International Journal of Sport Studies for Health

Journal Homepage

An Interactive Real-Time System for Pose Classification in Children's Yoga and Kavayat Exercises

Vanita. Babanne[1], Pankaj. Agarkar[2*]

[1] Department of Computer Engineering, VIIT, Savitribai Phule Pune University, Pune, India
[2] Department of Computer Engineering, Ajeenkya D Y Patil SOE Lohegaon, Pune, India

* **Corresponding author email address**: pmagarkar@gmail.com

A r t i c l e I n f o

Article type:
Original Research

How to cite this article:
Babanne, V., & Agarkar, P. (2026). An Interactive Real-Time System for Pose Classification in Children's Yoga and Kavayat Exercises. *International Journal of Sport Studies for Health, 9*(1), 1-9.
http://dx.doi.org/10.61838/kman.intjssh.4381

A B S T R A C T

Objective: Vision-based motion identification and categorization of human movement through machine learning is a crucial aspect of various applications, including healthcare, surveillance, and sports analysis.

Methods: Through the analysis of movement data, the system differentiates between diverse actions with high precision, offering valuable insights for monitoring and analysis in real-time.

Findings: This investigation demonstrates yoga and Kavayat (abbreviated as YK) that leverages machine learning techniques, specifically Logistic Regression, to precisely discern and categorize physical action patterns and classify with real-time feedback and inform the accuracy of the posture. For children, Yoga and Kavayat, sometimes called mock drill, improve the physical as well as mental health.

Conclusion: The lightweight model called PoseHeatMap achieves a remarkable 98.00% accuracy, demonstrating its capability to effectively detect and classify patterns of physical action and give real-time feedback.

Keywords: *Logistic Regression, Drill Exercises, Environmentally Integrated Technology, Machine Learning, Computer Vision, Human Action Recognition.*

1. Introduction

For maximizing fitness benefits and reducing the risk of injury, correct physical exercises execution is necessary. In particularly significant for children, structured activities such as yoga and Kavayat (mock drills) contribute to improved endurance, posture, overall health, and confidence. The World Health Organization endorses that children engage in at minimum 60 minutes of moderate to vigorous physical activity each day (1). However, the absence of trained trainers often restricts the provision of

Article history:
Received 20 August 2025
Revised 13 October 2025
Accepted 29 November 2025
Initial Published 02 December 2025
Final Publication 01 January 2026

real-time guidance, and even when supervision is available, corrections can be inconsistent due to human subjectivity and time limitations. Meet the challenge head-on with current technology-enabled intelligent solutions. Developments in computer vision and machine learning allow systems to monitor body movements and deliver immediate feedback.

Despite these advances, most prior studies focus on adults or controlled fitness environments, while children's exercises such as yoga and kavayat remain underexplored. Children present unique challenges, including differences in body proportions, irregular movements, and varied attention spans, which reduce the generalizability of existing adult-focused models. Moreover, many prior deep learning–based approaches are computationally heavy, making them unsuitable for real-time classroom or home deployment.

Frameworks like OpenPose (2) and AlphaPose (3) basic models enable automatic detection of body joints from video which create opportunities for structured exercise monitoring.; for instance, Biswas et al. (4) approach adapted for fitness and yoga applications proposed a CNN-based model for yoga posture recognition, while authors (5) developed a machine learning–based exercise feedback system to guide users toward correct form. Additionally, In (6) demonstrated 3D skeletal pose estimation via VideoPose3D indicating that accurate action recognition can be achieved even with limited data.

This research aims to tackle limitations by introducing a lightweight pipeline that integrates PoseHeatMap representations of MediaPipe landmarks with a Logistic Regression classifier. Unlike complex deep architectures, Logistic Regression is computationally efficient, interpretable, and well-suited for real-time use. Combined with spatial heatmaps of skeletal landmarks, this approach enables high accuracy in recognizing children's poses while maintaining low resource requirements. This novel framework contributes toward bridging the gap in child-specific exercise recognition, offering a practical solution for real-time correction and monitoring of yoga and kavayat drills.

