INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact Factor: 6.078
(Volume 11, Issue 3 - V11I3-1257)
Available online at: https://www.ijariit.com

Real-Time Bicep Curl Tracking and Pose Detection Using OpenCV and Media-Pipe

Shiv Arora
shiv.arora.cse.2021@miet.ac.in
Meerut Institute of Engineering and Technology, Meerut, Uttar Pradesh

Drishti Sharma
drishti.sharma.cse.2021@miet.ac.in
Meerut Institute of Engineering and Technology, Meerut, Uttar Pradesh

Shubh Mudgal
shubhmudgal369@gmail.com
Meerut Institute of Engineering and Technology, Meerut, Uttar Pradesh

Sudhanshu Chaudhary
sudhanshuror1@gmail.com
Meerut Institute of Engineering and Technology, Meerut, Uttar Pradesh

ABSTRACT

Human pose estimation is crucial for enabling real-time monitoring of physical exercise via the analysis of movement and orientation of the body. However, existing pose estimation techniques are prone to major flaws such as mislocalization of joints, occlusion issues, and misrecognition of repetition of exercises. Such flaws undermine the efficacy and reliability of fitness tracking systems. In an attempt to address these flaws, the present study proposes a real-time bicep curl tracking system based on OpenCV and MediaPipe. The proposed system is designed to accurately estimate human pose, calculate joint angles, and provide automatic user feedback. One of the system's basic features is that it uses a state-based repetition counter, which improves accuracy in repetition detection by eliminating false positives caused by minor landmark placement variation. The system only detects repetitions when the form is proper and the range of motion is full. In addition to providing real-time feedback on posture changes and detecting improper exercise form, the system effectively eliminates the risk of injury during the execution of strength training exercises. It provides real-time feedback on posture changes and incorrect exercise form. Through empirical analysis, the system proposed has a remarkable accuracy of 96% in quantifying repetitions, which outperforms the performance of the traditional pose tracking models. The high accuracy verifies the system's robustness as well as its usability in real-world fitness applications. Findings indicate that the integration of AI-driven pose estimation and feedback mechanisms can potentially make personalized fitness training much more effective. Together with real-time correction and individualized data, these technologies can improve efficiency in training while motivating safer training habits. This work contributes to the growing field of AI-driven health and fitness technology and opens the door to more advanced and responsive physical activity monitoring devices

Keywords: *Human Pose Estimation, AI-Based Fitness, Bicep Curl Tracker*

1. INTRODUCTION

Correct posture while strength training is essential for peak performance and reducing the risk of injury. Poor form in exercises such as weightlifting can lead to muscle imbalances, joint stress, and chronic health issues [1]. The introduction of artificial intelligence has seen major advancements in real-time pose estimation and motion tracking approaches [2]. Due to these advances, it is now possible to develop intelligent fitness systems that provide immediate feedback without the need for human observation.

Bicep curls, one of the basic upper body strength exercises, require accurate elbow and wrist placement. This can be challenging for beginners to achieve. Conventional training practices are based on subjective self-evaluation, which can be unreliable or inaccurate [3]. The demand for AI-based fitness systems that can provide real-time feedback on posture, movement accuracy, and repetition quality is on the rise [4].

This project demonstrates a real-time bicep curl monitoring solution using OpenCV and MediaPipe for pose detection and motion analysis. The system detects major body landmarks, computes joint angles, and provides real-time corrective feedback to maintain proper exercise form. The primary objective is to provide a smart, autonomous device for trainers to monitor posture, monitor repetitions, and maximize training efficiency.

With the incorporation of deep learning methods into pose estimation, the system is able to successfully overcome the challenge of real-time tracking and monitoring of human movement. The primary contributions of the work are:

- A vision-based system to detect and analyze bicep curl motions.
- Real-time form correction feedback through AI-driven motion tracking.
- Automated repetition counting for consistent workout tracking.

This research contributes to the emerging discipline of AI-assisted exercise assessment by offering a scalable approach geared toward enhancing performance while reducing injury threats.

