International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Exercise Movement Detection Using Spearman Correlation-based Sliding Window Technique

[1]Anant Kaulage, [2]Deepak Mane, [3]Gopal Upadhye, [4]Satpalsing D. Rajput, [5]Satish Kale, [6]Bhushan Zope

Submitted: 17/07/2023 **Revised**: 05/09/2023 **Accepted**: 24/09/2023

Abstract: The lack of an efficient and fast way to track and monitor exercise types and the number of actions completed during workout sessions can lead to incomplete or inaccurate data for fitness enthusiasts, coaches, and healthcare professionals. In this paper, we proposed a system that uses the existing sliding window algorithm; in addition, we modified it with Spearman correlation using machine learning. Also proposed is the developing of a software solution that utilizes advanced algorithms to detect and count the total number of actions completed. Our solution aims to address this issue by providing accurate and comprehensive data, which can be used by fitness enthusiasts, coaches, and healthcare professionals to make informed decisions about workout routines. The proposed solution will provide an innovative and accurate way to track exercise and help individuals achieve their fitness goals. No standard dataset was available for exercise poses, so we created and tested the model on our dataset. The hyperparameter gamma tuned to optimize the classification accuracy on our dataset produced 81%, which is better than other approaches. The experiment results demonstrate the effectiveness of the system in tracking and correcting exercise poses, and its potential to enhance the quality of exercise practice.

Keywords*: Pose Estimation, Spearman Correlation, Euclidean Distance, Sliding Window, OpenCV, Mediapipe, SciPy*

1. Introduction

Regular exercise is crucial for maintaining a healthy lifestyle and being happy. However, it is equally important to perform exercises with correct posture to avoid the risk of injury. Unfortunately, many individuals forget to keep track of the number of repetitions they perform during a workout, leading to overexertion and decreased muscle hypertrophy. The rise of at-home workouts has made receiving proper guidance and supervision even more challenging. Personal trainers can be expensive, and not everyone can afford them. In this context, Exercise Pose

Detection becomes essential. Such a system will provide real-time feedback, allowing users to adjust their workouts accordingly to maximize muscle hypertrophy. Specific pose detection models using deep learning and machine learning models are already in place. The general methodology of such a project includes detecting a human face and body, drawing a skeleton of points on the various joints, as shown in Fig. 1, and then training a machine learning model on media input to recognize the exercise effectively.

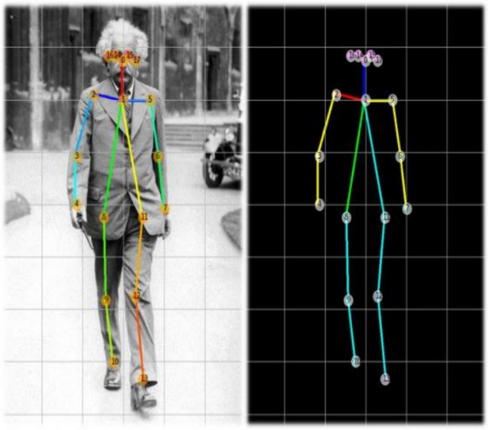


Fig 1.A skeleton drawn on input image

However, they use a certain amount of computational power and use neural networks throughout. This paper presents a system that can algorithmically check for the correct pose while limiting the requirement of neural networks only for detection. The purpose of this modified system is to track and monitor all parameters of a workout, including counting sets and repetitions accurately, while supplanting the use of multiple machine learning models to carry out all the processes. It uses similarity and distance-based techniques to recognize human actions accurately. This project aims to provide a cost-effective alternative to having a personal trainer, making it easier to apply such a system in a real-life environment without particular computational time and space constraints.