Proper execution of physical exercises is essential for children's physical and mental well-being. Structured activities such as Yoga and Kavayat (mock drills, YK) improve endurance, posture, and confidence, but real-time guidance from instructors is often limited and inconsistent. Advances in computer vision and machine learning, particularly pose estimation frameworks like MediaPipe, enable automated analysis of skeletal landmarks, allowing accurate recognition of movement patterns. However, most existing approaches focus on adults or controlled settings and often require computationally intensive deep learning models, limiting real-time deployment for children. To address these challenges, this study proposes PoseHeatMap, a lightweight system that combines MediaPipe landmark-based heatmaps with Logistic Regression to classify children's exercise actions efficiently. This system leverages to develop i) a real-time, interpretable recognition system for Yoga and Kavayat, (ii) evaluate its accuracy across multiple children-specific poses, and (iii) provide a practical AI-based feedback framework for instructors and caregivers. This novel approach bridges the gap in child-focused exercise monitoring, offering high-accuracy action recognition with low computational overhead.

The primary objective is to accurately identify real-time yoga and kavayat posture of children by providing the video as input using the video frames passes to the lightweight system Medipipe, and generate the heatmap, and then classify variability and imperfect poses. Research focuses on the to enhance the movement analysis and correcting posture with advanced techniques. To evaluate this framework's performance in terms of accuracy. This study aims to contribute an intelligent system, child centered physical education tool that engages yoga and kavayat (YK) training environment.

2. Literature Review

The utilization of accelerometers and gyroscopic sensors (7), in smartphone-based human action recognition. It analyzes the features and discusses the results regarding recognition accuracy. Human activity recognition (8), is provided by using different types of sensors. The system underwent training and evaluation using a dataset that is publicly accessible. Authors of (9), model utilizes processor image and deep learning for live processing posture detection, guiding users towards correct postures by comparing them with standard yoga poses. Additionally, the system tracks gym activities, integrating features from diverse posture detection applications onto a unified interface.

It is deep learning based computer vision technologies. The cameras are required for the capturing the video. It uses CNN and ConvLSTM (10) It describes about novel logistic regression algorithm. It takes samples among that it divides in 2 groups. The original algorithm named logistic regression demonstrates higher accuracy as compared to the linear regression algorithm (11) In the study (12), a new architecture called YOLACT-TRN (You Only Look at Coefficients - Temporal Relation Network) is presented. This architecture is specifically designed for real-time understanding of hand actions. It combines in real instance segmentation, YOLACT incorporating a Temporal Relation Network (TRN). The TRN to facilitate a more thorough understanding of hand actions.

In (13) authors utilize deep convolutional neural networks in two distinct approaches: firstly, for end-to-end training with our RGB image dataset, employing transfer learning with ImageNet weights to categorize five postures; secondly, we leverage a pre-trained deep convolutional network (pose estimator) which has demonstrated accuracy. Introducing a novel method for representation learning utilizing randomly combined, smoothly connected sets of three instances, termed blended triplets, which augment the distinguishing characteristics between recognized and unfamiliar activities (14) .This innovative method employs Skeleton Detection technology to measure changes in skeleton points during consecutive actions. It utilizes a nearest neighbor technique to classify velocity levels and directs the ST-GCN model to infer changes in states (15) The methodology (16), employs convolutional neural networks for pose recognition, incorporating a model for localizing human joints to identify discrepancies in posture.

The explores recent advancements in human pose recognition achieved through the application of (17), deep learning methods. Deep learning revolutionizes pose estimation by enabling the expansion of extremely precise and well-organized models that can handle complex poses and occlusions. Ihis paper (18), tried three experiments. This paper has trained the system which recognizes each pose of human. It is also helpful in the transformation of variations of different actions into consistent patterns. It (19), describes about action recognition and image processing. It takes one dataset and examine the performance and accuracy. The dark videos are divided into images. It states (20), gives training for videopose3D. It requires less amount of data for human activity. Action based and post estimation problems can be worked. In machine learning, Logistic Regression (LR), a far and wide used classification technique, is repeatedly applied on child pose and gesture approximation (21). The LSTM (Long Short-Term Memory network) (22), a category of recurrent neural network which surpasses to typical consecutive information, making it particularly effective for capturing child pose and gesture info from skeleton key points.

