Conference Paper PDF Available

Pose Estimation and Virtual Gym Assistant Using MediaPipe and Machine Learning

September 2023
DOI:10.1109/NMITCON58196.2023.10275938
Conference: 2023 International Conference on Network, Multimedia and Information Technology (NMITCON)

Authors:

Urmi Dedhia
Carnegie Mellon University

Pratham Bhoir

Prateek Ranka
University of Southern California

Pratik Kanani
Dwarkadas J. Sanghvi College of Engineering

Citations (14) References (23) Figures (2)

Figures

Landmarks of Mediapipe's Bla...

Squats exercise implementation

Figures - uploaded by Pratik Kanani Author content

2023 International Conference on Network, Multimedia and Information Technology (NMITCON)

Pose Estimation and Virtual Gym Assistant using MediaPipe and Machine Learning

1st Urmi Dedhia
Department of Computer Engineering
Dwarkadas J. Sanghvi College of Engineering
Mumbai, India
dedhia.urmi20@gmail.com

2nd Pratham Bhoir
Department of Computer Engineering
Dwarkadas J. Sanghvi College of Engineering
Mumbai, India
prthmbhr28@gmail.com

3rd Prateek Ranka
Department of Computer Engineering
Dwarkadas J. Sanghvi College of Engineering
Mumbai, India
prateekranka1607@gmail.com

4th Pratik Kanani
Department of Computer Engineering
Dwarkadas J. Sanghvi College of Engineering
Mumbai, India
pratikkanani123@gmail.com

ITCON) | 979-8-3503-0082-6/23/$31.00 ©2023 IEEE | DOI: 10.1109/NMITCON58196.2023.10275938

***Abstract*—This research paper demonstrates a virtual fitness trainer, using Google's Mediapipe library, which was made for various multimodal machine learning and deep learning pipelines. By monitoring the user's every step with posture estimation algorithms, the suggested system can provide useful feedback and insights as they work out. Mediapipe's deep algorithms and posture estimation module are utilized to develop this novel system, which captures user movements by identifying unique body landmarks for each exercise. The system also keeps track of the total amount of reps quite precisely. The system determines the user's performance by computing angles and landmarks and feeding this data into a machine-learning model, which then classifies the correct posture and the number of repetitions of the exercise. The system uses many fitness datasets in addition to user-supplied data for thorough training and assessment. A detailed comparison of machine learning models like Logistic Regression, Support Vector Machine, Naive Bayes, Decision Tree, and Artificial Neural Network is included in this research. The findings prove that our suggested method**

indispensable for our topic as the paper aims to estimate tl correct posture of a particular exercise. This paper also us machine learning and deep learning to assess pose accura and exercise technique and com- pares machine-learni accuracy. Unique datasets are used to test strategies with a without hyperparameter adjustment. Artificial Neur Network (ANN) is employed to calculate accuracy a compared all methods to find the best for this project.

The following sections are provided in chronological der in the paper: Literature Survey, Proposed Methodolo Algorithm Summary, Dataset, Implementation, Results a Discussion, Conclusion and Future Scope, and Reference

II. Literature Survey

Mediapipe-based pose estimation research is emerging field. Arpita Halder et al. compared SVM, KN and other ma- chine learning models to detect vernacul

2023 International Conference on Network, Multimedia and Information Technol

Keywords— Mediapipe, Computer Vision, Pose Estimation, Fitness, Exercises, Deep Learning, Machine Learning

I. INTRODUCTION

Physical activity is essential in today's society since it helps build stronger bones and muscles and enhances your capacity to carry out daily tasks. People today are more health- conscious and exercise daily, either at home or at the gym. People exercised at home during the COVID-19 pandemic. But this presents some issues. Injuries and other issues result from people not knowing how to work out. People also choose home workouts since gyms are expensive.