[1,2,3]Vishwakarma Institute of Technology, Pune-411037, Maharashtra, India
[4]Pimpri Chinchwad College of Engineering Pune-411044, Maharashtra, India
[5]AISSMS Institute of Information Technology Pune-411001, Maharashtra, India
[6]Symbiosis Institute of Technology, Symbiosis International (Deemed University) (SIU), Lavale, Pune 412115, India
[1]anant.kaulage@gmail.com, [2]dtmane@gmail.com, [3]gopalupadhye@gmail.com, [4]rajputsatpal@gmail.com, [5]satish191@gmail.com, [6]bhushan.zope@hotmail.com

The objectives of the paper are,

- Our primary objective is to handle this problem using Python and OpenCV to support individuals receiving personalized training from their trainers.
- The main goal of the work is to create and execute a precise model that can effectively track motion and teach various physical exercises.
- The model will assist in identifying movement patterns during workouts and categorizing them based on the user's specific requirements.

The overview of this paper is, in section II, the literature survey represents the literature on exercise detection and correction systems in the context of exercise practice. It discusses the limitations of existing systems and highlights the need for a more accurate and real-time solution. The proposed architecture, section III, provides a detailed explanation of the Machine Learning based exercise pose detection and correction system. It covers the steps involved in training the Machine Learning model, including data preprocessing, feature extraction, and model selection. It also explains how the model is used to detect and correct different exercise poses in real-time. The results of experiments conducted to evaluate the system's performance are represented in sections IV& V. It includes a comparison of the system's accuracy to other models and a discussion of the real-time performance of the system. Conclusion with future direction explained in section VI.

2. Literature Review

The scope of this literature review includes academic journals, conference proceedings, and other relevant sources that discuss the detection of exercise types and the number of actions completed.

during workout sessions. The search strategy used to identify relevant literature included searching through various academic databases, such as Scopus, IEEE Xplore, and Google Scholar, on various pose estimation and human exercise detection models for a thorough analysis for building the prerequisite knowledge for this project The selection criteria for the literature review included articles and studies related to the detection of exercise types and the number of actions completed during workout sessions. The decision of which papers to review or discard was taken based on consideration of technology and scope [1].

Multiple studies were selected for this literature review: Studies by Sun, Chen and Ye around 2022 implementing the same application were reviewed. These proposed a model made in two stages; a deep learning modeland heatmaps of the body, for human action recognition, which could be used to detect exercise types during workout sessions. Chen et al. developed a deep learning model based on pre-trained models that could detect fitness movement types and completeness. Ye et al. proposed an exercise identification method built by aggregating the score of various modules based on a time related algorithm. All these studies showed promising results in detecting exercise types and the number of actions completed during workout sessions. Deep learning networks with neurons have been used for pose estimation for humans in most regression and classification approaches.[2][3] The 2D and 3D points on the human are predicted using machine learning networks in these studies whereas a classification models and papers on them use classes of various poses and output their projection according to the model. Most of these methods have architectural design of multi layered neural networks interloped with heavy datasets and have limited flexibility for working with web or graphic interface options for the purpose of actual application .In conclusion, the studies reviewed demonstrated the potential of using advanced algorithms, such as deep networks which are either pre-trained or use deep learning technologies, to detect exercise types and the number of actions completed during workout sessions accurately. However, there is still a need for further research in the application of such methods and their working and compatibility with frontend or GUI frameworks.[4]

3. Proposed System

Here, we proposed a system that uses the existing sliding window algorithm. In addition, we modified it with Spearman correlation using machine learning libraries and PyQT5 GUI to implement exercise pose detection. No standard dataset for such exercise poses is available, so we created our own dataset. The system architecture of the proposed model is represented in Fig.2. Here, media is an input of both format image and video by using OpenCV with a human being doing an exercise. Mediapipe with Blazepos model libraries is used to carry out the preliminary task of detecting the human being and detecting points on the skeleton of the body. The drawn skeleton has certain points on important joints for each exercise; these positions are extracted, and then the angle between them is calculated to check if the pose during the exercise is proper. If the exercise is done properly, the counter is increased, and the result is stored. Exercises, we use Spearman correlation to store the values from a video and image input given by the user and then store the values of angles