One approach (23), utilizes a stateless Input data generator that adapts dynamically to varying video lengths, while the other investigates the stateful ability of recurrent neural networks, particularly utilized in HAR. This approach enables the prototypical to gather to extract discriminatory designs from prior frames while preserving reminiscence integrity.

In our study find the new approaches towards the physical exercise for child YK(drill) with the logistic regression proposes the recursive method for face landmark, pose landmark, left hand landmark, right hand landmark to select the 33 features.

Table 1. HPE Models using Deep Learning

Pose estimation model	Keypoint	Functioning	Approach	Single/Multi
Openpose	25	Real-time	Bottom-Up	One person, Many Person
Mask-RNN	11	Real-time	Bottom-Up	Many Person
AlphaPose	12	Real-time	Top-Up	Many Person
Deepcut	17	Real-time	Bottom-Up	Many Person
Iterative error Feedback	13	Real-time	Top-Up	One Person, Many Person
PoseHeatNet	33	Real-time	Top-Up	One person

3. Methods and Materials

This system is grounded on apprehending the video frame and detection the correct human action using machine learning approach. Mock drill(kavayat) proposed system designed with the help of computer vision to simulate real-world states for evaluation and estimation purposes. This system leverages computer vision techniques to monitor and recognize actions during drills, allowing actions for automatic assessment and its performance metrics.

3.1 Method

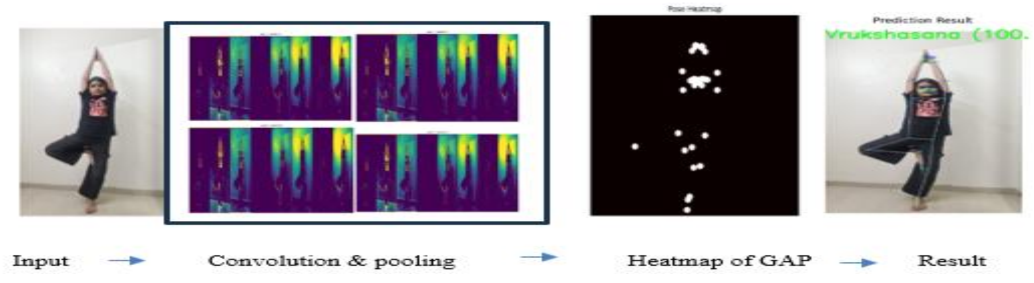

Figure 1. System Framework

This study divided in to a two-stage framework. First, a PoseHeatMap-based CNN is used to extract rich spatial representations of skeletal key points. Every key point is (x,y,z) and confidence score and to reduce variability across subjects and camera positions, coordinates are centred relative to a reference joint (e.g., hip centre) and scaled to maintain consistent proportions. Key points with low detection confidence are down-weighted or excluded to minimize the influence of noisy or occluded landmarks. The processed key points are transformed into PoseHeatMap, which encode spatial relationships between joints, providing rich information about posture and body configuration. These heatmaps as input to the CNN for feature extraction. This CNN acts primarily as a *feature extractor* rather than a classifier. Second stage, A Convolutional Neural Network is applied to PoseHeatMap to extract high-dimensional spatial features. Significantly, the CNN functions primarily as a feature extractor, rather than performing classification. This design leverages the CNN's ability to capture intricate spatial relationships among joints while keeping the classification layer flexible. These extracted features are fed into a Logistic Regression (LR) model that performs the final classification. This study deliberately selected Logistic Regression at decision layer as it is lightweight, understandable, and suitable for real-time deployment in classroom or low-resource environments. Thus, while CNNs provide the necessary expressive features, the novelty of our work lies in coupling them with a simple but effective Logistic Regression classifier, contrasting with prior studies that rely solely on deep end-to-end CNN pipelines.