This study will show how to exercise at home without going to the gym every day. The number of reps performed were counted by identifying human body landmarks such shoulders, elbows, knees, wrists, ankles, etc. using the Mediapipe library, which Google developed. This paper includes rep counts for popular workouts like bicep curls, squats, front kicks, and jabs. Biomechanics is extremely

979-8-3503-0082-6/23/$31.00 ©2023 IEEE

Agarwal et al. briefly describe fitness-related AI applications. GOQii, Cure. fit, HealthifyMe, and others offer gym alternatives and promote fitness on smartphones. The authors suggested using machine learning to evaluate asanas. Mediapipe's BlazePose model detects postures by identifying and analyzing 33 pairs of critical spots on the human body. A pose-analyzing tool calculates landmark angles heuristically [4]. X. Li, M. Zhang, et al. used Mediapipe to identify basic fitness movements. Up and down detection states are divided into picture samples in the first of three fitness action recogni- tion algorithm phases.

splitting for training and testing. Machine learning mode use preprocessed data. According to the results, SV outperformed other algorithms [1]. In a paper (A. Singh al.), the challenges surrounding estimating a person's po are covered, along with an overview of significant po estimation research, covering both deep learning metho and conventional image-based approaches. The authors ha created a straightforward model using a convolutional neu network that estimates the postures and exemplifies t potential of CNNs after examining numerous findings a identifying the constraints [2]. Several works focus solely yoga and gym workouts. The writers of Chen, Steven et a built a gym exercises assistant model employing tw methods. The first method sets geometric heuristics threshold values manually using logical deductions. The se ond machine learning strategy uses the closest neighbor cla sifier and dynamic time warping. OpenPose, whi estimates posture using part affinity fields is used for po estimation. Vector fields encode limb location a orientation [3].

system (Pauzi, A.S.B. et al.) monitors body activity an adds labelled skeleton joints to a video. This solution use deep learning for sports and physically demanding job with PoseNet dataset and Mediapipe Blazepose algorithm Accuracy is found within 10% of IMU-based motio capture. The system labels and calculates joint movemen velocity and angles [12].

Ashish Ohri et al. showed how to implement real-ti pose estimation and correction in a user-friendly ap OpenPose, Pifpaf, Tensorflow-Lite, and TensorflowJS a

categorization findings by counting fitness activities. The Mediapipe Blazepose model is superior for customized AI fitness on mobile devices than Openpose, Alphapose, etc. [5]. A paper (by Kumar D. et al.) proposes an android application for yoga pose estimation using the OpenPose model for key landmark detection. The pipeline proposed uses a dataset of yoga poses which are preprocessed using OpenPose and stored on the local machine and then the real-time captured event of the pose is compared with the local machine data to determine the pose [6].

G. Taware et al. used occlusion-simulating augmentation to test the posture detector in strong occlusions, which are different from typical tests. Before employing user alignment, the estimator locates the user's 33 critical points. It uses regression and heatmaps. The test model's layers are trimmed after using the training model's algorithms. The heatmap was also used by the authors to analyze the encoder's lightweight integration [7]. A paper by Kanase et al. presents a method for pose estimation using the pre-trained OpenPose model, a multistage CNN model. This model finds key points in videos. Key points are normalized and cleaned. The authors suggest using a heuristic method and machine learning for pose estimation. The heuristic method uses the OpenPose model's landmark angles. The model uses dynamic time warping to classify videos. Euclidean distance between key points is computed and represented in a distance matrix. Best match points are chosen, and distance is calculated.[8]. In a paper, Anilkumar et al. used the BlazePose model by Mediapipe to obtain body joint coordinates. Coordinates are analyzed and compared with yoga position data. Users can set the flexibility threshold variable as needed. Feedback is given to users based on results [9].

Another paper also describes this same approach for AI-fitness repetition capabilities. The authors used the BlazePose model to get the 33 landmark coordinates and then computed the angles of the required joints and compared them to the geometric heuristics of the exercise [10]. In another study (Chen KY et al.), a fitness database was created using deep transfer learning. A neural network was trained to identify fitness movements. Researchers used Yolov4 and Mediapipe for quick fitness movement

yoga self-coaching system using transfer learning. The stuc started by collecting yoga posture data using a webcam, the using data aug- mentation techniques. Study analyz transfer learning using MobileNet model. They created a AI yoga system with a real- time prediction model. The yo posture classification method had a 98.43% accuracy rate our self-coaching system. It predicts outcomes and giv guidance for incorrect postures. Incorrect posture is detecte using the Mediapipe algorithm for joint angle calculati [14].R. Gadhiya et al. analyzed pose estimation models lil OpenPose, BlazePose, DeepPose, and Hourglass to discu their pros and cons in different scenarios. The analys included push-up exercise with OpenPose and BlazePos Results were compared and analyzed [15].