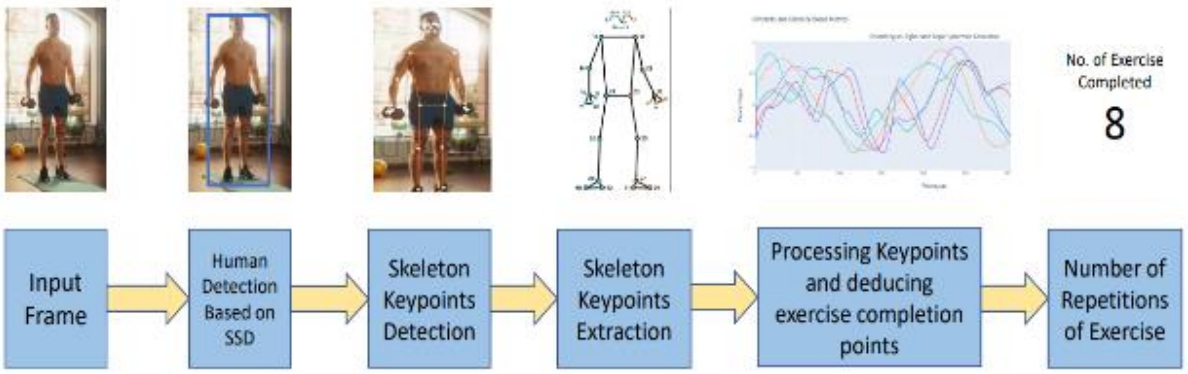


Fig 2. System Architecture

Between important points for reference later. All this is displayed in a GUI created by PyQT5.The machine learning library Mediapipe uses an already-trained model to identify the points of the skeleton on an image or video. Both these libraries are easy to interface with Python and Computer Vision. The process marks the coordinates of joints and the points of those joints in certain variables: each joint of the elbow, ankle, shoulder, etc., is stored and can be referenced later. The skeleton created by these points is shown below in Figure 3. The calculation of the angle between the joints in the region that we need, e.g., for curls, we need elbow, shoulder, and wrist, is done, and the necessary variables are extracted to iterate upon and check if the pose is correct. For an image, this is straightforward and mostly requires a simple algorithm for an input image. For video and live video using a webcam, we have to use a continuous loop where variables are updated continuously and live counters are ticked for the completion of one set of exercises and detecting points on the skeleton of the body.

```
Data: A: The array
n: Length of the array
k: Maximum number of minutes to beused
Result: Returns the maximum length of range
whose sum is at most k
Algorithm:
answer <-- 0;
sum + <-- 0;
R+ 1;

for L1 to n do
|  if L> 1 then
|  |     sum + sum - A[L-1];
|  end
|  while R≤n AND sum + A[R] ≤ k do
|          sum + sum + A[R];
|          R+R+1;
|   end
|   answer + maximum (answer, R-L);
end
```

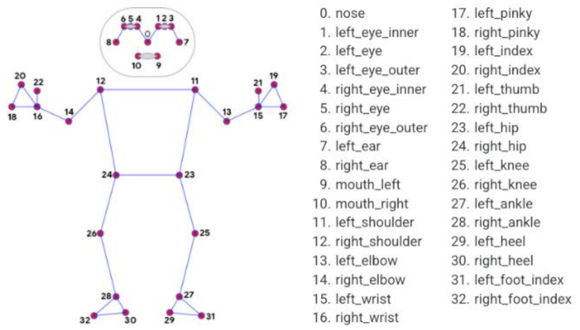


Fig 3. Skeleton diagram

Figure 4 illustrates the drawn skeleton has certain points on important joints for each exercise, these positions are extracted, and then the angle between them is calculated to check if the pose during the exercise is proper. If the exercise is done properly, the counter is increased, and the result is stored [10].