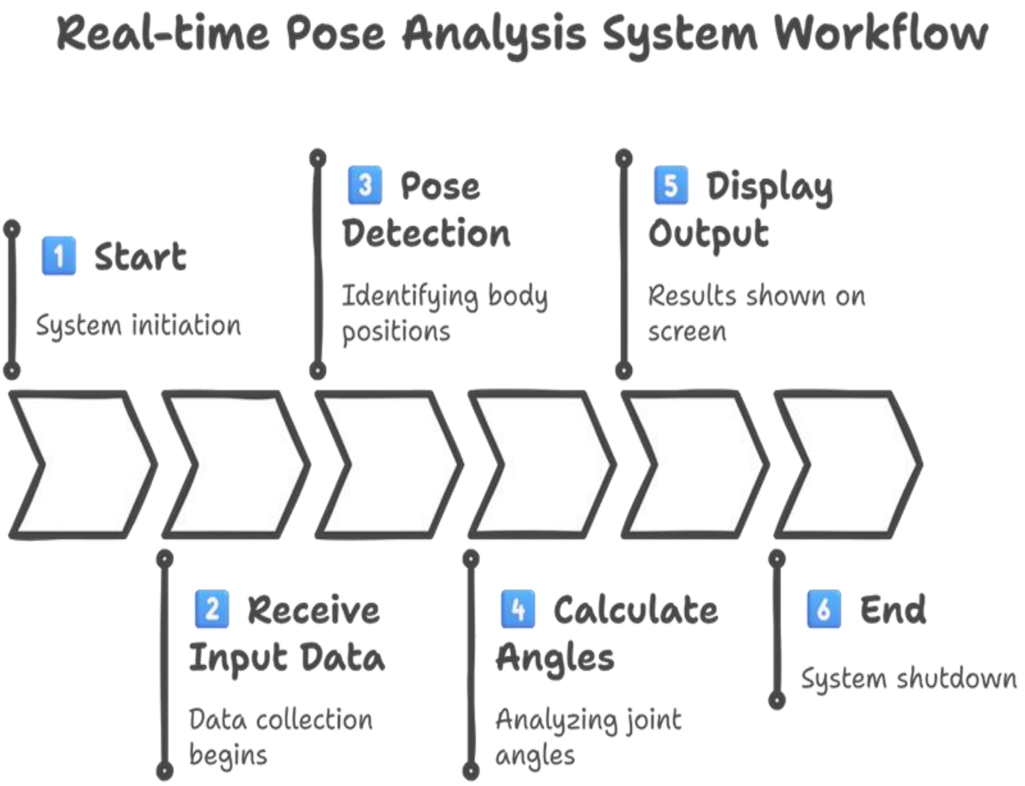

Fig. 1
Flow chart of pose detection process

2. OBSTACLES IN REAL TIME POSE ESTIMATION

Despite the significant progress in computer vision and deep learning, the issue of real-time human pose estimation is still beset by various challenges that impact its accuracy and reliability. These challenges are most relevant to fitness tracking use cases, where immediate and accurate feedback is critical.

2.1. *Occlusion and Self-Occlusion:*

Occlusion or body part blocking from the camera's view, is still a common problem. Self-occlusion occurs when one body part blocks another's view, thus occluding keypoints from being recognized correctly. It is common in dynamic poses, e.g., bicep curls, where arm movement can occlude joint visibility. Existing work has addressed the need for models capable of inferring occluded keypoints in a bid to mitigate the occlusion effect [1].

2.2. *Sudden motion and motion blur:*

Rapid movements create motion blur, resulting in fuzzy images and making it difficult to identify the proper joints. This is a major challenge in real-time applications where rapid movements are common. Work has focused on creating models that are robust to the artifacts caused by motion and are accurate [2].

2.3. *Depth Ambiguity in Monocular Configurations:*
Monocular camera systems, which employ one camera, experience depth perception limitations, thereby rendering the precise joint three-dimensional position localization more difficult. Such depth uncertainty can affect pose estimation errors, especially in the distinction between similar poses positioned at different depths. Experiments have revealed that in the absence of extra depth cues, monocular systems suffer from the inability to resolve such ambiguities [2].

2.4. *The Incorporation of Spatial and Temporal Data:*
Efficient combination of the spatial and temporal information is essential to precise pose estimation. Spatial information offers information regarding joint positions within single frames, while temporal information records motion between frames. Incorporating the two forms of information remains challenging to accomplish since the model has to find a compromise between the real-time spatial precision and continuity offered by the temporal information [2].

2.5. *Computational Limits for Real-Time Processing:*
To be able to perform in real time, models must be both accurate and efficient in computation. But such accurate models usually require enormous computational power, which would be too inconvenient to be used in processing-limited devices. A balance between efficiency and accuracy is an important consideration in the context of real-time applications [3].

