BITLIS EREN UNIVERSITY
JOURNAL OF SCIENCE AND TECHNOLOGY

E-ISSN: 2146-7706

ACCURACY DETECTION IN SOME SPORTS TRAINING USING COMPUTER VISION AND DEEP LEARNING TECHNIQUES

Nurettin ACI[1], Muhammed Fatih KULUÖZTÜRK[2,*]

[1] *Bitlis Eren University, Graduate Education Institute, Turkey,* nurettin.aci@adalet.gov.tr
[2] *Bitlis Eren University, Electrical-Electronics Engineering Department, Turkey,* mfkuluozturk@beu.edu.tr
* *Corresponding author*

KEYWORDS

Deep learning
Sports biomechanics
Machine learning algorithms
Performance evaluation

ARTICLE INFO

Research Article
DOI: 10.17678/beuscitech.1330481

Received	31 July 2023
Accepted	23 October 2023
Year	2023
Volume	13
Issue	2
Pages	133-158

ABSTRACT

In this study, the performance of the MediaPipe Pose Estimation model in estimating body position in different sports activities was investigated in the light of biomechanical parameters. Additionally, the performance of the model was evaluated by comparing the real-time data obtained from the camera with different machine learning algorithms (regression, classification, etc.). The results showed that the MediaPipe Pose Estimation model is a suitable and effective tool for sports biomechanics. The model was able to estimate body position with high accuracy in different sports activities. Additionally, the performance of the model was improved by using different machine learning algorithms. This study is a pioneer research on the applicability of computer vision-supported deep learning techniques in sports training and pose estimation. The model has been developed into an application that can be used to improve the performance of athletes.

1 INTRODUCTION

Every year, a large number of athletes are injured as a result of participation in strenuous physical activity. These injuries can be costly to treat and can also lead to long-term health problems [1]. The frequency of these injuries can be reduced by proper training, technique, and fitness, which can help to identify factors that may make an athlete susceptible to injury [2]. Additionally, efforts to reduce these injuries require both efforts to ensure the long-term health of athletes and efforts to reduce medical costs [3].

The movements of living things are of interest to many researchers. Studies in this field include biomechanical analyses, performance analyses, person recognition, identification of movement disorders, and virtual human animation in computer graphics. These studies are based on the biomechanical properties of sports activities such as walking, running, and jumping. Biomechanics has emerged with the use of structural, physical, and mathematical models for the understanding and analysis of human movements [4].

The emergence of biomechanics has also been influenced by the challenges of numerically expressing and giving meaning to the factors that cause movement, as well as the high degree of freedom in the human body. The analysis and definition of human movements is a very challenging process and the development of biomechanics has been necessary [5].

The development of biomechanics has been greatly influenced by the integration of mathematics, physical principles, and engineering methodologies. The combination of these disciplines has contributed to the diversification and development of the application areas of biomechanics. The main purpose of biomechanics is to understand, analyse, and optimize human movements. In this direction, simple models can be used to represent the movements and explain some basic mechanical properties, while sports biomechanics is used for more complex motion structures. Sports biomechanics plays an important role in understanding the human movement capacity, the evaluation of the principles of injury and prevention mechanics, and the treatment of muscle and skeletal problems. The benefits of technology are also used to get the maximum output from the movements made with sports biomechanics [6].

Human movements are studied by using kinematic and kinetic data of body members. For this purpose, computer vision-based human pose estimation models have been developed. Human pose estimation aims to find human body parts with the data obtained from images and videos, and to create a three-dimensional human body skeleton [7].

The analysis of human movements is not limited to the evaluation of large data sets obtained by motion analysis systems with basic statistical methods [8]. Researchers perform motion analysis based on the observation that the positions of the joints that make up the human body parts do not change randomly over time and show a regular behaviour, even though they move at the point of high degree of freedom using different techniques [9].

Human pose estimation has been attracting increasing interest in recent years. Although deep learning-based solutions provide high performance in human pose estimation, there are still challenges such as insufficient training data, depth uncertainties, and occlusions [10].

"There are people familiar with using YouTube content and fitness applications to exercise at home. However, these individuals may exercise based on incorrect information from non-professionals, which can increase the risk of injury by performing exercises with incorrect posture or without considering their physical abilities [11]. While exercising at home may be convenient, it can be harmful to the body as it is difficult to monitor proper posture. Therefore, this article proposes a program that guides users while exercising at home without the need for a fitness instructor. This program utilizes MediaPipe Pose's landmark model to estimate body posture and displays instructions on the screen when the user's posture does not conform to proper exercise criteria. If the user's posture is correct, the number of exercises performed according to the displayed instructions increases."

The aim of this article is to systematically analyse the accuracy rates of sports activities performed using real-time images through computer vision-based deep learning solutions for human pose estimation from the perspective of sports biomechanics.

1.1 Relationship between Computer Vision and Deep Learning

Computer vision is a field that aims to transfer the vision processes of the human brain to the machine. The main goal in this field is to be able to identify and group objects in digital images by understanding the content of the images. In this way, the machine can perform extraction of an object, text or model on the image [12].

Deep learning is a type of machine learning. This method performs feature extraction and transformation operations using many nonlinear processing unit layers. Each consecutive layer takes the output of the previous layer as input. The algorithms can be used for both supervised learning (such as classification) and unsupervised learning (such as pattern analysis) [13].

Deep learning algorithms are used to distinguish objects in the field of computer vision and image processing. The most commonly used deep learning algorithm is the convolutional neural network algorithm, which provides an effective method for distinguishing objects in images. The basic components of this algorithm are convolutional, pooling and fully-connected layers. In the model training process, weights and bias values are optimized with repeated layers, forward and backward propagation operations, and object recognition tasks can be performed [14].

There are many algorithms used in the field of computer vision and deep learning, in addition to convolutional neural networks (CNN). Algorithms such as artificial neural networks (ANN), recurrent neural networks (RNN), long short-term memory (LSTM), deep belief networks (DBN) and learning-based classifier systems are preferred for different tasks. The choice of algorithm depends on the properties of the problem to be applied and the structure of the data set. The deep learning field is rapidly developing and new algorithms are constantly being discovered or existing algorithms are being developed [15].

1.2 Human Pose Estimation with Computer Vision

Human pose estimation is a computer vision technique in which a person's movements and positions are represented graphically using deep learning techniques. This technique uses a model-based approach to estimate the positions of a person's body parts or joint positions using 2D or 3D data. Pose estimation models are able to detect human poses using 2D and 3D pose estimation techniques. 2D pose estimation estimates the positions of body joints using X and Y coordinates, while 3D pose estimation estimates the object's true spatial location using an additional Z dimension. 3D pose estimation poses a significant challenge for machine learning engineers due to the need to create data sets and algorithms [16].

There are three main methods for solving the human pose estimation problem: Absolute Pose Estimation, Relative Pose Estimation, and Proportional Pose Estimation. Pose estimation algorithms are designed to estimate a person's position relative to the background using the human pose and direction. Models can be designed using bottom-up or top-down approaches, and the encoder-decoder architecture is frequently used. There are five popular deep learning-based pose estimation libraries for human pose estimation, and these libraries are used in the development of customized human pose estimation applications [17].

OpenPose: OpenPose is a free human joint detection library that can run in real time. It can simultaneously detect a total of 135 key points by detecting important points for body, face, hand, and foot estimation. It uses a bottom-up approach and provides an open source API, supporting different hardware architectures. OpenPose is widely used in various application areas [18].