Fig 4. Display of skeleton on human body

Teaching a new exercise by user input requires the user to input a video and image in proper format and input necessary joint locations for the system to store the variables properly. The system then takes in the input variables. It correlates two variables: the one of the input images in the future and then of the input data beforehand to identify the exercise properly. After that, the processing of key points to draw graphs for analysis is done. The entire system is wrapped in GUI created by PyQT5, allowing us to have our entire project in one language and straightforward design of a frontend for the application.[5]

4. Methodology

A. *Sliding window Similarity Check*

In this paper, a modified sliding window algorithm is proposed which helps to classify the exercise poses based on joint angles. Proposed approach for classification involves a uncommon technique, unlike used in image processing to analyze and process images in small, overlapping rectangular sections. The basic idea behind the algorithm is to apply a function or operation to each of these small sections, known as windows or patches, in order to extract useful features or perform some form of analysis on the image. The sliding window algorithm involves moving a rectangular window across the image, one small step at a time, and processing the image within the window. The window typically has a fixed size and can be moved in any direction depending on the requirements of the algorithm. The window is moved one pixel or a few pixels at a time until it have covered the entire image.

In feature extraction, the sliding window algorithm can be used to extract features from an image by analyzing the pixels in each window. This can be done using various techniques such as edge detection, colour histogram analysis, or texture analysis. By sliding the window over the entire image, a set of features can be extracted that can be used to classify or identify the image. In our case we are using it to extract the region and points of interest to calculate the angle between them for our algorithm. The output generated by sliding window similarity check contains noise which can interfere with peak estimation thus a smoothing algorithm called moving average is applied on the sinusoidal wave.

B. *Spearman Correlation*

Spearman correlation is a statistical value that is commonly measured for evaluating a relational monotonic function of variables

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$$

Equation 1: Spearman Correlation

Where,

r_s = Spearman correlation coefficient

Σd^2 = sum of the squared differences between the ranks of the two variables

n = number of observations

The Above Equation No. 1is commonly used in data analysis to explore the relationship between two variables that may not have a linear relationship.

In image processing, the Spearman correlation can be used to compare the similarity between two images or between different regions within the same image. Spearman correlation in this system is to compare different regions within the same image. Regions with high correlation coefficients can then be grouped together.

To use the sliding window algorithm with correlation for pose estimation, we first need to create a template image that represents the object of interest in a specific pose. This template image can be generated using various techniques such as 3D modelling, manual image annotation, or machine learning. Once the template image is created, we can slide a window of the same size over the input image or video frame and calculate the correlation between the window and the template image.[8][9]

The correlation coefficient is a measure of similarity between two images, and it ranges from -1 (completely dissimilar) to 1 (identical).

By comparing the Spearman correlation coefficients at different window positions, we can identify the position where the window best matches the template image, which corresponds to the estimated position of the object in the image or video. In our case we use it to compare the region in which the exercise is performed and then correlate it with the input exercise to check if they are similar to validate the correctness of the performance.

C. *Data Analysis*

A dataset was primarily created for analyzing variation in angle. It was difficult to get the dataset we needed on the internet, so we had to create the entire dataset on our own. Each individual video was run using OpenCV and while to track and collect joint coordinates Mediapipe library was used. Then the entire coordinates were imported to a .csv file for data analysis.

The experimental results demonstrate that our approach achieves a high accuracy of 81%, which outperforms other state-of-the-art methods on the same dataset.

5. Results and Discussions

The paper uses machine learning techniques to describe a system for exercise pose classification and correction. The system extracts key points from an exercise pose performed by a practitioner using an image or a video and matches them to other techniques. Once the pose is classified, if the exercise pose is incorrect, the system will give instructions for correction of the exercise pose so that they can learn and correct their posture accordingly. The system achieved an accuracy rate of 81%, indicating its potential as a helpful tool for exercise practitioners, which is better than existing approaches.

However, limitations and challenges may exist, and further research may be needed to improve the accuracy and usefulness of the system.