2.6. *Anatomical Variability Among Individuals:*
Anatomical differences between individuals, like variations in body proportions and length of limbs, create difficulties for pose estimation models trained on standardized data. Such models may not be able to learn the idiosyncrasies of individual users and can fail in estimations with decreased accuracy. For overcoming such variability, customized calibration processes or adaptive algorithms need to be employed [4].

2.7 *Environmental Factors and Background Clutter:*
Environmental conditions can affect the quality of pose estimation systems negatively, such as lighting variability and the presence of background clutter. Poor lighting can conceal important features, while intricate backgrounds can introduce noise that results in false joint detection. Therefore, it is important to develop strong models that can work well under different environmental conditions [5].

2.8 *Real-Time Performance on Edge Devices:*
The deployment of pose estimation models on edge devices like smartphones and embedded systems is confronted with the problem of computational resources and energy consumption. Models should be performance-efficient but not compromise on accuracy to enable real-time processing on these devices [2].

2.9 *Latency Issues and Instant Feedback:*
The capacity to provide real-time feedback in pose estimation applications is critical, especially in rehabilitation and fitness contexts. Latency caused by computationally intensive tasks can, however, impact the responsiveness of pose estimation systems. Optimizing algorithms to minimize latency without trading off accuracy is an issue of high priority [6].

3. ADDRESSING REAL TIME POSE ESTIMATION ISSUES

3.1 *Mitigating Occlusion and Self-Occlusion:*
Self-occlusion, most evident in movement actions like bicep curls, is a major challenge to precise joint detection [1]. As a compensation for this:

- Redundant Landmark Use: MediaPipe gives 33 precise landmarks, with several points surrounding the joints. By considering nearby landmarks (i.e., shoulder, elbow, and wrist), we approximated the location of the occluding joints.
- Temporal Smoothing Techniques: Temporal filtering enabled the prediction of occluded joint locations from their past locations to provide continuity and avoid abrupt detection changes.

3.2 *Handling Motion Blur and Rapid Movements:*
Repetitive arm movements during exercising might create motion blur, which influenced landmark detection accuracy [2]. Solutions of ours were.

- High Frame Rate Video Capture: In recording video at higher frame rates, we minimized blur motion between frames and thereby allowed sharper images to be processed subsequently.
- Adaptive Thresholding: Adaptive detection thresholds were varied according to movement speed so that high-speed movements would not affect detection accuracy.

3.3 *Resolving Depth Ambiguity in Monocular Environments:*
Monocular camera setups lack depth vision, leading to ambiguities in 3D pose estimation [2]. In order to mitigate this:

- Z-Axis Data Integration: MediaPipe delivers normalized z-axis values per landmark. By analyzing these values, we estimated joints' relative depths, which improved 3D pose reconstruction.

- Fixed Camera Angle Assumption: The uniformity of the same camera angle and distance enabled the standardization of depth estimations throughout the various sessions.

3.4 *Integrating Spatial and Temporal Information:*

Accurate estimation of pose involves the integration of spatial and temporal data. Our approach included:

- Sequential Frame Analysis: Through frame sequence analysis, we were able to represent temporal dynamics of movements, enabling us to better estimate poses over time,.
- Kalman Filtering: Kalman filter usage allowed for the estimation of future joint positions from their current and historical states, hence enhancing pose trajectory smoothness.

3.5 *Enabling Generalizability to Different Conditions:*

Model performance can be influenced by lighting, background, and wardrobe changes. Generalization can be improved by:

- Data Augmentation: It was also trained on augmented data to mimic varied light conditions, backgrounds, and patterns of clothes, thus becoming stronger.
- Enhancement Methods: The application of contrast normalization to input frames enabled consistent landmark detection regardless of changing visual conditions.

3.6 *Enhancing Efficiency for Immediate Processing on Peripheral Devices:*

Optimization was required to provide real-time performance on systems with limited computational resources.

- Model Quantization: We used MediaPipe's light models that are optimized for mobile and edge devices, providing faster inference without compromising too much on accuracy.
- Effective Memory Management: Employing memory-conservative data structures and eliminating redundant computation eased the processing load.

3.7 *Precise Repetition Counting and Form Evaluation:*

Measuring repetitions and form checks are crucial in fitness applications. Research that we conducted included:

- Angle Calculation Between Landmarks: By angle calculation between prominent landmarks (shoulder-elbow-wrist), we calculated the bicep curl phase to enable proper repetition counting.