Pose Detection: Pose detection is an open source library that can detect human poses in real time. This architecture, built on TensorFlow.js, allows the detection of body parts such as elbows, hips, wrists, knees, ankles, etc. for single or multiple poses. It is efficiently designed for light devices (browsers, mobile devices) and offers three different models, such as MoveNet, BlazePose, and PoseNet [19].

DensePose: Dense human pose estimation is an open source library that can match all human pixels in RGB images with the body's 3D surface-based model in real time. It runs on a Caffe2-supported detector framework and is used for single or multiple pose estimation problems [20].

AlphaPose: Alphapose is an open source, real time, and multi-person pose estimation library. It uses a top-down approach and helps to detect poses in case of misdetection of human bounding boxes. It uses the most appropriate architecture by determining human poses with the detected bounding boxes. It also provides

PoseFlow, an online pose tracker that can correctly detect single-person and multi-person key points in real time [21].

HRNet (High Resolution Network): HRNet is an architecture used for human pose estimation. It is used to determine key points based on specific objects or people in the image and estimates a very accurate key point heat map by preserving high resolution representations. HRNet is also suitable for detecting human poses in sports broadcast on television and has also been used in other dense estimation tasks [22].

1.3 Selection of Appropriate Human Pose Estimation for Sports Training

2D artificial learning models are generally successful, but sharper inferences are required for the perception of 3D objects. Mediapipe library, offered by Google as open source, is a tool that can be used in real time in many different areas that can make more detailed inferences such as size, position and orientation. This library can perform many operations from face detection to iris tracking, and works in a way that is compatible with different devices, resources and hardware.

Mediapipe library classifies human poses by detecting human body. This allows the accuracy of sports training to be determined using mathematical methods. The detection of body lines provided by Mediapipe is an important tool that can be used to evaluate the accuracy of sports training [23].

The key points of the user's body can be detected using the MediaPipe library. A list corresponding to the X, Y and Z coordinates of these points indicates the location of the body parts in the input image. MediaPipe's output contains the coordinates of a total of 33 key points (Figure 1). These points can be used to create a skeleton orientation. MediaPipe is an open source and customizable machine learning solution for real-time media streaming. The library is supported on different platforms such as Android, iOS, Python, JavaScript. MediaPipe's output provides a rough estimate of the human body structure and orientation in the given image or video stream. This output is obtained at a speed of 30 frames per second as specified by the frame rate [24].

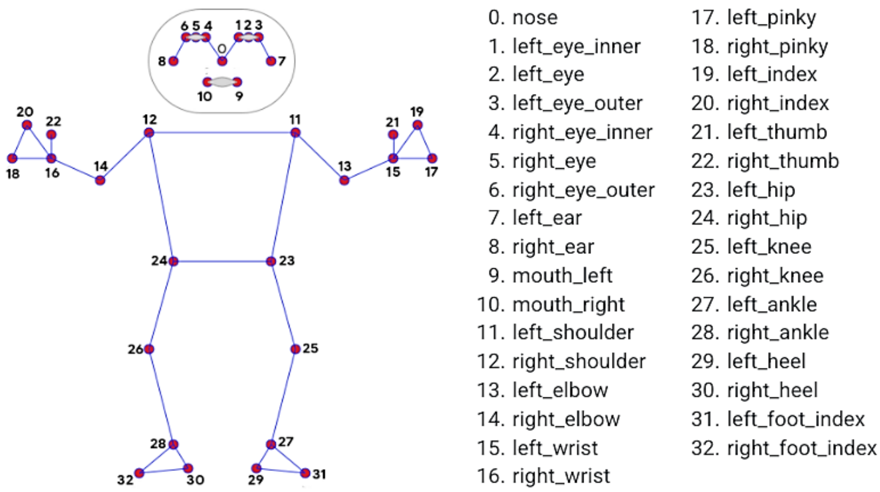

***Figure 1**. Key points noted in the MediaPipe Library.*

1.4 MediaPipe Library Model Architecture

The MediaPipe library uses heatmaps for image processing and analysis. Heatmaps are a method of visually representing how a certain feature or measurement changes over time. MediaPipe creates various models using heatmaps, especially for complex objects such as the human body and face. These models can be used for tasks such as hand gesture recognition, face recognition, body motion tracking, and image classification.

The MediaPipe library provides various tools and methods, such as data collection, data preprocessing, and machine learning techniques, to create models that achieve high accuracy and performance within the scope of deep learning algorithms, while it is creating models with the heatmaps method.

The models created are used to detect, track, and analyse specific features and objects with the combined use of heatmaps. The MediaPipe library is also useful for the detection and tracking of detailed objects, especially the human body (Figure 2). In addition to offering an easy-to-use and customizable interface, it allows users to combine and configure different models that support different functions according to their needs. These models offered by MediaPipe can be used in a variety of application scenarios, as well as being an ideal tool for the detection and tracking of objects with various details. [23]

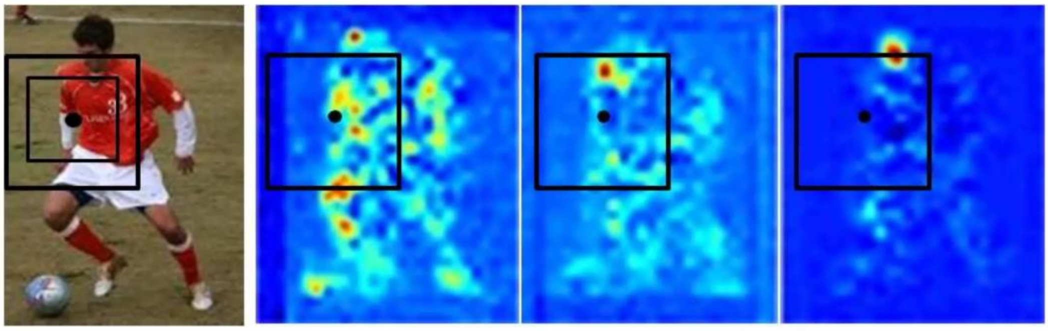

***Figure 2**. Use of Mediapipe Library Heatmaps.*

The MediaPipe library offers a new topology of 33 key points for the detection of the human body. This topology is a superset of the key points used in the BlazeFace, BlazePalm, and COCO datasets, and provides consistency between different datasets and models [24].

The MediaPipe topology allows it to estimate the rotation, size, and location of the relevant region with high accuracy and efficiency, using a minimum number of key points on the face, hands, and feet, compared to other popular topologies such as OpenPose and Kinect. The model reduces computational complexity and increases detection speed, as it uses only the necessary key points[25].

The MediaPipe topology is designed to provide suitable, accurate, and efficient person detection for a variety of applications, from augmented reality to sports analytics [26].

In contrast to other pose estimation solutions that detect key points using heatmaps, the underlying solution in MediaPipe is to initially perform a pose alignment.

The MediaPipe training dataset contains 60,000 images with one or several people, as well as 25,000 images of a single person doing sports exercises. The dataset is limited to cases where the entire person is visible or the hip and shoulder key points are clearly obtained. In addition, significant occlusion simulation augmentation is used to ensure that the model supports heavy occlusions that are not present in the dataset [27].