1. Data and Input

Due to the lack of videos or records on the internet of people performing exercises in both proper and improper ways, we took most of the self-collected videos. Figure No. 5 is taken as an input in image or video format, and the system requires png and JPEG images for good validity. In addition to the image format, the pose estimation algorithm may also require the image to have certain characteristics, such as a specific resolution or aspect ratio. The resolutions and aspect ratios should be moderate. A general ‘good resolution like 1200x800 should be sufficient for good results. Further, the color format, contrast, or lighting should be in the normal visual range. Images in both RGB and B&W formats are supported, but in certain color combinations, the contrast or lighting may cause problems with accuracy.[11][12]

Fig 5. Example of a clear image

2. Accuracy and Analysis

Figure 6 illustrates the confusion matrix generated by the classifier, which summarizes the classification outcomes for each class. The matrix represents the number of instances classified into each class, with the true labels displayed on the y-axis and the predicted labels on the x-axis.

The accuracy of pose detection of exercise is calculated and displayed using confusion matrices for checking validity.

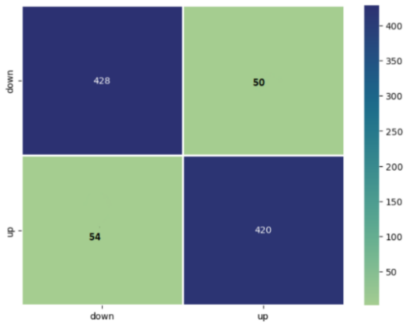


Fig 6. Squats Confusion Matrix

The above figure 6 shows the confusion matrix for the squats pose of exercise.

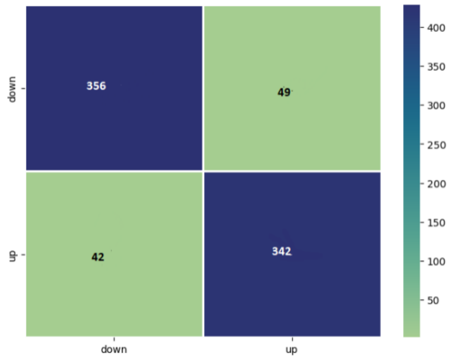
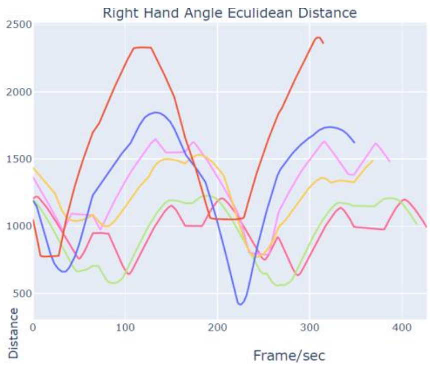


Fig 7. Curls Confusion Matrix

The above figure 7 shows the confusion matrix for the curls pose of exercise.

Spearman correlation is done for the data generated through data analysis the angle between the joints and its fluctuations is plotted against the times or frames per second spent.

(a)



(b)

Fig 8. (a)(b) Graphs for correlation of right-hand curl

The comparison of our model was done with three other open-source models. For proper comparison of models reviewed the media given to the model was similar and the average accuracy over multiple body parts was takenas shown in Figure 8 given above.

Table 1. Accuracy of Curls, Lunges , Squats and Raises for our model

Exercise	Accuracy
Curls	89%
Lunges	72%
Squats	79%
Raises	84%

Table 1. Accuracy of Curls,Lunges,Squats &Raises

Table 2. Accuracy of OpenPose, and MoveNet and our model

Model	Average Accuracy
OpenPose	78%
Movenet	78%
Proposed Model	**81%**

Table 2. Result Comparisons

6. Conclusion

This paper proposed a modified sliding window algorithm with Spearman correlation algorithm-based smart exercise assistant for exercise pose classification using machine learning techniques that hold great potential for enhancing the tracking and monitoring exercises using similarity-based metrics and is a promising tool. By analyzing movement patterns and providing real-time feedback, this program has the potential to improve exercise adherence and effectiveness, reduce the risk of injury, and enhance overall fitness outcomes. Such software can be new to exercise, have limited access to personal trainers, or are recovering from an injury. An additional advantage of using the sliding window algorithm with Spearman correlation for pose estimation is that it is more robust to outliers and non-linear relationships between the window and the template image than the standard correlation coefficient.