Key point detection of each key point returns normalized coordinates like n belongs to

$$p_n = (x_{n,}y_n,z_n) \in [0,1] \ (1)$$

In equation (1)

xn: horizontal (width-wise) location

yn :vertical (height-wise) location

zn: relative depth or distance from the camera

∈[0,1]: These coordinates are normalized:

Instead of being in pixels they're scaled between 0 and 1 with respect to the image frame size.

For example:

xn=0 means distant left, xn=1 means distant right

yn=0 means top, yn=1 means bottom

zn=0 is the closest depth; positive values indicate increasing distance from the camera.

Normalized coordinates mean the device independent works on different images which support real time application.

3.2 *Input Representation*

Each pose is represented by a set of N = 33 2D landmarks:

$$P= \begin{bmatrix} x1 & y1.... & c33 \\ x2 & y2.... & c33 \\ x33 & y33... & c33 \end{bmatrix} \in \mathbb{R}^{33*3} \quad (2)$$

Where (x_i, y_i) coordinates in 2D of the i[th] joint and c_i is the confidence score from MediaPipe.

To make the model invariant to different body scales and positions:

- Centering (around mid-hip or nose joint):

$$x_i' = x_i - x_{center} \qquad y_i' = y_i - y_{center} \qquad (3)$$

- Scaling (dividing by distance between shoulders or hips):

$$x_i'' = \frac{x'i}{s} \qquad y_i'' = \frac{y'i}{s} \qquad (4)$$

3.3 *CNN-Based Pose Classification*

The CNN processes the normalized keypoint matrix or heatmap through a few convolutional layers:

$$F_{i=Conv1(P\ Or\ H),}\ F_2{=}Conv2(F1),.........$$

$$Z = Faltten(F_n) \rightarrow Dense\ Layer \rightarrow y^{\wedge} \quad (5)$$

Where $y^\wedge \in \mathbb{R}^c$ $y^\wedge$ is the predicted class probability vector over C poses.