2 MATERIALS AND METHODS

2.1 Application for Movement Analysis for Sports Biomechanics

This study presents an application that detects the accuracy rate of sports activities using human pose estimation data obtained from real-time images. The application analyses the data collected using the MediaPipe Pose Estimation model in the scope of machine learning algorithms.

The application converts the input images created by a computer webcam or ready-made video files to RGB images from BGR images using OpenCV. Using the MediaPipe Pose Estimation model, it detects and records the body landmarks required for postural analysis. The position of the analysed position is compared with the pre-trained model in accordance with the x, y, z coordinates, the angle of the posture is calculated and it is verified whether it meets the criteria of the correct exercise posture. If the user's posture criteria are not met, appropriate warnings are displayed on the screen. If the posture is correct, the number of correction trainings is increased and the output is given on the screen (Figure3).

This study has shown that deep learning techniques can be used in addition to traditional methods in the field of sports biomechanics. This new approach aims to increase the usability and accuracy of athletes' performance and injury risk assessment. For our program, the most suitable exercises, squat (squat) and deadlift, were selected.

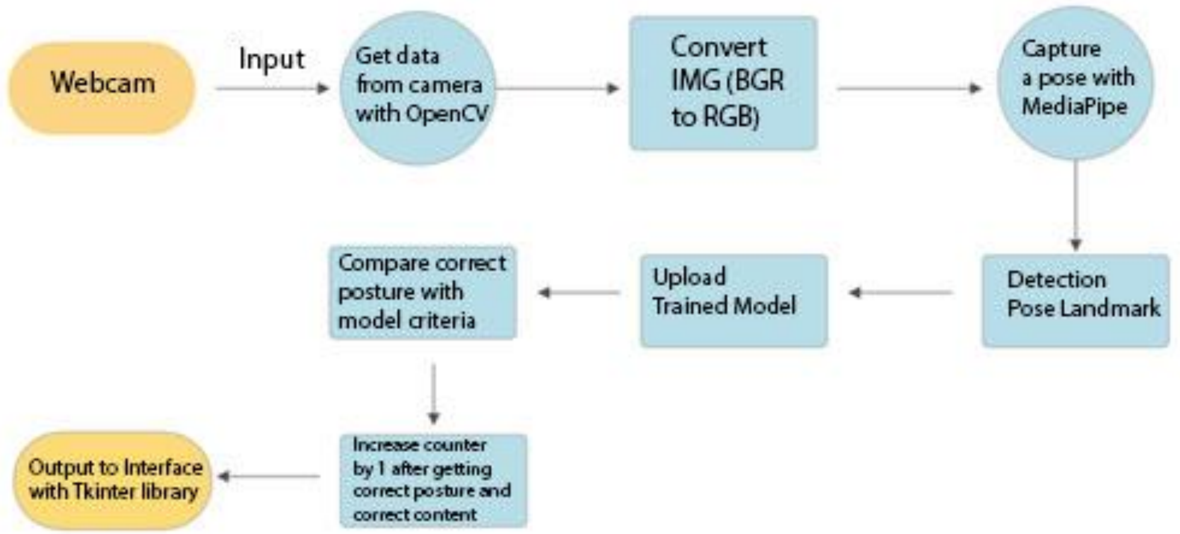

Figure 3. Correct position detection application processes from camera or video file.

2.2 Creating a Dataset from Video Files

Deep learning models often require large amounts of labelled data. This data is usually called a dataset and is used to train the model. However, it can sometimes be difficult to find a suitable dataset [28]. In these cases, the method of creating a dataset from video files can be used.

In this study, the following steps was followed to create a dataset from video files:

1. The images were extracted from the video recordings frame by frame.
2. The correct label was determined for each image.
3. The images were saved as a comma-separated values (CSV) file.

When labelling the images, kinematic data was used determined by the **American College of Sports Medicine (ACSM)**. According to ACSM, the position where the hips are at the lowest level in the deadlift exercise is the most effective and the hip angle should be between 120-140 degrees in this position [29]. In addition, the knee angle should also be between 100-120 degrees. In the squat exercise, the hip angle should be between 120-130 degrees [30]. The knee angle should be between 100-110 degrees. The back should be kept straight and the chest open. In addition, the head should be prevented from bending forward [31].

The dataset created consists of 7795 lines. Each line is labelled with 33 joint points and related kinematic data. The dataset has been used to achieve the correct position in the deadlift and squat exercises (Figure 4).

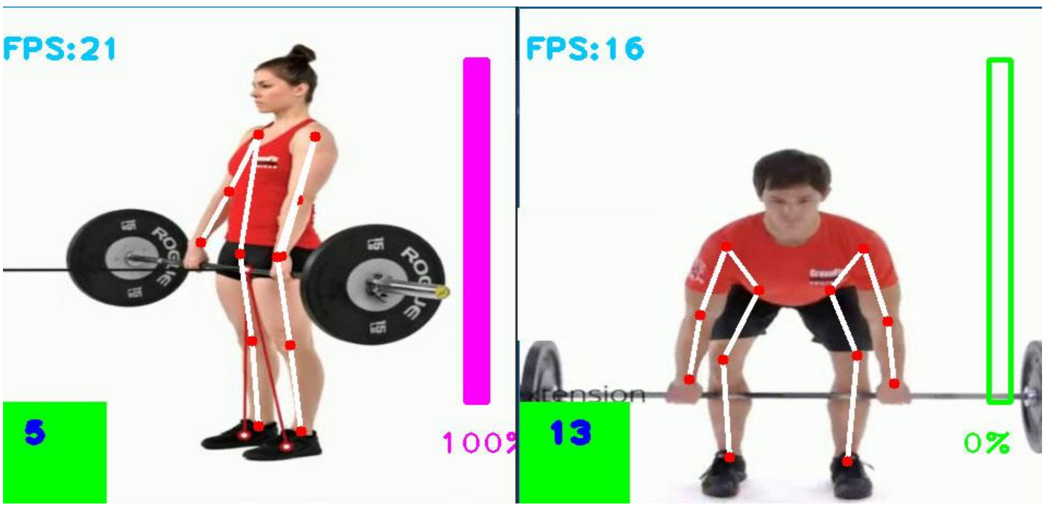

Figure 4*. Dataset creation within kinematic measurements.*

The MediaPipe joint points models used as base points to determine whether the correct posture conditions were met for both Deadlift and Squat. As can be seen in Figure 5, the camera angles of the right and left hips can be different, so the hip angle calculated by taking the average of the two values. Figure 5 shows the key points of the Mediapipe Pose Estimation model that used in the conditions to verify the correct squat and deadlift postures (Figure5).

Figure 5. *Kinematic angle values obtained from a certain camera angle.*

The hip angle (left knee (25), left hip (23), left shoulder (11)) and the right hip angle (right knee (26), right hip (24), right shoulder (12)) are calculated as follows based on the data obtained (Figure6):

Hip Angle = (Right Hip Angle + Left Hip Angle) / 2 **(1)**

The elbow angle (left wrist (15), left elbow (13), left shoulder (11)) and the right elbow angle (right wrist (16), right elbow (14), right shoulder (12)) are calculated as follows based on the data obtained (Figure6):

Elbow Angle = (Right Elbow Angle + Left Elbow Angle) / 2 **(2)**

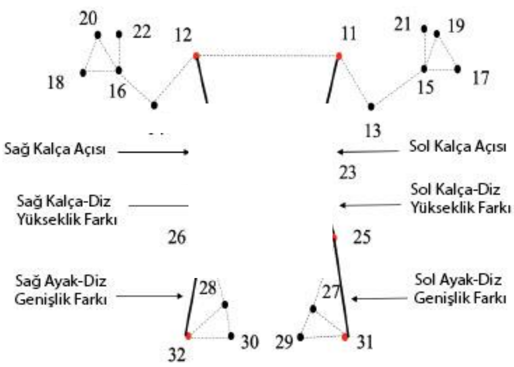

Figure 6. *Joint Points for Correct Posture Validity in Squat and Deadlift.*

For human pose estimation from a camera or video file, the x, y, and z coordinates are used to calculate the angle between three points. Figure 7 shows the angle between three points. Let the three points be A, B, and C. And the angle to be obtained is θ. The dot product of vectors was used to obtain θ.

