte, N.F., Rakovic, M., Marques, J., Santos-Victor, J.: Action alignment from gaze cues in human–human and human–robot interaction. In: Proceedings of the European Conference on Computer Vision (ECCV) (2018)

4.Zijie Ye, Haozhe Wu, Jia Jia: Human motion modeling with deep learning: A survey(2021)

5. https://en.wikipedia.org/wiki/Analysis_of_variance

6.https://cmu-perceptual-computing-lab.github.io/openpose/web/html/doc/md_doc_00_index.html

7.Elisha Odemakinde: Mask R-CNN: A Beginner’s Guide

8.https://github.com/MVIG-SJTU/AlphaPose

9. YU XINBO Practical Algorithms for Vision-Based Human Activity Recognition and Human Action Evaluation

10.Gladstone, D.J., Danells, C.J., Black, S.E.: The fugl-meyer assessment of motor recovery after stroke: a critical review of its measurement properties. Neurorehabil. Neural Repair 16(3), 232–240 (2002)

11.Zhe Cao, Tomas Simon, Shih-En Wei, Yaser Sheikh:Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields

12.Bishop: Pattern Recognition and Machine Learning

13.Fundamentals of Deep Learning: Designing Next-Generation Machine Intelligence Algorithms 1st Edition

14.Yuanpeng Zhang: Computer Vision-Driven Evaluation System for Assisted Decision-Making in Sports Training