# POSE ESTIMATION AND JOINT ANGLE DETECTION USING MEDIAPIPE MACHINE LEARNING SOLUTION

**Katarina Mitrović, Danijela Milošević**

[1] Department of Information Technologies, Faculty of Technical Sciences, University of Kragujevac, Svetog Save 65, 32000 Čačak, Serbia
e-mail: katarina.mitrovic@ftn.kg.ac.rs, danijela.milosevic@ftn.kg.ac.rs

**Abstract**:

Health is one of the central aspects of life due to which innovative ways for its improvement are constantly being studied. Artificial intelligence has an extensive application and its contribution to health and medicine is widely recognized. In this paper, the application of machine learning algorithms in the field of health care is presented. A model for physical activity injury prevention based on the MediaPipe solution for body pose tracking has been developed. The solutions for pose estimation and detection of joint angles and angles relative to the horizontal are integrated into a comprehensive system that detects all key body landmarks and angles during the movement of the observed person. In addition, one of the goals of this research is to develop a flexible system with the ability to process a variety of inputs in terms of video content and format. The system is trained and tested on video inputs, and it can process front, left, and right perspectives. In the processing phase, a graph of posture and angle estimation is generated. The graph represents detected joints and the corresponding angles that vary depending on the observed perspective. The input is integrated with the graph and thus provides valuable information about body posture and alignment. The results provide support to professionals in physical activity monitoring and injury prevention.

**Keywords**: artificial intelligence, machine learning, MediaPipe, pose estimation, angle detection

## 1. Introduction

Throughout history and in all parts of the world, health has always been considered to be a central aspect of life. A great number of health conditions are affected by physical activity [1]. It brings benefits to the cardiovascular system, skeletal muscles, psychological function and so on [2]. Physical activity can be categorized into sports activity, habitual activity and work-related activity [3]. Guided practicing of physical activities is one of the preferred methods to organize a routine for injury prevention [4]. Papers [5], [6] and [7] propose intelligent fitness trainer systems based on human pose estimation without the help of a trainer that provide instant feedback about the posture. A lightweight 2D human pose estimation for a fitness coaching system is described in [8]. The paper [9] produces results for the detection of specific technique-related issues which are highly associated with risk of injury during common exercises. One of the important additions to pose estimation is detecting joint angles. In the research [10] a wavelet neural network is proposed for joint angle estimation. Paper [11] compares the performance of multilayer perceptron, long short-term memory, and CNNs for the prediction of joint kinematics and kinetics. In the research [12] a recurrent neural network is used for joint position estimation.

The main goal of this paper is to create a solution for exercise monitoring and delivering comprehensive but legible and user-friendly system. The main contribution of this paper is integrating various techniques such as pose estimation and angle detection as well as extending the current pose estimation model with multiple solutions. Additionally, the software provides identification of critical landmarks and joint angles depending on the perspective of the person performing the exercise, which addresses some of the problems faced by researchers on this topic. Determination of these elements is defined in cooperation with domain experts. In addition to the basic joint angles, the angles relative to the horizontal are also identified in order to complete the image of the body pose and perform a proper estimation of the posture. Another contribution of this paper is defining equations for scaling the graph to input data.

## 2. Methods

### *2.1 Input Data*

This paper uses videos that capture a person performing a specific physical activity as the input data. The model is not sensitive in terms of the environment and the appearance of the person in the video. Flexibility is achieved by creating three different modules: front, right and left side view. For each perspective, different landmark and angle segmentation is performed. Videos can also vary in terms of format, dimensions, fps and other format and quality determinants. All output components are scaled to different dimensions of the videos to make the output visible and the display design proportional, clear and user-friendly.

### *2.2 Pose Estimation*

Pose detection is preformed using MediaPipe Pose Landmark Model. MediaPipe is an open-source framework that offers fully extensible and customizable ML solutions [13]. The solution developed in this research includes three modules with different perspectives: front, left and right. Depending on the observed perspective, different pose landmarks are extracted. These points are crucial for each of the perspectives and provide the basis for angle detection.

### *2.3 Angle Detection*

For a comperhensive interpretation of the correct posture, the measurement of key angles was performed using custom functions. Measured angles can be divided into two groups: joint angles and angles relative to the horizontal. The angles are extracted depending on the observed perspective, the type of the angle and calculated based on the three points that make up the angle.

### *2.4 Design Scaling*

As input videos can vary in terms of dimensions, the graph design must be adaptable to the input for better readability and visibility of the results. This is achieved by scaling all display elements in relation to the dimensions of the input video. Line thickness *lt* is determined using the following equation:

$$lt = \max\left(\frac{\min(w,h)}{500}, 1\right) \tag{1}$$

where *w* represents image width, and *h* represents image height. The dimension of the landmarks is obtained by using the formula (1) and multiplying the result with the number 3. Text size scale *ts* is calculated using the equation:

$$ts = \frac{\min(w,h)}{2500} \tag{2}$$

In addition, scaling was performed over the background that visually distinguishes this segment of the display. The starting coordinates are determined using the following formula:

$$s = \left(p_x + lt * 6, p_y + \frac{th}{2}\right) \tag{3}$$

where $p$ is the point denoting the joint for which the angle estimation is made, while $th$ is the text hight. The ending coordinates are established using the following formula:

$$e = (s_x + tw + \frac{bl}{2} + \max\left(\frac{lt}{2}, 1\right) * 2, s_y - th\ ) \tag{4}$$

where $tw$ represents text width, and $bl$ is text baseline.

