Pose Detection using OpenCV and Media Pipe

Deepak Rai
School of CSE and Technology
Bennett University
Greater Noida, India
0000-0002-7635-3801

Anjali
School of CSE and Technology
Bennett University
Greater Noida, India
e21cseu0182@bennett.edu.in

Ankit Kumar
School of CSE and Technology
Bennett University
Greater Noida, India
e21cseu0185@bennett.edu.in

Anjali Baghel
School of CSE and Technology
Bennett University
Greater Noida, India
e21cseu0183@bennett.edu.in

***Abstract*—Whether doing out at home or at a gym, users can follow the computer vision-based program recommended in this study report. The software estimates user-body landmarks using OpenCV and Media Pipe to compute body angles and other numerical data in real time and determine the best posture for each workout. If a user assumes the wrong posture, the application provides guidance on how to fix it; if a user assumes the right posture, the application provides guidance and increases the quantity of exercises. The model will give you visual feedback and shows you how to adjust your posture on the screen so we can do it more accurately. This helps people prevent from any injury while doing any wrong exercise. Basically, this model will work as an instructor for the user and maintain their posture according to the exercise and prevent injury. All things considered, pose estimation whether it is for posture detection, hand tracking, holistic tracking and many more are very useful tools which help to meet individuals to their fitness goals. The application is also used in healthcare like hand gesture tracking such as in grabbling exercise for tracking patients hand movement by providing real time feedback.**

***Index Terms*—Computer Vision, posture correction, OpenCV, Media Pipe, Exercise guidance, Hand gesture, Real-time feedback**

I. Introduction

As an individual's survival, a healthy lifestyle plays an important part in our life. As in past year a pandemic COVID-19 has forced people to stay at their homes and the trend of online has been increased whether it is a work from home or doing exercise from the YouTube videos [1]. On the other hand, an individual which is new to exercise or facing difficulty in doing exercise with the help of digital videos needs proper instruction or a trainer to help. Without instruction it can lead to poor posture and can cause injuries.

To solve this problem, researchers have developed multiple models that will instruct users to how to do exercise and the system will correct the posture of the users while it is capturing real time performance of the user and provide real time instruction or feedback. It uses coordinates x, y and z to evaluate the body posture. It provides real- time feedback so the user can correct its posture.

The system not only detects posture for exercise, but it can be useful for many industries such as in health care, gaming and entertainment, real estate, education, automotive, manufacturing and many more [2]–[4]. Exercise daily helps individuals to recover from depression, but correct posture is always important while doing exercise. A bad posture can be a cause of back pain or any injury. An instructor always needs to instruct at some steps while doing the exercise so it can adjust the posture as per the exercise. However, everyone does not have access to a personal trainer. In this scenario this system will be helpful as their personal trainer.

This system is based on machine learning and artificial intelligence. In recent years artificial intelligence and machine learning developed with neural networks. Also, deep learning models can perform a wide range of tasks. The purpose of this study is to discover how effective a pose estimation method can be in various industries such as in detecting poses. We use OpenCV for real time webcam feed and it detects the body landmark for estimation the posture. It can also calculate the angels between different landmarks. Finally, using artificial intelligence and machine learning techniques the system provides instant feedback and correct the posture as needed. The study may help in improving the accuracy and efficiency of the system.

II. Literature Review

Studies on posture recognition have gained popularity. Deep learning-based posture estimation, sensor-based posture recognition, and wearable device research are all related fields of study. Here are some examples of research, specifically. identifying bad posture from an image using an Arduino-based postural pressure measuring device, then suggesting an activity to improve it. Using a human skeleton model and a CNN-based deep learning model, detect incorrect yoga poses and offer feedback. Give input for improvement using Arduino-controlled vibrations in clothing and recommend a smartphone app that can monitor warning signals through Bluetooth. Apply a recognition algorithm to wearable technology-collected data on a human motion to wirelessly recognize poses [5].

Many fields use posture recognition in addition to healthcare, such as in devices that detect elderly fallers and offer monitoring information. People can play virtual games by identifying hand movements in games, while governments may stop crime by seeing aggressive behavior in CCTV footage. Due to learning through a significant amount of data for posture categorization and determination, however, current deep learning-based research takes a long time for posture identification, estimation, and feedback. Additionally, users need separate equipment for studies based on wearable technology or sensor-based data-gathering systems [6].