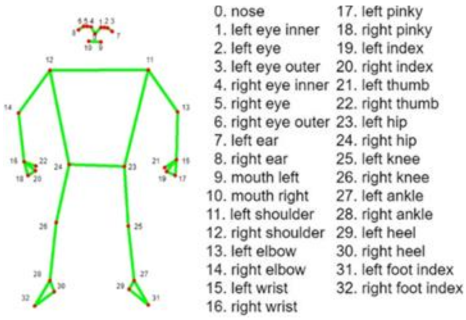

Fig. 2

33 Landmarks detected on the human body using MediaPipe

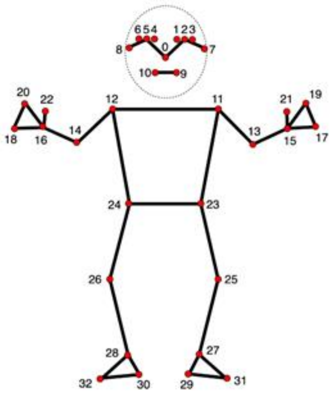

Fig. 3

Pose Landmarks Indicies

4 METHODOLOGY AND IMPLEMENTATION

This section explains the methodological setup utilized for the real-time pose estimation of the bicep curl. The combination of real-time pose estimation with the MediaPipe pre-trained model and a joint angle analysis module has been used for improving the accuracy of motion tracking [11]. A systematic process has been followed with video frame preprocessing, landmark detection, joint angle calculation, and repetition counting, thus enabling proper and timely monitoring of the exercise. The system is particularly designed to be executed in real-time, enabling instant posture assessment and feedback to the users, thus helping users achieve proper form. The following subsections explain each of these components in detail, including their implementation and impact.

4.1. *Pose Estimation and Bicep curl tracking method:*

Correct form must always be maintained when doing strength training exercises to prevent injury and get the optimal performance. Joint movement tracking issues due to occlusions, landmark detection issues, and low light conditions, as identified by most of the current literature [12–15] using pose tracking models in fitness use cases, are particularly critical in the tracking of bicep curls, where keypoint detection issues at the elbow or wrist joints may lead to erroneous repetition counting and poor form discrepancy classification. To address such issues, our work integrates real time pose estimation using MediaPipe pre trained model and an angle based analysis module to ensure improved tracking accuracy.

4.2 *Camera Input and Video Frame Processing:*

Real-time video frames are obtained from an external camera, providing real-time tracking of user motion. Every frame is preprocessed to convert the color format from BGR (used by OpenCV) to RGB, thus making it compatible with MediaPipe's pose estimation model. The conversion is performed to improve landmark detection accuracy by making the input format compatible with MediaPipe's pre-trained models.

4.3 *Landmark Detection and Tracking:*

After preprocessing the frames, MediaPipe Pose Estimation is utilized to detect 33 key points on the human body [12]. Of the detected landmarks, the position of the elbow, shoulder, and wrist are specifically highlighted in order to enable accurate tracking of the motion of the bicep curl.

4.4 *Joint Angle Calculation and Form Analysis:*

To quantify the user motion quality, joint angles are calculated from shoulder, elbow and wrist positions. Joint angle theta between joints is calculated from trigonometric functions, useful when computing the different phase of an exercise. DOWN phase is indicated by an elbow angle of greater than 160 degrees, i.e., maximum arm extension. UP phase is indicated when the elbow angle is less than 35 degrees, i.e., maximum arm flexion.

4.5 *Hyperparameters for Pose Detection:*

For enhancing the system's pose and exercise stage estimation accuracy and confidence in detection and tracking, some configurations were selected. The confidence threshold of detection and tracking was both set to a value of 0.5. This configuration is meant to eliminate low-confidence detection of landmarks and provide stability to tracking from one frame to the next. Angle-based constraints were also incorporated to differentiate between the downward and upward of the bicep curl. Table I shows the selected configurations and the rationale behind each.

5 TECHNOLOGIES AND TOOLS UTILIZED

The use of the real-time bicep curl monitoring system was made possible by combining multiple technologies and tools, each of which contributed to various aspects of the system's functionality.

5.1 *MediaPipe Pose:*

For human pose estimation, MediaPipe Pose was utilized. The MediaPipe Pose framework offers a machine learning solution to high-fidelity body pose estimation, estimating 33 3D landmarks from RGB video frames. The landmarks include prominent body points such as shoulders, elbows, and wrists, which are essential for movement analysis of the bicep curl. MediaPipe's real-time performance on multiple platforms made it appropriate for this use case [16].