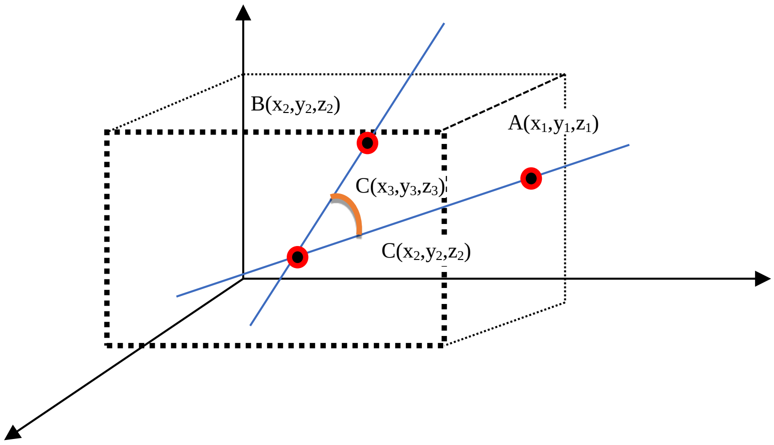

Figure 7. *Angle between three points.*

The vector with the viewpoint at point C and the endpoint at point A is denoted by $\underset{CA}{\rightarrow}$, and the vector with the viewpoint at point C and the endpoint at point B is denoted by $\underset{CB}{\rightarrow}$. The vectors $\underset{CA}{\rightarrow}$ and $\underset{CB}{\rightarrow}$, are defined as (3.2.1).

$$\underset{CA}{\rightarrow} = (x1 - x3, y1 - y3, z1 - z3)$$
$$\underset{CB}{\rightarrow} = (x2 - x3, y2 - y3, z2 - z3) \quad \textbf{(3)}$$

So θ is obtained by substituting the inner formula of the vector. In this way, the angle between three points can be defined as (3.2.2).

$$\underset{CA}{\rightarrow} \cdot \underset{CB}{\rightarrow} = \left|\underset{CA}{\rightarrow}\right| \left|\underset{CA}{\rightarrow}\right| \cos(\theta)$$
$$\theta = \cos^{-1} \frac{\underset{CA}{\rightarrow} \cdot \underset{CB}{\rightarrow}}{\left|\underset{CA}{\rightarrow}\right| \left|\underset{CA}{\rightarrow}\right|} \quad \textbf{(4)}$$

2.3 Data Modelling with Dataset

Dataset modelling is an approach that is often used in areas such as machine learning and artificial intelligence. This approach is the process of creating a model using the examples in a dataset. This model is then used to classify, predict, or explore new data [32].

In this study, the following steps was followed to do dataset modelling with the dataset prepared in csv file format in the light of kinematic data:

1. Our dataset was divided into 80% training and 20% test data.
2. The training data was used in the model's learning process.
3. The test data was used to measure the performance of the model.
4. A model was created using the features of the dataset.
5. The best performing algorithm was obtained by using different classifier algorithms such as Logistic Regression, Ridge Classifier, Random Forest Classifier and Gradient Boosting Classifier.
6. The performance of the model was calculated with performance metrics such as accuracy, precision and recall (Table 1).
7. The model was saved to be suitable for future predictions and to be used later as a Pickle file.

Table 1. *Model training dataset validation metric values.*

Algorithm	Accuracy	Precision	Recall
Logistik Regression	1.0	1.0	1.0
Ridge Classifier	1.0	1.0	1.0
Random Forest Classifier	0.998	0.996	0.994
Gradient Boosting Classifier	0.955	0.994	0.988

Pickle is very useful in cases where objects need to be serialized and saved. For example, the training of a machine learning model can take a long time, and the model obtained as a result of the training may be desired to be saved for future predictions. In this case, after the training of the model is completed, the model can be serialized using the Pickle module and can be used again later [33].

After completing the training of the model with the pickle library used for serialization of objects in the Python programming language, the csv file, which contains the attributes of **33 joint points** obtained using the Mediapipe pose estimation model and **7795** lines indicating whether the movement is upward or downward, and occupies approximately **8.22 Megabytes of space**, was saved in a smaller format using the Pickle module. The file obtained in this way was made suitable for the machine learning model by taking up **219 Kilobytes of space**.

2.4 Model Testing Phase

The model was tested our model in Python code after converting it to an appropriate format taking into account factors such as the quality, size and balance of the dataset. Our Python code uses the MediaPipe Pose Estimation model to detect the position of the human body by taking the video stream and converts this positioning into a data frame to test a model that recognizes body language. Then, this data frame is given to the classification model, which is pre-trained and saved as a pickle file, and tries to predict to which class the body language belongs.

This specific code provides a mechanism that counts the number of times the user stands up and squats properly in the deadlift or squat exercise defined above, in line with the movements the user makes. It detects the body language that the user shows during the "up" and "down" movements and keeps the number of "up" movements. This information is displayed in a box displayed on the screen. During the video stream, the class predicted by the model as well as the prediction probabilities are displayed (Figure 8).

***Figure 8**. Testing the obtained model.*

2.5 Interface Development and Accuracy Evaluation

In this study, an interface has been developed to evaluate the accuracy of sports exercises. The interface uses machine learning algorithms that provide real-time feedback by processing data from a camera or video file.

The interface uses a webcam to view the user's movement in real time and defines the body position through the Mediapipe Library. Then, it evaluates the predictions using a previously created machine learning model. The components, labels, buttons, and images in the interface are designed in a specific layout so as not to distract the user and to display the feedback and statistics when the user does not make the correct movements.

The graphic interface code includes the Tkinter, Pandas, Numpy, Pickle, Mediapipe, OpenCV and PIL modules. The program takes the video stream, defines the body structure, and predicts the accuracy of a specific posture. The results are shown in the user interface with the appropriate screen instructions, whether the movement is "up" or "down" based on the 33 joint points of the Mediapipe Pose Estimation model, with a similarity of 70% or more.

The interface is designed to help users perform their exercises correctly. The interface can help users learn the correct posture and reduce the risk of injury (Figure9).