So, in this study, we propose a posture correction program that can shorten the time needed to analyze posture and not cost extra money for analysis equipment. Only the most popular exercises, the squat, and push-up, are covered in the posture correction suggestions displayed on the screen. To assist the user in exercising with the proper posture, the number of exercises is delivered by voice and text as the activity is being performed in real time.

A. Physical fitness and its importance

The most widely reported comparison of kids in the United States and Europe was made using standardized fitness centers [7]. 42% of American teenagers were physically fit to the same extent as 92% of their European counterparts. Nearly half of American children are not in good physical shape, and many are becoming mushy. Every person's softness will result in the depletion and destruction of the country's vitality, vigor, and spirit. The hostile media condemned this technique and concentrated on maintaining low fitness ratings for schoolchildren. The adage used by detractors demonstrates how awkward, uneasy, and meaningless fitness testing felt. Researchers recommended rewarding physical fitness and viewing it as a form of academic achievement to increase its psychological significance [8].

Then, for pupils with all levels of potential, a performance indicator and a reward tag can serve as a motivating factors. The cognitive assessment theory keeps young people inspired and motivated to engage in regular physical activity. We are aware that the generation before us was more physically fit. The current generation is still working to increase both their physical and mental health, nevertheless. A better diet with more nutrients and regular exercise will improve one's fitness. The recent resurgence of both proponents and opponents of the necessity of physical fitness and possible testing raised considerable public awareness of the emergency in the neighborhood.

The author has outlined the key ideas for a thorough discussion of physical fitness in this. The idea focuses on the evaluation and how it relates to fitness for health and varying levels of health with age. Early physical assessment debates questioned the value of fitness assessments. However, there is a consensus among public health and medical professionals as to the value of the evaluation as well as previously completed fitness testing in academic school physical education programs. The author also emphasizes the subject's aerobic fitness, which may be age-independent. Another finding was that the amount of oxygen the body uses and absorbs directly correlates with aerobic capacity [8]. These ten fundamental ideas of the fitness assessment can be used to determine an individual's strength. For optimal coordination and fitness programming, the significance of fitness evaluation and its integration into daily activities is taken into account.

B. Python and its use

The author describes an image processing course that should be broken down into modules like Python fundamentals, image processing and its principles, digital and contemporary techniques of image processing, Image, and video acquisition, detection of the main object of the model in motion, Image and pixel formation, borders, image segmentation, and optic instruments. Students are more engaged in the learning process because of the use of PBL and GUI (Graphical User Interface) representations. One could ask questions like "What," "How," and "For what?" as they are closely related to the problem viewpoints.

Teaching in particular fields makes it possible to build problem-solving methods and approaches as well as specific talents. Part of the teaching process includes organizing workshops in class and setting up a conducive environment for the new technology. One will be forced to quickly adjust to emerging technologies [8].

C. Image and Video Processing

The author used pre-trained models for 19 different types of objects on the "travel" data for the PASCAL Visual Object Classes Challenge 2009 [9]. The robust approach applied provides easy object detection for static images, further increasing the efficiency of video detection. The model was trained by selecting one frame per second and running an object detector for each of them to collect the scores. The probability distribution is calculated using Platt scaling and map, then calibrated probabilities. The author applied four different approaches: windowing, Pos, which is a parser of tagging, Parsing, and Parsing II. The model works on rules and neural networks [8]. To obtain an effective test for object recognition, the data set must be split up into numerous distinct training sets. The set recognizer is not the only restriction on activity detection. It added text using only the datasets we had and trained the model using those datasets. Objects in the images that were not specified in the data sets could not be detected by the model.

The author outlines the ergonomic assessment of posture, which finally aids the workers in realizing their poor postures. This will lessen the impact of back pain and ache on the employees and help build workstations that are appropriate for them. The fundamental purpose of ergonomic approaches is to test or evaluate workplace risk. Varied specialists have different viewpoints on this technique for evaluating posture. The MOCAP (Motion Capture) apparatus analyses and distinguishes between different postures [10]. Human data can be used to train the AI model, and humans can help with its

predictions. Humans are a significant resource for businesses, and if they are physically capable of doing the job, they can be kept with good efficiency.