5.2 *OpenCV:*

OpenCV (Open Source Computer Vision Library) was used for image and video processing operations. It supported operations like reading frames from the webcam, conversion of color space, and visualization of the detected landmarks. OpenCV's rich functionality and compatibility with Python made it a suitable option for the computer vision needs of the system [17].

5.3 *Python Programming Language:*

Python was used as the main programming language for system development. Its simplicity and the existence of libraries made the pose estimation pipeline operate efficiently. In addition, Python's compatibility with OpenCV and MediaPipe made the direct incorporation of the various components [18] possible.

5.4 NumPy:

NumPy was utilized for numerical computations, particularly to calculate angles between joints to check the precision of bicep curl exercises. Its array operations allowed efficient processing of the landmark coordinates obtained from MediaPipe.

6 EXPERIMENTAL RESULTS AND OBSERVATIONS

6.1 System Configuration

The experiments conducted for this study were carried out on a system with the following hardware and software specifications:

Operating System: Windows 11 (64-bit)

Processor: Intel Core i7 (2.3 GHz, 14 cores)

Memory: 16 GB DDR5 RAM

Graphics Processing Unit (GPU): NVIDIA RTX with 6 GB VRAM

Programming Environment: Python 3.10.12

Deep Learning and Computer Vision Libraries: MediaPipe, OpenCV, and NumPy.

6.2 Findings

The bicep curl monitoring system was tested in real-world usage scenarios. The emphasis was on pose estimation accuracy, joint angle detection, repetition count consistency, and real-time performance.

6.2.1 Pose estimation and landmark detection:

The system was able to accurately capture upper-body joint movement, with an emphasis on the shoulder, elbow, and wrist landmarks. Fig.4 illustrates that the MediaPipe-based detection pipeline was able to accurately detect key points and track arm movement consistently throughout repeated bicep curls.

FIG.4 Key point detection and tracking

6.2.2 Joint Angle Measurement and Range of Motion Analysis:

Joint angles were specified in relation to the shoulder, elbow, and wrist positions. As shown in Fig. 5, real time assessment of exercise technique was made possible by the system. Movement phases were labelled by threshold angles:

Down Position: Arm extended completely (angle > 160°)

Up Position: Arm maximally flexed (angle < 35°)

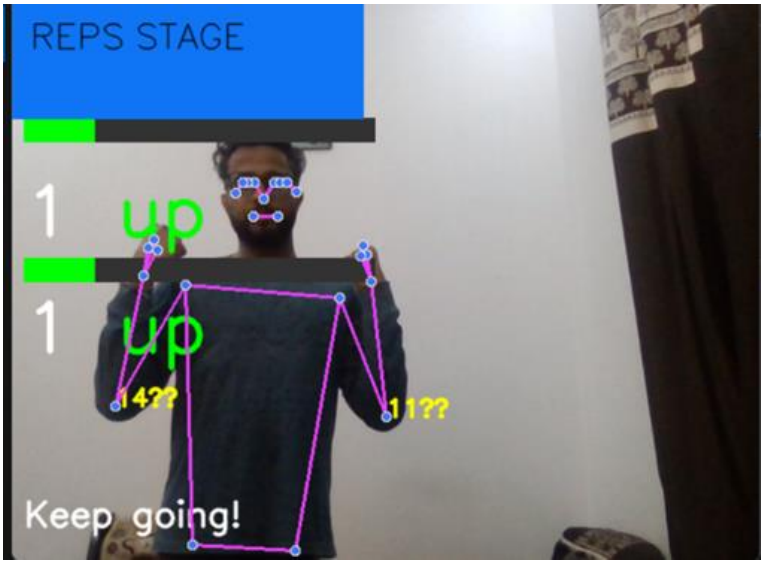

FIG.5 Real time computation of elbow angle

6.2.3 Set Completion and Reset Mechanism:

As can be observed in Fig. 6, the system had a visual cue for set completion upon a user attaining a fixed number of repetitions. The feedback system significantly enhanced user motivation and interest. The system also had a reset system whereby users could trigger a new set by pressing the 'r' key, thus facilitating easy tracking across sets.

FIG.6 Set Completion Notification Display

6.2.4 Immediate Response:

The system gave instant feedback on posture and repetitions achieved. The real-time factor is one of the most important requirements in fitness apps, where instant feedback significantly helps in enhancing performance and minimizing the risk of injury. User Insights and Real-World Implementation User reviews emphasized the usefulness of providing instantaneous feedback to enhance the efficacy of exercises. The reminders of set completion acted as motivational benchmarks, and the reset feature enabled uninterrupted workout sessions. On the whole, the system exhibited profound utility for systematic exercise routines.