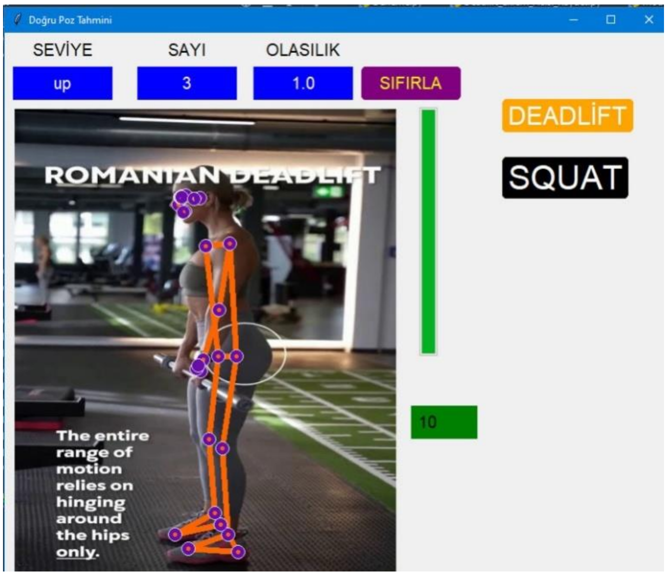

***Figure 9**. Model interface design screen.*

3 RESULTS AND DISCUSSION

3.1 Applicability of Mediapipe Human Pose Prediction to Sports Biomechanics

Sports biomechanics is not only a branch of science that investigates the mechanical aspects of the most complex movements of athletes and moving objects by examining them to fine details, but also measures the changes that occur in the body during movement by examining movements specific to the characteristics of athletes [34].

In addition, sports biomechanics is applied for different purposes by using it in fields such as sports sciences, physiotherapy, rehabilitation, orthopaedics in the study of human movement [35].

There has been an interest in the study of human-specific movements since ancient times and many researches have been conducted in this field. In the 1830s, under the leadership of the Weber brothers, the modern study of human movements began. The Weber brothers played an important role in the first gait analysis studies by using modern methods [36]. Most of the research carried out today continues under the influence of Winter's work [37].

Braune and Fischer, who were among the pioneers of biomechanical studies, proposed the use of high-speed recording systems for the study of human and athlete movements. In parallel with the advances in technology, the advancement of electronic and computer systems has enabled the development of new methods in the analysis of athlete movements [38].

In the field of sports biomechanics, there are many software and commercial motion analysis systems for athlete health and sports studies. Examples such as TASS (TNO Automotive Safety Solitiouns), LifeMOD, The AnyBody Modelling System, OpenSim and CATIA are some of the systems used as ergonomic design and analysis modules. Systems developed by APAS (Ariel Dynamics, Inc.), CODA (Charnwood Dynamics Ltd.), ELITE (Bioengineering Technology and Systems), OPTOTRACK (Northern Digital, Inc.), PEAK (Peak Performance Technologies, Inc.), QUALISYS (Qualisys Medical AB) and VICON (Vicon Motion Systems Ltd.) are among the most widely known in the field of motion analysis [39].

Again, Sim Mechanics is a software that combines Simulink and MATLAB (The MathWorks) tools to analyse uncomplicated athlete movements. HUBAG, designed by Hacettepe University biomechanics research group, is a Three Dimensional Motion Analysis software designed for academics, engineers and physicians. The software, which runs in MATLAB environment, is designed for users who want to analyse musculoskeletal systems. HUBAG aims to analyse the movements of athletes and improve athlete technique [40].

However, since the terms and methods used to describe and analyse the movements performed during exercises have an important place in movement analysis, it has become possible to perform these operations with computer vision in order to use these processes more effectively and efficiently with today's technologies. When the concept of computer vision is included in the field of sports biomechanics, it has become clear that even the most complex movements can be analysed and physiological changes can be measured during movements.

With technological advances, machine learning algorithms have made human pose detection and tracking possible. With the widespread commercialization of human pose estimation technology, it has significant implications in various fields such as security, business, health, entertainment and autonomous driving.

In addition, various applications of human pose estimation include human activity and motion estimation, augmented reality (AR) and virtual reality (VR) applications, robotics, animation and games. Human pose estimation has a wide range of application potential, from sports training to sign language communication.

While machine learning models developed today generally focus on 2D perception, Google's open-source Mediapipe library, a deep learning-based model pipeline tool for 3D object perception, is used in the application part of our study as it can perform many operations such as real-time pose estimation applications, face detection, iris tracking.

Deep learning models usually require large amounts of labelled data. This data is often referred to as a dataset and is used for model training. However, sometimes it can be difficult to find a suitable dataset [41]. In these cases, the method of creating a dataset from video files can be used. In this study, this method was used to create our own dataset.

While creating the dataset from video files, images were extracted from video recordings frame by frame and labelled these images. In addition, this method is especially used in areas such as object recognition, face recognition, human activity recognition.

Using the characteristics of the dataset, different classifier algorithms were used different classifier algorithms such as Logistic Regression [42], Ridge Classifier [43], Random Forest Classifier [44], Gradient Boosting Classifier [45], which are popular today and generally perform the best, to obtain our own model and transfer it to the application interface.

Our application shows that deep learning techniques can be used in addition to traditional methods in the field of sports biomechanics. Since it aims to increase the usability and accuracy of this new approach to assess athletes' performance and injury risk, a new approach was presented to the usability of human pose estimation data obtained using the MediaPipe Pose Prediction model in the field of sports biomechanics by determining the accuracy rate of sports activities.

3.2 Human Pose Estimation Libraries Comparison

Nowadays, as the demand for Human Pose Estimation has increased, many skeleton-based Human Pose Estimation algorithms have been developed and Human Pose Estimation algorithms have been packaged into libraries to provide ease of use for researchers. Since the performance of Human Pose Estimation libraries is important to ensure the reliability of the different practical applications they are integrated into, Jen-Li Cung et al. in their study ***"Comparative Analysis of Skeleton-Based Human Pose Estimation"*** conducted a comparative analysis of four state-of-the-art Human Pose Estimation libraries, namely PoseNet, MoveNet, OpenPose, and MediaPipe Pose, based on images and videos, despite challenges such as camera position and self-shutdown [46].

In this study, the performance of human pose estimation algorithms oflibraries is crucial in terms of the reliability of different applications. To this end, four state-of-the-art human pose estimation libraries, namely PoseNet [47], MoveNet [48], OpenPose [49], and MediaPipe Pose [50] have been comparatively analysed. These libraries employ the keypoint detection approach, which is divided into two classes:

top-down and bottom-up methods. In the top-down method, the number of individuals is determined from the input, and each person is assigned to a separate bounding box sequentially [51]. Subsequently, unlike the top-down method where keypoint estimation is performed in each bounding box, the bottom-up method performs keypoint detection in the initial step. After this stage, keypoints are grouped according to human instances [52].

The datasets used in the experiment for image and video sources are the Microsoft Common Objects in Context (COCO) [53] and Penn Action [54] datasets, respectively. COCO is a commonly used dataset for human pose estimation, containing 330,000 images and 1.5 million object instances. It also provides additional annotations for body keypoint detection, where each person is labelled with 17 keypoint locations. The COCO dataset and the Penn Action video dataset were used to compare the performance of the human pose estimation libraries. Data pre-processing was performed to eliminate irrelevant data in the experiments. The quality of the human pose estimation libraries was measured using the formula of Percentage of Detected Joints. The formula is as follows:

$$d(x-y) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}\ pixel \quad \textbf{(5)}$$

$$body\ diameter = \sqrt{(x_{ls} - x_{rh})^2 + (y_{ls} - y_{rh})^2} \quad \textbf{(6)}$$

$$percentage\ of\ detected\ joints = \frac{\sum_{i=1}^{n} bool(d_i < 0.05 * body\ diameter}{n} \quad \textbf{(7)}$$

The Euclidean distance calculation, denoted as d(x, y), between the ground truth key points (x_1, y_1) and the estimated key points (x_2, y_2) is shown in Equation (5). A threshold of 0.05 is used for the detected joint percentage based on the body diameter value. The body diameter is calculated as the Euclidean distance between the left shoulder and the right hip, represented by the coordinates (x_{ls}, y_{ls}) and (x_{rh}, y_{rh}), respectively, as shown in Equation (6). If the Euclidean distance d(x, y) between the estimated key points and the ground truth key points is below the threshold, the estimated key points are considered correctly detected. Therefore, the detected joint percentage is calculated as shown in Equation (7). The variable "n" represents the total number of estimated joints.