*2.5 Hardware and Software Requirements*

The proposed software was developed in the Python programming language using TensorFlow, MediaPipe, ffmpeg, opencv, NumPy, glob and pathlib libraries. The hardware configuration includes NVIDIA GeForce GTX 1650 Ti graphics processing unit, AMD Ryzen 5 4600H 3.00 GHz central processing unit and 8 GB of installed physical memory (RAM). It can be concluded that with the improvement of the processing configuration, there can be a significant improvement in the processing time, and additional optimization could lead to real time processing and display of posture estimation and joint angle detection during physical activity.

## 3. Results

Depending on the chosen perspective, different points on the body are detected and a custom view is generated. Also, in addition to the estimation of the pose, the measurement of joint angles is performed. This system is the implementation of a comprehensive model that combines several different methods for evaluating posture and thus provides complete and highly accurate information to the person monitoring physical movements. This system can be widely used in medicine during physiotherapy and similar activities, as well as in fitness during exercise supervision. Therefore, coaches and medical staff are assisted in correcting the improper posture of the supervised person during physical activity, which prevents potential injuries and leads to the better overall health of the individual.

Another benefit of this paper is reflected in the flexibility of the system in terms of input. The software is applicable to inputs of a variety of content and formats, both in terms of reliable landmark prediction and in terms of processing and displaying the resulting output. The result is a fast, flexible and comprehensive software that detects key points on the body and provides insight into body alignment for computer-aided posture estimation.

The software has been tested on numerous inputs, including various examinees, perspectives, physical activities and other elements of the video.

## 4. Conclusions

In this paper, a system that combines several different methods for injury prevention is presented. It applies the most modern techniques of AI in order to provide reliable information about the posture of a person during exercising. In this study, the integration of posture estimation and joint angle detection was performed in order to identify potential problems that could lead to injuries. In addition, the detection of key points on the body in relation to the horizontal is performed, which provides valuable information about body alignment.

The proposed software was developed locally using an average hardware configuration. In the following stages of development, the system can be web-based and adapted to real-time video processing in order to provide feedback on body posture during physical activity without delays. In addition, integration with the incorrect exercise performance detection can be done, which implies indicating the critical points and providing advice for correction. This would enable the system to provide complete assistance without the supervision of professionals.

**Acknowledgement:** This study was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia, and these results are parts of the Grant No. 451-03-68/2022-14/200132 with University of Kragujevac - Faculty of Technical Sciences Čačak.

## References

[1] Vuori I, Does physical activity enhance health? Patient Education and Counseling, 33, pp. S95-S103, 1998.
[2] Health Education Authority, Sports Council. Allied Dunbar Fitness Survey. London: Belmont Press, 1992.
[3] Schmidt SC, Tittlbach S, Bös K, Woll A, Different types of physical activity and fitness and health in adults: an 18-year longitudinal study. BioMed research international, 2017.
[4] Mendonça LD, Ley C, Schuermans J, Wezenbeek E, Witvrouw E, How injury prevention programs are being structured and implemented worldwide: An international survey of sports physical therapists. Physical therapy in sport, 53, pp. 143-150, 2022.
[5] Agarwal S, Gupta M, Khandelwal S, Jain P, Aggarwal A, Singh D, Mishra VK, FitMe: A Fitness Application for Accurate Pose Estimation Using Deep Learning. 2nd International Conference on Secure Cyber Computing and Communications, IEEE, pp. 232-237, 2021.
[6] Chen S, Yang RR, Pose Trainer: correcting exercise posture using pose estimation. ArXiv preprint arXiv: 2006.11718, 2020.
[7] Zou J, Li B, Wang L, Li Y, Li X, Lei R, Sun S, Intelligent fitness trainer system based on human pose estimation. International Conference on Signal and Information Processing, Networking and Computers, Springer, Singapore, pp. 593-599, 2018.
[8] Jeon H, Yoon Y, Kim D, Lightweight 2D human pose estimation for fitness coaching system. 36th International Technical Conference on Circuits/Systems, Computers and Communications, IEEE, pp. 1-4, 2021.
[9] Eivindsen JE, Kristensen BY, Human Pose Estimation Assisted Fitness Technique Evaluation System, Master's thesis, NTNU, 2020.
[10] Sivakumar S, Gopalai AA, Lim KH, Gouwanda D, Chauhan S, Joint angle estimation with wavelet neural networks. Scientific reports, 11(1), pp.1-15, 2021.
[11] Mundt M, Johnson WR, Potthast W, Markert B, Mian A, Alderson J, A Comparison of Three Neural Network Approaches for Estimating Joint Angles and Moments from Inertial Measurement Units. Sensors, 21(13), p. 4535, 2021.
[12] Koch P, Mohammad-Zadeh K, Maass M, Dreier M, Thomsen O, Parbs TJ, Phan H, Mertins A, sEMG-Based Hand Movement Regression by Prediction of Joint Angles with Recurrent Neural Networks. 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society, pp. 6519-6523, 2021.
[13] MediaPipe Pose. https://google.github.io/mediapipe/ [Accessed: 07.03.2022.]