The study intends to evaluate the benefits of better posture as well as the drawbacks of bending the back and many other structures by combining the usage of MOCAP and AI technologies. The evaluation approach used by OWAS (Ovako Working Posture Assessment System) is arbitrary. To assist the observer and make more reliable classification predictions, it develops an AI-based approach. The development of techniques for measuring objects placed in front of the optic detector was the main emphasis of the paper. Calculating the distance between them is made simpler if both the actual size and the size that was captured in the image are known. However, the precise size is frequently omitted. As a result, whenever the real size of an object is not stated, the paper provides a way to determine the distance from the object. To gather more precise real-time data, the demonstration camera needs to be quick. The single-shot and multiple-shot possibilities depended on how accurate and precise the desired solution was. The fundamental ideas behind how a camera functions are used to illustrate the distance between the camera and the item. To arrive at a calculated solution, the author additionally applied the fundamentals of optics and the laws of reflection [11] [12].

III. Study Area and Data Resources

A. Computer vision

The aim of the computer vision within artificial intelligence and computer science is to make it possible for machines to interpret digital pictures and images from outside the world. It involves models and algorithms which can automatically analyze, assess and interpret the visual data taken from the actual world. Computer vision involves many applications such as image and video recognition, facial recognition in photos and in videos, object tracking and object identification and analysis of medical pictures. It is also important for machines to sense and engage with the surrounding [13].

B. OpenCV

Real time webcam feed, image processing and video processing is helped by OpenCV. OpenCV is an open-source library for computer vision and machine learning. It also provides the standard framework for computer vision applications. It was created by Intel 1999, and Willow Garage and later it was supported by Itseez. OpenCV was written in C++. It includes interfaces in Python, java and in MATLAB. It has over 2,500 effective algorithms which are applied in computer vision applications. It is also used in broad range of applications which includes robotics, augmented reality, medical imagining and even in film productions.

As shown in Fig. 1, in exercise that involves posture detection, OpenCV can be utilized by incorporating its image processing and computer vision functionalities. OpenCV can be used to carry out the following tasks in particular:

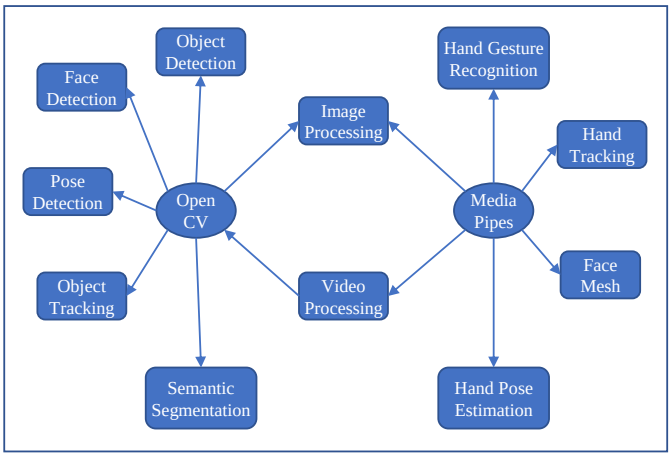

Fig. 1. Ready to use models for different tasks.

- Image Acquisition: With an OpenCV, one can acquire pictures or even video frames from cameras and videos respectively. The user may use this information to live stream his/her posture and exercise performances.
- Pre-processing: Before processing captured images, noise reduction, enhancement of contrast and lighting improvement should be performed. This operation can apply several image processing methods available from OpenCV.
- Feature extraction: OpenCV can then be applied from the pre-processed images to retrieve important information, e.g., body landmarks and joint locational spaces. For instance, there are techniques such as matching of features, detection of edges and outlines, line drawing, etc.
- Pose estimation: With captured features using OpenCV, a person can determine the body pose and angle of joints of the user. Some examples of machine learning algorithms include ANN and SVM.
- Posture Evaluation: With the predicted pose and joint angle, OpenCV can be used to estimate posture of the user. Comparing the user's posture and ideal exercise posture, this could be achieved.