Table 2*. Number of images detected by human pose estimation in each specific proportion of detected joint percentage*

	0% < DJP ≤ 25%		25% < DJP ≤ 50%	50% < DJP ≤ 75%	75% < DJP ≤ 100%
MediaPipe Poz	240	342	116	127	275
OpenPoz	8	87	85	252	668
PoseNet	30	128	185	323	434
MoveNet	5	33	68	247	747

The performance of human pose estimation libraries is affected by challenges such as image clarity and inappropriate camera positioning. In this experiment, the performance of three different libraries, MoveNet, MediaPipe Pose, and PoseNet, was compared (Table 2). MoveNet achieved a higher value for the Detected Joint Percentage (DJP) compared to the other libraries. MediaPipe Pose obtained the highest maximum and median values for the Detected Joint Percentage. Consequently, MediaPipe Pose performed well on the video dataset, while MoveNet showed the best overall performance (Figure10).

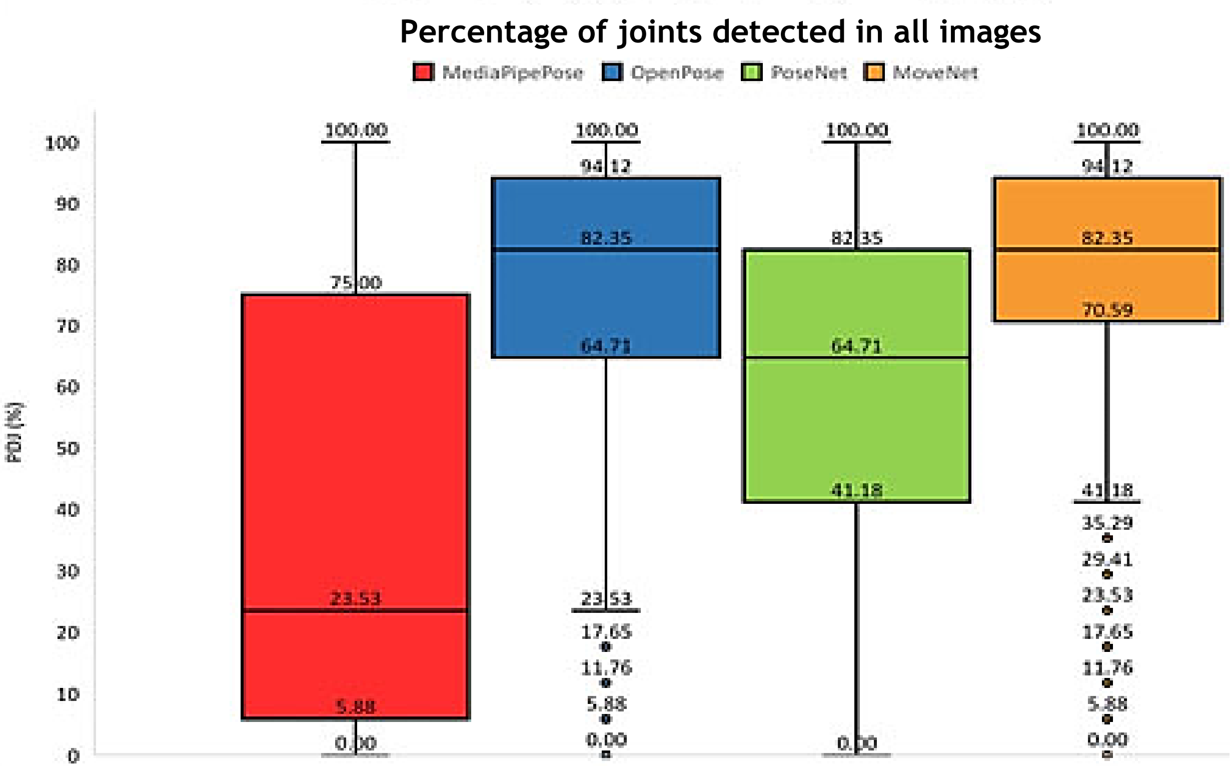

Figure 10*. Box plot of the number of joints detected for the images in the dataset.*

Also, in another study the Mediapipe library trained two models using a subset of the Augmented Reality (AR) and Yoga datasets to compare and evaluate the quality of the BlazePose model against OpenPose. Both models were trained using the common 17-point MS Coco topology, and their performance was assessed using the Percentage of Correct Points formula. On the AR dataset, the BlazePose model performed slightly worse than the OpenPose model. However, in Yoga/Fitness use cases, the BlazePose Full model outperformed OpenPose. Furthermore, the BlazePose model demonstrated significantly faster processing speeds, ranging from 25 to 75 times faster than a 20-core desktop processor on an intermediate-level phone CPU, depending on the desired quality (Table 3).

Table 3. *BlazePose and OpenPose comparison.*

Model	FPS	AR Dataset, PCK@0.2	Yoga Dataset, PCK@0.2
OpenPose (only body)	0.4	**87.8**	83.4
BlazePose Full	10	84.1	**84.5**
BlazePose Lite	**31**[2]	79.6	88.6

4 RESULTS AND DISCUSSION

In recent years, the use of artificial intelligence, machine learning, and deep learning technologies has been increasing in the analysis of athlete movements. These technologies enable a more detailed and accurate analysis of athlete movements. In this context, the deep learning-based MediaPipe Pose Estimation model is used to analyse real-time images and obtain human pose estimation data.

In this study, the accuracy rate of deadlift and squat activities were aimed to predict in the field of sports biomechanics. For this purpose, it was used the MediaPipe Pose Estimation module, Numpy, Pandas, and Scikit-Learn Python libraries to process human pose data from real-time or video and images, calculate the angles, and use machine learning algorithms to select the most suitable classifier model and analyse. In addition, this study is an example of the joint use of the MediaPipe Pose Estimation module and Python libraries for human pose estimation in the field of sports biomechanics.

In this context, our application is a program that calculates the kinematic angles for deadlift and squat sports exercises and records the "up" and "down" movements as a CSV dataset consisting of 7795 rows and 33 columns. The program was trained on four machine learning algorithms and compared with performance measurements such as accuracy, precision, and recall scores. The classification algorithm that gave the best result in the comparison was used to evaluate the accuracy of sports exercises. In addition, the processing and memory performance of the machine learning model created using the kinematic angles was tested. The analysis results show that the machine learning method performs better than the processing and memory performance of the kinematic data. These results are visually presented in Figure 11 and Figure 12.