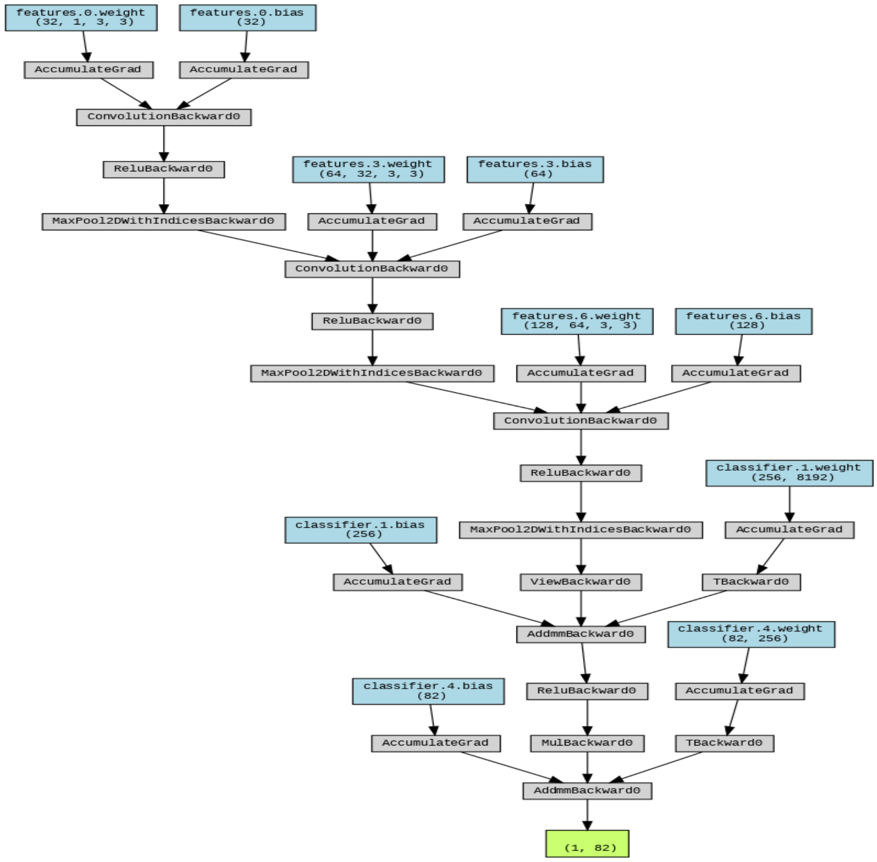

Figure 2. Layer of convolution

Table 2. Comparison of the dataset used

Dataset	MediaPipe Keypoints Applied	Labels	Pose Type	Expected Accuracy
COCO (Keypoints)	Use MediaPipe on each person	Different poses like standing, sitting etc.	Many person, real-world	70–85%
MPII	Use MediaPipe on each person	Different poses (e.g., standing, sitting lying, walking)	One person	75–90%
Custom Yoga (e.g., Yoga-82S)	Now works with MediaPipe	Yoga postures	One person	85–95%

It shows that the dataset and their comparison, in this system, uses the mediapipe with 33 key points and a light CNN with GAP classification, which is designed for only one person pose classification, such as yoga and kawayat datasets created in real-time using the mediapipe. For each frame, match with key points k(n). Where k(n) represents the function of the pose estimation technique, which extracts coordinates (x, y, z) of thirty-three joints for each frame fy. For a particular frame fy the xi, yi, zi coordinates match with every frame (fy) (x, y, z) a particular video Vn. This gives an accuracy of 92 to 99 % as it varies from conditions like clean and clear landmarks accuracy is 99 % but for bad and poor landmarks, it gives 70 to 85 %. The performance of the proposed PoseHeatMap-CNN combined with Logistic Regression was assessed using commonly adopted classification metrics, including accuracy, precision, recall, and F1-score. These measures offer a comprehensive evaluation of the model across all yoga pose categories. Metrics were calculated for each class and then averaged using the macro-average approach to reflect overall performance. The results achieved an accuracy of 94.6%, with mean precision, recall, and F1-score values of 94.2%, 93.8%, and 94.0%, respectively.

3.4 Dataset

A total of 120 child participants (Table 3) were enlisted for this work from local schools. All participants were healthy and free from musculoskeletal or neurological conditions that might interfere with physical activity.

Table 3. Age and gender distribution vs number of participants

Age	No. participants & (%)	Gender	percentage
6 to 8	42(35%)	Male	62(51.7%)
9 to 11	48(40%)	Female	58(48.3%)
12 to 14	30 (25%)		

3.5 Data Collection Protocol

The data acquisition was conducted in a controlled indoor setting to reduce background and illumination variability. A Logitech C920 HD Pro RGB camera was employed, capturing videos at a resolution of 1920 × 1080 pixels and a frame rate of 30 fps. To avoid visual distractions during landmark detection, recordings were performed against a plain, texture-free wall with uniform ambient lighting that minimized shadows. Participants were positioned approximately 2.5–3.0 meters from the camera, with adjustments made for individual height to ensure the entire body was within view. The camera was fixed on a tripod at a height of around 1.2 meters, corresponding roughly to the participants' midline.4.3 Children were instructed to stand on a marked spot on the floor to maintain a consistent distance and positioning relative to the camera. Each participant was asked to perform a predefined set of yoga postures and kavayat drills (10 exercise classes in total). Movement was performed three times per posture/exercise to account for intra-participant variability. Teachers and supervising staff ensured that exercises were conducted safely, with adequate rest between repetitions. Participants wore comfortable sports attire that did not obstruct joint visibility.