References

[1] Sun, R. et al. (2022) "Human action recognition using a convolutional neural network based on skeleton heatmaps from two-stage pose estimation", Biomimetic Intelligence and Robotics, Volume No 2 Issue 3, pp. 100062.Availableat:https://doi.org/10.1016/j.birob.2022.100062.

[2] Chen, K.-Y. et al. (2022) "Fitness Movement Types and Completeness Detection Using a Transfer-Learning-Based Deep Neural Network" Sensors, Volume 22, 5700. https://doi.org/10.3390/s22155700

[3] Ye et.al (2022) "Human action recognition method based on motion excitation and temporal aggregation module", Heliyon, Volume 8 Issue (11). Available at:

https://doi.org/10.1016/j.heliyon.2022.e11401.

[4] Pavllo et. al. (2019) "3D Human Pose Estimation in Video With Temporal Convolutions and Semi-Supervised Training." 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). doi:10.1109/cvpr.2019.00794

[5] Rogez et. al. (2019). LCR-Net++: "Multi-person 2D and 3D Pose Detection in Natural Images". IEEE Transactions on Pattern Analysis and Machine Intelligence.

[6] Schober et. al. (2018) "Correlation Coefficients: Appropriate Use and Interpretation". Anesthesia& Analgesia 126(5):p 1763-1768, May 2018.

[7] N.I. Glumov et. al. (1995) "Detection of objects on the image using a sliding window mode, Optics & Laser Technology", Volume 27, Issue 4,1995,Pages 241-249,ISSN0030-3992 https://doi.org/10.1016/0030-3992(95)93752-D.

[8] Song, Liangchen et. al. (2021). "Human pose estimation and its application to action recognition: A survey", Journal of Visual Communication and Image Representation. 76. 103055. 10.1016/j.jvcir.2021.103055.

[9] Jo, Beomjun et. al. (2022). "Comparative Analysis of OpenPose, Pose Net, and MoveNet Models for Pose Estimation in Mobile Devices". Traitement du Signal. 39. 119-124. 10.18280/ts.390111.

[10] Mane, D. et. al. (2023), "Smart Yoga Assistant: SVM-based Real-time Pose Detection and Correction System", International Journal on Recent and Innovation Trends in Computing and Communication, volume 11, Issue 7s, pp. 251–262.

https://doi.org/10.17762/ijritcc.v11i7s.6997

[11] Badiola-Bengoa et. al. (2021) " A Systematic Review of the Application of Came" Ra-Based Human Pose Estimation in the Field of Sport and Physical Exercise". Sensors 2021, 21, 5996. https://doi.org/10.3390/s21185996

[12] Sudharshan Chandra Babu et. al.(2019) "A guide to human pose estimation with deep learning".https://nanonets.com/blog/human-pose-estimation-2d-guide/, 2019.

[13] Andrew Hernandez, Stephen Wright, Yosef Ben-David, Rodrigo Costa, David Botha. Risk Assessment and Management with Machine Learning in Decision Science. Kuwait Journal of

Machine Learning, 2(3). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/196

[14] Sharma, R., Dhabliya, D. A review of automatic irrigation system through IoT (2019) International Journal of Control and Automation, 12 (6 Special Issue), pp. 24-29.

[15] Purnima, T., & Rao, C. K. . (2023). CROD: Context Aware Role based Offensive Detection using NLP/ DL Approaches. International Journal on Recent and Innovation Trends in Computing and Communication, 11(1), 01–11. https://doi.org/10.17762/ijritcc.v11i1.5981