III. Proposed Methodology

The complete process pipeline is separated into tw stages which are described below.

A. Landmark Detection Stage

The subject is required to place themself in front of t camera so that the regions of interest are visible to it. Wi the help of Mediapipe, we first detect the subject/region interest(ROI) [15]. Using Mediapipe's Pose Landma Model, we detect the landmarks of various body parts, whi are discussed later.

B. Rep Counting Stage

- The landmarks of a particular frame are extract and the angles between the desired referen landmarks are calculated. The inputs are th passed to a threshold function based on the ang heuristic, which is different for each exercise. If t angle extracted satisfies the threshold, the rep cou is incremented by 1; otherwise, it is incremented 0.
- If the angles satisfy the threshold, the extracte landmarks are fed to a machine learning mode which classifies the form of the pose to determin how accurately the exercise is being performed. I the confidence score is above 60 percent for th correct form, the system validates the current re and increments the rep count. Otherwise, i

pleteness classifications had positive outcomes [11]. A has implemented and compared various ML model

and DL models for the form classification.

C. *Pose Detection and Angle calculation*

- Pose detection makes it easier to keep track of the posture during any exercise set. Pose detection models have a set of predefined landmarks which help in identifying various body parts easily.

Fig.1. Process Pipeline

Some examples of pose detection models include PoseNet, AlphaPose, OpenPose, and many others. After some research, BlazePose was the model selected for this proposed application. Fig. 1 illustrates the pipeline of the proposed posture estimation, rep counting, and

Google Research that can detect 33 key points of the hum body, including the head, torso, arms, and legs.

Fig.2. Landmarks of Mediapipe's Blaze Pose Model [16]

2) Algorithm used for detecting Squats: Squat analy can be performed by dividing motion into three compl domains. Upper body, lower body, and moveme mechanics. The lower body evaluates the alignment of t hip, knee, and ankle landmarks, the upper body focuses stability and posture of the head, neck, and stomach co and movement mechanics evaluate workout time a coordination. The shoulder, knee, hip, ankle, and foot joi and their spatial angles are used for posture or attitu calculation. The back and the thigh landmarks calculate t

detail.

1) The BlazePose Model: Mediapipe's pose detection API uses BlazePose. The BlazePose model performs slightly poorer than OpenPose but is suitable for Yoga/Fitness positions. The Pose Detection API detects landmarks in the proposed system. BlazePose's landmark coordinates are shown in Fig.2. It is a human pose estimation model developed by Google Research that can detect 33 key points of the human body, including the head, torso, arms, and legs. BlazePose is fast and accurate, making it ideal for real-time fitness tracking, virtual try-on, and gaming. BlazePose uses a lightweight neural net- work architecture with convolutional and depthwise separable convolutional layers for high accuracy and low computing complexity. The model learns from a huge dataset of annotated photos using supervised and self-supervised learning. We will extract landmarks to calculate angles for specific exercises, then predict the accuracy, and count reps [16].

Mediapipe's pose detection API implements the BlazePose model. The BlazePose model performs slightly worse compared to the OpenPose model, but it is an ideal choice for Yoga/Fitness poses. The proposed system uses the Pose Detection API for landmark detection. Fig. 2 below depicts the landmark coordinates from BlazePose's pose and landmark detector model which provides human pose tracking. It is a human pose estimation model developed by

BlazePose's pose and landmark detector model whi
provides human pose tracking of 33 landmarks [16][1
Given below is the heuristic value of the various angles us
to check for a correct and accurate squat posture. Fig.
below shows the squats exercise being done and the desir
landmarks being tracked by MediaPipe in real-time.