3.6 Limitation

Even though the new PoseHeatMap combined with Logistic Regression demonstrates encouraging results for recognizing children's yoga and Kavayat activities, certain constraints cognizance. The training dataset is comparatively small, which may reduce the system's capability to generalize to a wider variety of children or exercise variations. The evaluation was carried out under controlled conditions, meaning factors such as poor lighting, background clutter, and occlusion were not fully addressed. This condition is common in real-world environments and may influence recognition accuracy. Moreover, Logistic Regression, while efficient and easy to interpret, may not capture highly complex, non-linear movement patterns as effectively as more advanced deep learning methods. Finally, the study does not include long-term testing, so it remains unclear how well the system adapts to natural variations in children's growth, posture, and consistency over time.

4. Results

A built-in camera of a laptop or connected externally is used to capture live pose sequences. The system processes these video inputs in real time and classifies activities by comparing detected skeletal patterns with the reference dataset. In addition to identifying the performed action, the system evaluated pose correctness with an accuracy between 97% and 99% across all trained classes.

The new model PoseHeatNet, which uses grayscale heatmaps generated from 33 anatomical landmarks extracted via MediaPipe, achieved stable convergence within 10 training epochs. Loss curves and validation metrics confirmed that the network effectively captured discriminative features from joint patterns. The model consistently achieved above 90% classification accuracy on clean datasets, demonstrating the robustness of skeletal landmark representations over raw RGB images in controlled environments such as yoga and kavayat drills.

For the detection of human action recognition, a camera is utilized. The system processes a live dataset to detect human poses. The camera employed can be embedded in a laptop, or an external camera can be used. It identifies the output of live poses by comparing them with the original dataset. Additionally, the system provides the name of the activity being performed. The model on Human Action Pattern Recognition in Physical Training for children's environments explored the integration of progressive technologies, with computer vision, deep learning, to recognize and classify diverse human actions within

Physical Training settings. Leveraging this system, which is accomplished of real-time action recognition, provided feedback in the form of the forecast workout name. By positively developing an intelligent system accomplished of identifying humanoid actions in real time, the system has paved the way for transformative applications in the fitness industry. This technology not only improves the physical training involvement for entities but also enhances the capacity for applications in sports training, healthcare. The executed system is for a solo user at a period by using logistic regression with MediaPipe Framework with state of art on training and testing model on eight different classes with 0.99 accuracy, which is only for children. This system is a combination of technology-guided human movement. In the future, integrating temporal modelling like Transformers for video-based classification.

Figure 3. Pose in real-time

5. Discussion and Conclusion

The results indicate that structured skeletal heatmaps provide a reliable, background-invariant representation of body movements. Unlike methods relying on deep CNN architectures or raw image frames, our lightweight CNN layers achieve efficiency while maintaining strong accuracy, making the model suitable for real-time positioning on schoolroom or home devices. A key distinction of this work is the utilization of all 33 MediaPipe keypoints, whereas benchmark datasets such as MPII and COCO typically use only 16–17 landmarks. This richer representation improves the precision of pose characterization. Furthermore, the focus on children's physical activities introduces practical value, since children's movements are often irregular, and body proportions vary significantly compared to adults. By addressing these challenges, our approach offers an accessible solution for digital yoga training and school-based physical activity monitoring, while remaining computationally lightweight. A pilot usability study with children demonstrated positive engagement and responsiveness to real-time corrective feedback. Teachers noted that the system effectively reinforced proper form and complemented traditional supervision. These observations highlight the practical and educational value of the approach, beyond its quantitative performance metrics. Kavayat was included because it is a culturally relevant set of children's physical drills designed to promote health, coordination, and discipline. Unlike typical yoga postures, Kavayat involves dynamic, repetitive movements that emphasize rhythm, whole-body coordination, and basic strength, providing a complementary biomechanical challenge. Potential deployment scenarios for the proposed system in real classroom settings. The system can be implemented using standard webcams, tablets, or mobile devices, enabling real-time monitoring without requiring specialized hardware. We also address practical considerations, including cost-effectiveness, safety, and privacy: all video data can be processed locally to avoid sharing sensitive information, and the lightweight nature of the PoseHeatMap system ensures low latency and minimal computational requirements, making it suitable for widespread educational use. The system's real-time corrective feedback aligns with the cognitive and associative stages, helping children recognize errors, refine movements, and gradually achieve autonomous execution. By linking the performance improvements observed in Yoga and Kavayat exercises to established feedback-based learning principles, the discussion now provides a stronger conceptual grounding and highlights the pedagogical relevance of the PoseHeatMap system.