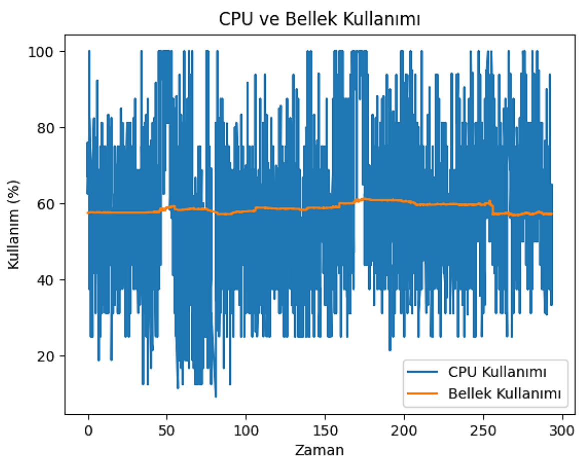

***Figure 11**. Kinematic angle measurements performance.*

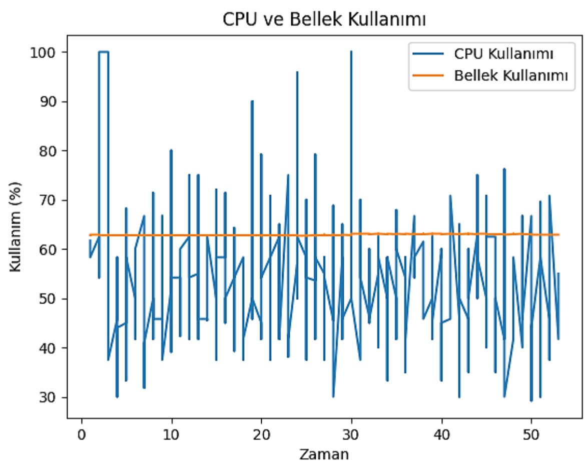

***Figure 12**. Machine learning performance.*

Thanks to the program created in this way, people will be able to do sports activities with the correct posture at home or in any suitable environment without the help of a professional trainer and get the economic benefit of sports activity.

In addition, it has the advantage that the time required for posture analysis is short and the user can receive immediate feedback.

Our system created in this direction will make it easier to exercise without the need for a special trainer and will reduce injuries due to incorrect technique. In addition, by comparing the BlazePose model used by the Mediapipe library with other models, it is suggested that this library can be used effectively in the field of sports biomechanics due to its advantages as opposed to its disadvantages. By moving the system to an Android or IOS application and creating a personalised account to monitor progress, a wider audience be able to be reached in the future. In addition, the emergence of depth sensors in smartphones will also help to improve the accuracy of detecting the pose much more precisely, which will help to create a more accurate posture for comparison.

Conflict of Interest

There is no conflict of interest between the authors.

Authors contributions

[Muhammed Fatih KULUÖZTÜRK]: Played a leading role in identifying the main design and research questions of the article. He was involved in the data collection process and provided management of the analyses. He also played a key role in making sense of the methodology and results.

[Nurettin ACI]: Conducted in-depth studies during the literature review phase and contributed significantly to the organization and style of the manuscript. He led the critical appraisal of the results of the analysis.

Task Sharing:

Design and Concept: [Muhammed Fatih KULUÖZTÜRK]

Data Collection and Analysis: [Muhammed Fatih KULUÖZTÜRK]

Literature Review: [Nurettin ACI]

Writing and Editing: [Nurettin ACI]

Methodology Development: [Muhammed Fatih KULUÖZTÜRK]

Interpretation of Analysis Results: [Nurettin ACI]

Statement of Research and Publication Ethics

The study is complied with research and publication ethics.

REFERENCES

[1] Ö. Çevik, Oğuzhanoglu, N. K., & Gülbas, Z., "Spor Yaralanmaları: Önleme ve Tedavi," *Türk fiziksel tıp ve rehabilitasyon dergisi,* vol. 64, no.3, pp. 266-273, 2018.

[2] I. Shrier, "Strategic Assessment of Risk and Risk Tolerance (StARRT) framework for return-to-play decision-making," *British journal of sports medicine,* vol. 49, no.19, pp. 1311-1315, 2015.

[3] R. Bahr, & Holme, I, "Risk factors for sports injuries-a methodological approach," *British journal of sports medicine,* vol. 37, no. 5, pp. 384-392, 2003.

[4] Y. Çakır, "Biyomekaniksel Analizler," İstanbul Medipol Üniversitesi, Sağlık Bilimleri Enstitüsü, Fizyoterapi ve Rehabilitasyon Anabilim Dalı, 2019.

[5] İ. Sarı, "Biyomekanik ve Fiziksel Performansın Biyomekanik Analizi," *Hacettepe Spor Bilimleri Dergisi,* vol. 25, no. 4, pp. 153-167, 2014.

[6] S. Mülayim, "Spor Biyomekaniği," *Ankara Üniversitesi Beden Eğitimi ve Spor Yüksekokulu Dergisi,* vol. 19, no. 3, pp. 183-192, 2019.

[7] Y. Güneri, "Biyomekanik, sağlık bilimleri ve spor bilimleri açısından önemi," *Türkiye Klinikleri Journal of Sports Sciences,* vol. 11, no.1, pp. 16-22, 2019.

[8] D. A. Winter, "Biomechanics and motor control of human movement," *John Wiley & Sons,* 2019.

[9] V. M. Zatsiorsky, & Seluyanov, V. N, "The mass and inertia characteristics of the main segments of the human body," *Biomechanics VIII-B,* pp. 115-122, 1983.

[10] F. Muradlı, "Derin Öğrenme Kullanılarak Görüntülerden İnsan Duruş Tespiti," *Sakarya Üniversitesi Fen Bilimleri Enstitüsü Yüksek Lisans Tezi,* 2021.

[11] A. Özdemir, & Özdemir, A, "MediaPipe Pose ile Evde Egzersiz Yaparken Duruş Tespiti ve Rehberlik Etme,"*International Journal of Informatics Technologies,* vol. 14. no. 2, pp. 123-132, 2021.

[12] M. Dersuneli, T. Gündüz, and Y. Kutlu, "Bul-Tak Oyuncağı Şekillerinin Klasik Görüntü İşleme ve Derin Öğrenme Yöntemleri ile Tespiti,"*Bitlis Eren Üniversitesi Fen Bilimleri Dergisi,* vol. 10, no. 4, pp. 1290-1303, 2021.

[13] L. Deng and D. Yu, "Deep learning: methods and applications,"*Foundations and trends® in signal processing,* vol. 7, no. 3–4, pp. 197-387, 2014.

[14] A. Yaman, S. Abdulkadir, B. Ümit, and E. Sami, "Deep learning-based face liveness detection in videos," in *Proceedings of the IEEEInternational Artificial Intelligence and Data Processing Symposium (IDAP), Malatya, Turkey*, 2017, pp. 16-17.

[15] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature,* vol. 521, no. 7553, pp. 436-444, 2015.

[16] C.-H. Chen and D. Ramanan, "3d human pose estimation= 2d pose estimation+ matching," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 7035-7043.

[17] C. Zheng *et al.*, "Deep Learning-Based Human Pose Estimation: A Survey," *ACM Computing Surveys,* 2019.

[18] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, "Realtime multi-person 2d pose estimation using part affinity fields," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 7291-7299.

[19] M. H. BC, R. Prathibha, and S. Kumari, "YOGA AI TRAINER USING TENSORFLOW. JS POSENET."

[20] R. Guler, N. Neverova, and I. DensePose, "Dense human pose estimation in the wild," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA*, 2018, pp. 18-23.