C. Media Pipe

Media Pipe, a cross-platform open-source framework for building multimodal machine learning (MMML). Google developed it and launched it to operation for the first time in 2019. This framework allows one to process different kinds of data such as images, videos, and audio, and then apply machine learning algorithms on them to obtain useful information. As shown in Fig. 1, it has some ready to use models for tasks like object detection, face detection, pose estimation, and hand tracking as well. Moreover, through this framework, Media Pose provides developers with APIs that can help on in changing as well as extending the framework according to needs. The media Pipe Stance is an out of the box model that applies machine learning to video frames for posture detection in exercises.

For instance, a position estimation model with the aspect of posture detection related to exercises is offered by Media Pipe Pose. The model operates on essential human markers, like joints and body parts, and finds the three-dimensional coordinates of these reference points. In such a way, sensitive spots are monitored over time, and it is easy for the device to determine posture of user at work out period as well as find deviation from standardized position.

The steps listed below can be used to include Media Pipe Pose into a posture detection system for exercise:

- Capture video input: Take a video using camera or a smartphone when the user is doing an exercise.
- Pre-process the video: To improve the scene quality as well as remove unwanted noises, the video frames must be processed in earlier stages. They can use processes like filtering, cropping, or resizing.
- Apply Media Pipe Pose: Media Pipe Pose should then process the pre-processed video frames, providing key points for every frame.
- Calculate angles: The system could subsequently use the key data gathered from the Media Pipe Pose to compute the angles formed by various body parts and ensure that they conform with the right posture requirements for a particular exercise.
- Display feedback: The system should display the necessary instructions on the screen to enable the user to correct their posture if the user has got the wrong position. If the user correctly positions him/herself, the system will be able to verbally generate the additional workouts and project them onto the screen.

IV. Methodology

A. System Architecture

The construction of our OpnCV and MdiaPip basd solution for postur dtction during xrcis is composed of numerous components.

The first part is the camera module that records live footage of the user performing different exercises. The latter module receives a video stream from the Open CV module. The OpenCV module detects the location of the human body in the video stream with the help of Python package embedded into Media Pipe.

The other element is the kNN algorithm module that differentiates between postures like the palm and the fist. The coordinates of different postures are used as learning data to train the kNN algorithm. The kNN algorithm module can predict the correct posture with a high degree of accuracy based on the coordinates retrieved by OpenCV module.

The next element is the calculation module that uses the coordinates provided by OpenCV module to determine the angles between the body's main articulations. Then, these angles are used to check whether the user is performing the exercise correctly. The system then provides details on how they can correct their stance if their practicing of the exercise is not proper. As shown in Fig. 2, there is an output module

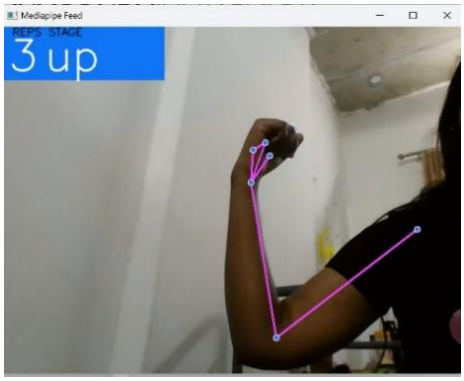

Fig. 2. Feedback module showing the results of exercise.

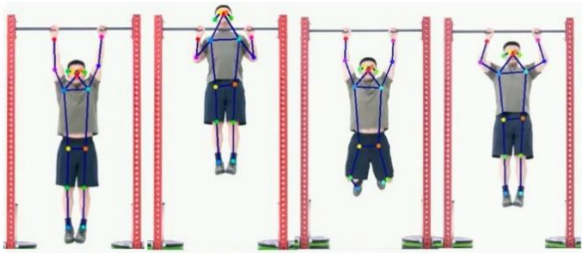

Fig. 3. Snapshots of webcam recorded live video.

(feedback module), which presents the user with the results of exercise. The user's proper posture throughout is shown as well as the total number of reps performed, and the calories expended. Our system proves accurate in the overall detection and evaluation of the user's postures during exercises by fusing computer vision techniques with machine learning.