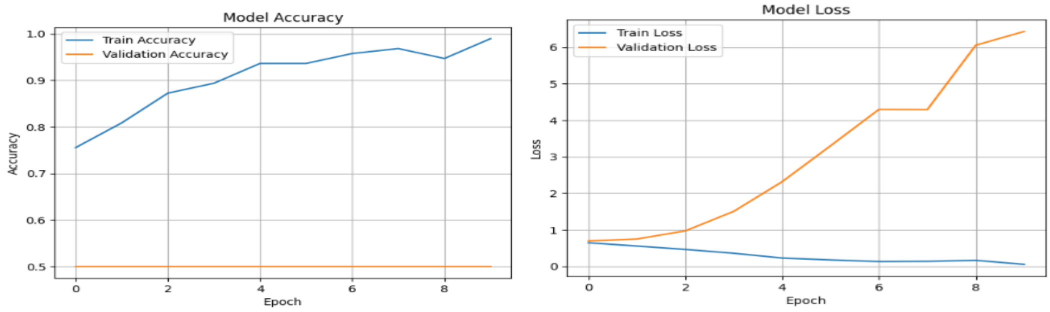

Figure 4. Model Accuracy and loss

Human action recognition in children physical training environments leveraging advanced technologies such as computer vision and deep learning to identify and classify a variety of pattern associated with human actions. Using a single-user setup with Logistic Regression and the MediaPipe framework called PoseHeatMap, this system achieved an accuracy of 0.99 across eight exercise classes, demonstrating the feasibility of intelligent, responsive feedback in physical training contexts. Despite these encouraging results, certain limitations remain. The current implementation is restricted to single-user scenarios and may encounter difficulties in multi-user or crowded environments. Furthermore, the robustness of the model under varying lighting conditions, occlusions, and different camera perspectives has not been extensively evaluated.

Future research can address these challenges by extending the system to support multiple users simultaneously, integrating multimodal inputs such as depth sensors or wearable devices, and optimizing the model for deployment on mobile or embedded platforms. Incorporating sensor fusion techniques, including inertial measurement units (IMUs), could improve motion capture accuracy and reliability. Additionally, embedding gamification elements may enhance engagement and motivation among children, making the system more interactive while preserving real-time monitoring capabilities. These advancements would increase both the pedagogical impact and practical applicability of the system, enabling safer, personalized, and interactive physical training experiences for children and adults.

Authors' Contributions

VB contributed to the concept, methodology, data collection, original draft writing, and review and editing of the manuscript. PA contributed to the concept, formal analysis, visualization, supervision, and review and editing.

Declaration

In order to correct and improve the academic writing of our paper, we have used the language model ChatGPT.

Transparency Statement

Data are available for research purposes upon reasonable request to the corresponding author.

Acknowledgments

The authors would like to thank their Family and friends, and thank you to Vishwakarma Institute of Information Technology for providing resources and support.

Declaration of Interest

The authors report no conflict of interest.

Funding

According to the authors, this article has no financial support.

Ethical Considerations

The study protocol adhered to the principles outlined in the Helsinki Declaration, which provides guidelines for ethical research involving human participants.