7 CONCLUSION

The system was shown to have potential in successfully tracking principal upper-body reference points—shoulders, elbows, and wrists—throughout bicep curls. With real-time joint angle calculation, it was able to identify phases of exercise and form deviations precisely, allowing for precise movement analysis.

The application of pose estimation technology was shown to have a great potential in improving sports performance while reducing the risk of injury. Real-time feedback was required in the optimization of techniques, and the integration of artificial intelligence-based systems enabled the development of adaptive, personalized training experiences. Integration of smart coaching features in the future might further refine the outcome of training through customized coaching. Future extensions might include incorporating additional exercises into the system and using predictive analytics to analyse and mitigate the risk of injury. Overall, the solution provides a robust and normalized approach to real-time exercise monitoring with important implications for optimizing exercise performance and preventing injury in the context of sports and fitness training.

REFERENCES

[1] Neupane, R.B., Li, K. & Boka, T.F. A survey on deep 3D human pose estimation. Artif Intell Rev 58, 24 (2025).

[2] A. Bazarevsky, I. Grishchenko, K. Raveendran, and M. Grundmann, "BlazePose: On-device Real-time Body Pose tracking," arXiv preprint arXiv:2006.10204, Jun. 2020.
[3] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh, "Deep Learning-Based Human Pose Estimation: A Survey," *arXiv preprint arXiv:2012.13392*, Dec. 2020

[4] M. Tatariants, "Challenges of Human Pose Estimation in AI-Powered Fitness Apps," InfoQ, Oct. 15, 2020.

[5] Fukushima, T., Blauberger, P., Guedes Russomanno, T. et al. The potential of human pose estimation for motion capture in sports: a validation study. Sports Eng 27, 19 (2024).

[6] S. Yuan and L. Zhou, "GTA-Net: An IoT-Integrated 3D Human Pose Estimation System for Real-Time Adolescent Sports Posture Correction," arXiv preprint arXiv:2411.06725, Nov. 2024.

[7] Google AI Edge, "MediaPipe Pose: High-Fidelity Body Pose Tracking," GitHub, 2023.

[8] R. Johnson, "Real-Time Pose Estimation with Python and OpenCV: A Hands-on Guide," Data Rodeo, Dec. 8, 2023.

[9] Hackers Realm, "Realtime Human Pose Estimation using Python," Hackers Realm, Jul. 1, 2022.

[10] Google Research, "On-device, real-time body pose tracking with MediaPipe BlazePose," Google Research Blog, Aug. 2020.

[11] Rama Bastola Neupane, Kan Li and Tesfaye Fenta Boka. "A survey on deep 3D human pose estimation". Artificial Intelligence Review, vol. 58, article number 24, Nov. 2024.

[12] Wenqiang Zhang, Jiemin Fang, Xinggang Wang, and Wenyu Liu "EfficientPose: Efficient human pose estimation with neural architecture search". Computational Visual Media, vol. 7, pp. 335–347, April 2021.

[13] Sihan Ma, Jing Zhang, Qiong Cao, and Dacheng Tao, "PoseBench: Benchmarking the Robustness of Pose Estimation Models under Corruptions," arXiv preprint arXiv:2406.14367, Jun. 2024.

[14] Honghong Yang, Longfei Guo, Yumei Zhnag and Xiaojun Wu "U-shaped spatial–temporal transformer network for 3D human pose estimation". Machine Vision and Applications, vol. 33, article no.82, Sep.2022.

[15] Yucheng Chen, Yingli Tian, and Mingyi He, "Monocular Human Pose Estimation: A Survey of Deep Learning-Based Methods," Computer Vision and Image Understanding, vol. 192, p. 102897, March 2020.

[16] Google Research, "On-device, real-time body pose tracking with MediaPipe BlazePose," Google Research Blog, Aug. 2020.

[17] Sudharshan Chandra Babu, "A 2019 guide to human pose estimation with deep learning," Nanonets, April 2019

[18] Ko, Yeong-Min & Nasridinov, Aziz and Park, So-Hyun, "Real-Time AI Posture Correction for Powerlifting Exercises Using YOLOv5 and MediaPipe," in IEEE Access. 12. 195830-195853, Dec. 2024.