B. Implementation Details

The setup for the posture detection system involves some phases when it comes to exercises using OpenCV or Media Pipe. As shown in Fig. 3, the system makes use of the OpenCV Python library for accessing the user's webcam to record live video. A Media Pipe module is subsequently able to access the primary joints of the human body through provision of joint coordinate information.

Fundamental trigonometry is then used to determine the angle between each joint. The computer records this data to determine if they are executing the exercise appropriately. For example, measuring the angle at the shoulder, elbow, and wrist while performing a bicep curl. The number increases by one for any angle ranging from 30 to 160 during that specific exercise.

Not only does the technology identify the posture but also instructs one's movements during an exercise session. There are three training routines within the system, which focus on pulling exercises, pushing exercises and leg exercises. Every single plan has exercises accompanied with directions on how to execute the work properly.

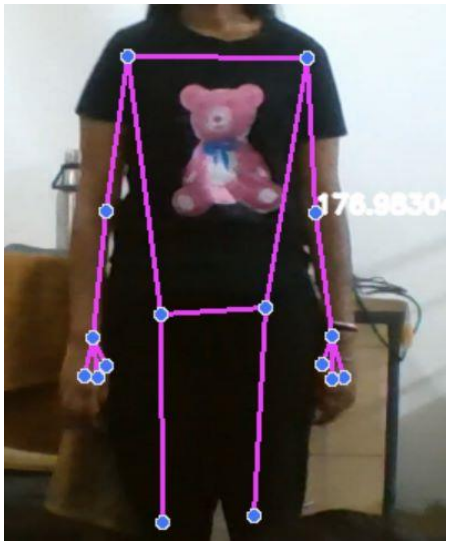

Fig. 4. Angle measurement between joints using fundamental trigonometry.

To further enhance the system's performance, a k-NN approach is used to develop more accurate motion detection. This algorithm has been educated to distinguish between the contour of a palm compared to that of a fist. It determines the label worthiness by comparing, at a real time, the location of the fist and paw shapes towards the reading data location. It has an interactive game that allows consumers to have a playful means of burning calories as well. The player should hit the target accurately and quickly to progress with the game. When the target is hit properly, the score line and the pace increase. In general, real-time video processing, angle computation, motion guide and the k-NN motion detection algorithms are used in a workout posture detection system based on OpenCV and Media Pipe.

V. Study of Proposed Algorithm

The system waits for the user to initiate it, then it grabs information from the camera device. He checks the frame's content and distinguishes individual body parts using the data Media Pipe Library [14]. Motion tracking a computer vision technology is used for tracking the movements of a person or object with positional estimation. Being presently applied to augmented reality, animation, gaming, and robotics, pose estimation is a part of the effective and popular technologies. Some of the products in the market are available to perform pose estimation.

The following steps are followed for taking input from media devices for making system ready:

- Ensure optic sensor device is available for access from the system for capturing.
- Video Capture(X):
  - X is the Vice code for your input-taking device.
  - **If** (Webcam is open)
  - **Then** Read the input.
  - Return the Image to the screen (frame).
  - Release the Webcam after taking input.
- Make detection from the feed on the criteria $Min_detection$ and $Min_tracking$.
- Highlight the points on Body Parts.
- Once the user has finished, close the system.

The overall posture tracking process is shown in Fig. 5.

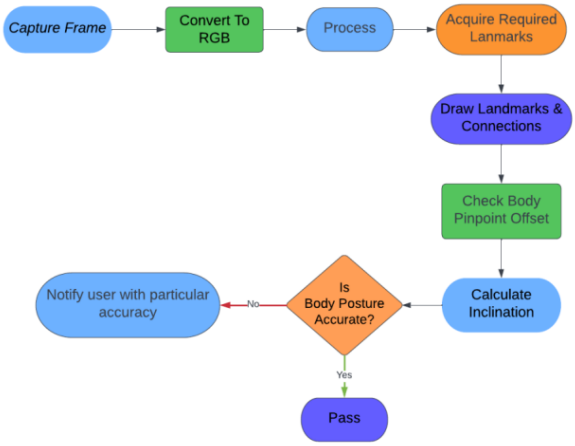

Fig. 5. Tracking posture.