Fig.3.Squats exercise implementation

TABLE I. TABLE OF THRESHOLD HEURISTIC ANGLE VALUE FOR SQUATS EXERCISE

Angles	Heuristic Values	Remarks
Hip Angle	50 - 71 degrees	less than 44 degre will cause imbalar
Knee Angle	55 - 68 degrees	greater than 75 degrees will reduc the effect of the r

Citations (14) References (23)

... Estimation is conducting on 18 body keypoints, including face features and hands. Both have found extensive applications in fitness, health, sport analytics, gaming, animation, and human-computer interaction due to their time efficiency and wide range of platform adaptability in harsh environmental conditions (Roggio et al. 2024; Dedhia et al. 2023) . ...

Toward Robust Human Pose Estimation under Real-World Image Degradations and Restoration Scenarios

Preprint Full-text available

Apr 2025

Nada E. Elshami · Ahmad Salah · Amr Abdellatif · Heba Mohsen

View Show abstract

... These days, the put of choice for exercise is the Exercise center. When working out at an exercise center, you will be able to discover a coach who can coordinate you to do the correct exercise, it isn't basic to do work out at the exercise center and you've got to alter your time to the exercise center [7] [8]. This is the reason why individuals are lazier to exercise because it takes a lot of time and cash. ...

Improve Exercise Movement: Detecting Mistakes on Yoga with Mediapipe and MLP

Article Full-text available

Mar 2025

Mahda Laina Arnumukti · Andi Prademon Yunus Yunus · Aliyu Suleiman Babale

View Show abstract

... Finally, BlazePose has been used to provide real-time feedback on a person's position and movement, with an accuracy of around 90% in Increased Reality (IR) environments [28]. As an alternative to improve latency, other frameworks such as OpenCV [29] or the Medi-aPipe feature extractor [30] have been used, achieving accuracies of up to 95%; however, these proposals impact the preprocessing of the images before being interpreted by the model [22]. The combination of OpenPose with Azure Kinect using the KinOp tool [31] has proven to be an accurate solution for

CNN 1D: A Robust Model for Human Pose Estimation

Article Full-text available

Feb 2025

Mercedes Hernández de la Cruz · Uriel Solache · Antonio Luna-Álvarez · Dante Mújica-Vargas

View Show abstract

Asana AI – Yoga Pose Detection Using Pose Estimation and Real-Time Feedback System

Conference Paper

May 2025

Swetha Dhanasekar · S. Sindhu

View

Smart Gym Assistant: Posture Monitoring and Rep Counting

Conference Paper

Feb 2025

Arya Jayasankar · Krishnendu M Uday · Ishaan Shokeen · Lalitha S

View

Analysis of Weight-Bearing Squat Posture based on Deep Learning and Sports Science

Conference Paper

Sep 2024

Ji-Yi Chen · Kuanyu Lin · Chia-Li Chi · Chun-Jung Lin

View

Intelligent Posture Monitoring System with Real Time Notifications Using MediaPipe and OpenCV

Chapter

Feb 2025

Sneha Sathish Kumar · S. Madesh · B. Nikitha · Maheshwari

View

Human upper limb kinematics using a novel algorithm in post-stroke patients

Article

Jan 2025 · P I MECH ENG H

Porkodi Jayavel · Hari Krishnan S · Varshini Karthik · Ashokkumar Devaraj

Advanced Human Activity Recognition Using Pose Estimation and Deep Learning Techniques

Conference Paper Full-text available

Oct 2024

Haider Rasheed Hassan · Dheyab Salman Ibrahim · Ziyad Tariq Mustafa Al-Ta'i

View

Enhancing fitness action recognition with ResNet-TransFit: Integrating IoT and deep learning techniques for real-time monitoring

Article

Dec 2024

Jijie Li · Ruyao Gong · Gang Wang

View

Show more

Smart Gym Assistant: Posture Monitoring and Rep Counting

February 2025

Arya Jayasankar · Krishnendu M Uday · Ishaan Shokeen · Lalitha S

Read more