[21] H.-S. Fang *et al.*, "Alphapose: Whole-body regional multi-person pose estimation and tracking in real-time," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 2022.

[22] K. Sun, B. Xiao, D. Liu, and J. Wang, "Deep high-resolution representation learning for human pose estimation," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 5693-5703.

[23] S. Garg, A. Saxena, and R. Gupta, "Yoga pose classification: a CNN and MediaPipe inspired deep learning approach for real-world application," *Journal of Ambient Intelligence and Humanized Computing,* pp. 1-12, 2022.

[24] H. H. Pham, H. Salmane, L. Khoudour, A. Crouzil, S. A. Velastin, and P. Zegers, "A unified deep framework for joint 3d pose estimation and action recognition from a single rgb camera," *Sensors,* vol. 20, no. 7, p. 1825, 2020.

[25] A. Howard *et al.*, "Searching for mobilenetv3," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 1314-1324.

[26] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, "Realtime multi-person 2d pose estimation using part affinity fields," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1302-1310.

[27] K. Su, D. Yu, Z. Xu, X. Geng, and C. Wang, "Multi-person pose estimation with enhanced channel-wise and spatial information," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 5674-5682.

[28] M. Orescanin, L. N. Smith, S. Sahu, P. Goyal, and S. R. Chhetri, "Editorial: Deep learning with limited labeled data for vision, audio, and text," (in English), *Frontiers in Artificial Intelligence,* Editorial vol. 6, 2023-June-13 2023, doi: 10.3389/frai.2023.1213419.

[29] A. C. o. S. Medicine, *ACSM's guidelines for exercise testing and prescription*. Lippincott williams & wilkins, 2013.

[30] R. F. Escamilla, G. S. Fleisig, T. M. Lowry, S. W. Barrentine, and J. R. Andrews, "A three-dimensional biomechanical analysis of the squat during varying stance widths," *Medicine and science in sports and exercise,* vol. 33, no. 6, pp. 984-998, 2001.

[31] R. F. Escamilla, A. C. Francisco, A. V. Kayes, K. P. Speer, and C. T. Moorman 3rd, "An electromyographic analysis of sumo and conventional style deadlifts," *Medicine and science in sports and exercise,* vol. 34, no. 4, pp. 682-688, 2002.

[32] Z. Cömert and A. KOCAMAZ, "A study of artificial neural network training algorithms for classification of cardiotocography signals," *Bitlis Eren University journal of science and technology,* vol. 7, no. 2, pp. 93-103, 2017.

[33] M. Pilgrim, "Serializing Python Objects," in *Dive Into Python 3*: Springer, 2009, pp. 205-223.

[34] N. Çetin, Biyomekanik, *Setma Baskı Ankara, vol.* 1, pp.4-41, 1997.

[35] C. Açıkada, H. Demirel, Biyomekanik ve Hareket Bilgisi, *AÜAÖFEskişehir,* pp. 15,1993.

[36] G. Yavuzer, The use of computerized gait analysis in the assessment of neuromusculoskeletal disorders, *Journal Of Physical Medicine And Rehabilitation Sciences,* vol. 10, no: 2, pp. 043-045, 2007.

[37] DA Winter, Biomechanics And Motor Control Of Human Movement, *2nd Edition, John Wiley & Sons Canada,* 1990.

[38] W. Braüne, O. Fischer, The Human Gait (Ceviri: Maquet P Furlong R), *Springer-Verlag Heidelberg Almanya,* 1987.

[39] YI Abdel-Aziz, HM Karara, Direct linear transformation from comparator coordinates ınto object space coordinates in close-range photogrammetry proceedings of the asp/ui, *Symposium on Close-range Photogrammetry American Society of Photogrammetry,* Falls Church VA, s 1-18, 1971.

[40] E. Civek, Comparison of kinematic results between Metu-kiss & Ankara University-vicon gait analysis systems, *Y Lisans Tezi, Odtü Makina Mühendisliği Bölümü*, 2006.

[41] M. Orescanin, L. N. Smith, S. Sahu, P. Goyal, and S. R. Chhetri, "Editorial: Deep learning with limited labeled data for vision, audio, and text," (in English), *Frontiers in Artificial Intelligence,* Editorial vol. 6, 2023-June-13 2023, doi: 10.3389/frai.2023.1213419.

[42] Ö. Çokluk, "Lojistik regresyon analizi: Kavram ve uygulama.", *Kuram ve uygulamada eğitim bilimleri*, vol.10, no: 3, pp. 1357-1407, 2010.

[43] C.F. İşçen et al, "Su Kalitesi Değişimine Etki Eden Değişkenlerin Lojistik Regresyon, Lojistik-Ridge Ve Lojistik-Lasso Yöntemleri İle Tespiti.", *Biyoloji Bilimleri Araştırma Dergisi*, vol. 14, no: 1, pp. 1-12.

[44] F. Erdem, et al., "Rastgele orman yöntemi kullanılarak kıyı çizgisi çıkarımı İstanbul örneği.", *Geomatik 3.2*,pp. 100-107, 2018.

[45] N. Chakrabarty, et al., "Flight arrival delay prediction using gradient boosting classifier.", *Emerging Technologies in Data Mining and Information Security: Proceedings of IEMIS* 2018, Volume 2. Springer Singapore, 2019, 2018.

[46] J.-L. Chung, L.-Y. Ong, and M.-C. Leow, "Comparative Analysis of Skeleton-Based Human Pose Estimation," *Future Internet*, vol. 14, no. 12, p. 380, 2022. [Online]. Available: https://www.mdpi.com/1999-5903/14/12/380.

[47] C. Lugaresi *et al.*, "Mediapipe: A framework for building perception pipelines," *arXiv preprint arXiv:1906.08172,* 2019.

[48] G. Papandreou *et al.*, "Towards accurate multi-person pose estimation in the wild," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4903-4911.

[49] B. Jo and S. Kim, "Comparative analysis of OpenPose, PoseNet, and MoveNet models for pose estimation in mobile devices," *Traitement du Signal,* vol. 39, no. 1, p. 119, 2022.

[50] F. Duman, T. D. İpek, and M. Saraçlar, "Unsupervised Discovery of Fingerspelled Letters in Sign Language Videos," in *2021 29th Signal Processing and Communications Applications Conference (SIU)*, 2021: IEEE, pp. 1-4.

[51] M. Mundt, Z. Born, M. Goldacre, and J. Alderson, "Estimating Ground Reaction Forces from Two-Dimensional Pose Data: A Biomechanics-Based Comparison of AlphaPose, BlazePose, and OpenPose," *Sensors,* vol. 23, no. 1, p. 78, 2022.

[52] L. Song, G. Yu, J. Yuan, and Z. Liu, "Human pose estimation and its application to action recognition: A survey," *Journal of Visual Communication and Image Representation,* vol. 76, p. 103055, 2021.

[53] L. Pishchulin *et al.*, "Deepcut: Joint subset partition and labeling for multi person pose estimation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 4929-4937.

[54] T.-Y. Lin *et al.*, "Microsoft coco: Common objects in context," in *Computer Vision-ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13*, 2014: Springer, pp. 740-755.

[55] D. C. Luvizon, D. Picard, and H. Tabia, "2d/3d pose estimation and action recognition using multitask deep learning," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 5137-5146.