VI. Results and Discussion

In this research, we applied the OpenCV and MediaPipe tools to develop a posture recognition system for physical exercises. During exercise time, this system can also track and determine if the users' posture is correct and notify them immediately. On a sample dataset of various fitness activities (pull-ups, push-ups and squats), we conducted studies to determine the accuracy of the system. Data was collected from each of the 10 volunteers who performed each exercise 10 times. We manually labelled each individual exercise with its corresponding position.

The k-NN algorithm is trained using these posture coordinators as a leaning data set. This means that it is possible to accurately predict the right posture using joint coordinates obtained from the OpenCV module in the k-NN algorithm module.

The second is the computation module of it that finds out the angles between the major joints of the body upon using the coordinates of it generated by the OpenCV module. The angles are then used in determining whether the person is doing it right. In addition, the system directs users as to what to do if they are not doing the exercise correctly about posture. Finally, there is the feedback section that displays the results of what the user has done. The total number of repetitions done is showcased in this segment. However, we also assessed the system's efficiency based on its ability to follow exercise repetition. Ninety two percent (pull-up count). The accuracy of counting for push-ups and squats, however, amounted to only 83% and 78%. Another factor that might explain the low

accuracy of counting repetitions for push-ups and squats is the variety of users' postures during activities. Unlike pull-ups, push-ups and squats can be performed with several variations thus making recognition of correct form for every repetition difficult.

The overall trial results indicate that a proposed posture detection system for physical activities works. The system is precise.

VII. Conclusion and Future Work

This paper have demonstrated the way of implementing OpenCV and Media Pipe into a system designed for recognizing posture in an exercise. The aim of this approach is to enable users to maintain correct shape during exercise for the prevention of injuries and maximizing the benefits of the workout. First, we gained access to the computer's webcam using OpenCV so as to record the live video of the user. Subsequently, Media pipe is applied to recognize the key joints of the body and determine the angle between them.

Several features which we have included in use have enhanced the user's training experience such as display of number of reps and calories burned besides instruction in each exercise. Research opportunities focusing on posture recognition for exercise-related applications through OpenCV and Media Pipe. However, the accuracy of the first pose estimation model could be improved. The present approach accurately identifies the joints but there is room for improvement on how joint angles and repetitions are determined.

Secondly, the scope and variety of exercises that the system can accommodate needs to be expanded. Instead of just concentrating on bicep curls, pushups and squats one should also include other types like Lunges sit ups Jumping Jacks etc. Finally, the platform is easy to connect with other fitness monitoring gadgets or applications that provide users complete information on their progress in developing physical fitness.

It also has the capacity to further development in order to serve other users' needs, including special-needs individuals and elderly people. Overall, much can still be done to enhance this field of exercise-based assessment systems and future work in the area may help further increase accuracy or usefulness. In future some face super resolution techniques [15]–[23] may also be integrated with the proposed model to get high quality resultant image even in case of low quality or low resolution recorded images.

References

[1] J. Wilke, L. Mohr, A. S. Tenforde, P. Edouard, C. Fossati, M. González-Gross, C. S. Ramirez, F. Laiño, B. Tan, J. D. Pillay *et al.*, "Restrictercise! preferences regarding digital home training programs during confinements associated with the covid-19 pandemic," *International journal of environmental research and public health*, vol. 17, no. 18, p. 6515, 2020.

[2] M. Chen and A. Hernández, "Towards an explainable model for sepsis detection based on sensitivity analysis," *IRBM*, vol. 43, no. 1, pp. 75–86, 2022.

[3] L. Zeng, H. Zhang, Q. Han, Y. Tang, L. Ye, Y. Wu, and H. Zu, "An lstm-based driving operation suggestion method for riding comfort-oriented critical zone," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–17, 2023.

[4] V. Gupta, N. K. Saxena, A. Kanungo, P. Kumar, and S. Diwania, "Pca as an effective tool for the detection of r-peaks in an ecg signal processing," *International Journal of System Assurance Engineering and Management*, vol. 13, no. 5, pp. 2391–2403, 2022.

[5] J.-S. Han, C.-I. Lee, Y.-H. Youn, and S.-J. Kim, "A study on real-time hand gesture recognition technology by machine learning-based mediapipe," *Journal of System and Management Sciences*, vol. 12, no. 2, pp. 462–476, 2022.

[6] Y. Kwon and D. Kim, "Real-time workout posture correction using opencv and mediapipe," *The Journal of Korean Institute of Information Technology*, vol. 20, no. 1, pp. 199–208, 2022.

[7] S. Kale, N. Kulkarni, S. Kumbhkarn, A. Khuspe, and S. Kharde, "Posture detection and comparison of different physical exercises based on deep learning using media pipe, opencv," *International Journal of Scientific Research in Engineering and Management (IJSREM)*, 2023.

[8] ——, "Posture detection and comparison of different physical exercises based on deep learning using media pipe, opencv," *International Journal of Scientific Research in Engineering and Management*, 2023.

[9] T. S. Motwani and R. J. Mooney, "Improving video activity recognition using object recognition and text mining," in *ECAI 2012*. IOS Press, 2012, pp. 600–605.

[10] V. Igelmo, A. Syberfeldt, D. Högberg, F. Rivera, and E. Luque, "Aiding observational ergonomic evaluation methods using mocap systems supported by ai-based posture recognition," *Adv. Transdiscipl. Eng*, vol. 11, pp. 419–429, 2020.

[11] C. Ma, W. Li, R. Gravina, and G. Fortino, "Posture detection based on smart cushion for wheelchair users," *Sensors*, vol. 17, no. 4, p. 719, 2017.

[12] J. Li, C. Xu, Z. Chen, S. Bian, L. Yang, and C. Lu, "Hybrik: A hybrid analytical-neural inverse kinematics solution for 3d human pose and shape estimation," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 3383–3393.

[13] N. Sarafianos, B. Boteanu, B. Ionescu, and I. A. Kakadiaris, "3d human pose estimation: A review of the literature and analysis of covariates," *Computer Vision and Image Understanding*, vol. 152, pp. 1–20, 2016.

[14] J. Wu and N. Dahnoun, "A health monitoring system with posture estimation and heart rate detection based on millimeter-wave radar," *Microprocessors and Microsystems*, vol. 94, p. 104670, 2022.

[15] D. Rai and S. S. Rajput, "Robust face hallucination algorithm using motion blur embedded nearest proximate patch representation," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, Art no. 5002510, 2023.

[16] S. S. Rajput, D. Rai, and B. Kumar, "Oeinr-rfh: Outlier elimination based iterative neighbor representation for robust face hallucination," *Expert Systems with Applications*, vol. 237 Part C, Art no. 121553, 2024.

[17] D. Rai and S. S. Rajput, "Low-light robust face image super-resolution via neuro-fuzzy inferencing-based locality constrained representation," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, Art no. 5015911, 2023.

[18] A. S. Tomar, K. Arya, and S. S. Rajput, "Attentive exfeat based deep generative adversarial network for noise robust face super-resolution," *Pattern Recognition Letters*, vol. 169, pp. 58–66, 2023.

[19] D. Rai and S. S. Rajput, "A new face reconstruction technique for noisy low-resolution images using regression learning," *Computers and Electrical Engineering*, vol. 107, p. 108642, 2023.

[20] A. S. Tomar, K. Arya, and S. S. Rajput, "Deep hyfeat based attention in attention model for face super-resolution," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–11, 2023.

[21] D. Rai and S. S. Rajput, "Estimation of darkness factor from low-light images based on adaptive neuro-fuzzy inferencing technique," in *2022 International Conference on Engineering and Emerging Technologies (ICEET)*, 2022, pp. 1–6.

[22] A. S. Tomar, K. Arya, and S. S. Rajput, "Noise robust face super-resolution via learning of spatial attentive features," *Multimedia Tools and Applications*, pp. 1–17, 2023.

[23] D. Rai and S. S. Rajput, "A gaussian process regression-based noise level prediction technique for assisting image super-resolution," in *International Conference on Intelligent Technologies*. Springer, 2022, pp. 111–119.