References

1. World Health O. Global Recommendations on Physical Activity for Health: WHO Press; 2010. 3_https://doi.org/DOI}
2. Cao Z, Hidalgo G, Simon T, Wei SE, Sheikh Y. OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part

Affinity Fields. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2021;43(1):172-86. [PMID: 31331883] [DOI]
3. Fang HS, Xie S, Tai YW, Lu C. RMPE: Regional Multi-Person Pose Estimation. Proc IEEE International Conference on Computer Vision (ICCV). 2017:2334-43. [PMCID: PMC5530199] [DOI]
4. Biswas S, Das P, Sinha S. Yoga posture recognition by extracting key joint points using OpenPose. Procedia Computer Science. 2020;167:2360-7.
5. Haque A, Milstein A, Fei-Fei L, editors. Measuring the Machine Learning Gap in Exercise Feedback. Proc Machine Learning for Healthcare Conference (MLHC); 2019. 3_https://doi.org/DOI}
6. Martinez J, Hossain R, Romero J, Little J, editors. A simple yet effective baseline for 3D human pose estimation. Proc IEEE Int Conf Comput Vis (ICCV); 2017. 3_https://doi.org/DOI}
7. Lanke WP, Bordoloi GR, Shrivastava A, Deshmukh SV, editors. Assessment of human activity recognition based on feature extraction prediction accuracy. Proc 4th Int Conf Intelligent Engineering and Management (ICIEM); 2023: IEEE. 3_https://doi.org/DOI}
8. Waheed S, Amin R, Iqbal J, Hussain M, Bashir MA. Deep learning-based framework for automated human action recognition and classification. Proc 4th Int Conf Comput Math Eng Technol (iCoMET). 2023. [PMCID: PMC10208255] [DOI]
9. Jo B, Kim S, editors. OpenPose, PoseNet, and MoveNet model comparison for mobile pose estimation2022. 3_https://doi.org/DOI}
10. Shah M, Gandhi K, Pandhi BM, Padhiyar P, Degadwala S, editors. Real-time and offline human pose estimation using computer vision and deep learning. Proc 2nd Int Conf Appl Artif Intell Comput (ICAAIC); 2023. 3_https://doi.org/DOI}
11. Chaudhary I, Singh NKT, Chaudhary M, editors. Real-time detection of yoga poses using OpenCV with MediaPipe. Proc 4th Int Conf Emerg Technol (INCET); 2023: IEEE. 3_https://doi.org/DOI}
12. Esmaeili B, Akhavanpour A, Bosaghzadeh A, editors. Surya namaskar: advanced real-time pose recognition and correction for healthcare. Proc Int Conf Mach Vis Image Process (MVIP); 2022: Univ. Tehran, Iran. 3_https://doi.org/DOI}
13. Chaudhari A, Dalvi O, Ramade O, Ambawade D, editors. Yog-guru: pose correction in yoga using deep learning. Proc Int Conf Commun Inf Comput Technol (ICCICT); 2021: IEEE. 3_https://doi.org/DOI}
14. Babanne V, Agarkar P, editors. Human pose estimation for yoga and kavayat drills: a comparative analysis. Proc Int Conf Emerg Smart Comput Inform (ESCI); 2024: IEEE. 3_https://doi.org/DOI}
15. Bai H. Training method for VideoPose3D leveraging action recognition concepts. Proc Int Conf Signal Process Mach Learn (CONF-SPML). 2021. [PMCID: PMC8173576] [DOI]
16. Gera R, Ambati KR, Chakole P, Cheggoju N, Kamble V, Satpute VR, editors. Activity classification in video streams using CNN with ConvLSTM. Proc 2nd Int Conf Paradigm Shifts Commun Embedded Syst Mach Learn Signal Process (PCEMS); 2023: IEEE. 3_https://doi.org/DOI}
17. Tsai MF, Chen CH. STV-GCN: spatial temporal variation graph convolution for skeleton-driven emotional action recognition. IEEE Access. 2021;9:1-12. [DOI]
18. Ji J, editor Action recognition based on skeletal pose estimation. Proc 2nd Int Conf Comput Data Sci (CDS); 2021: IEEE. 3_https://doi.org/DOI}
19. Lee M, Kim SB. Open-set human activity recognition via sensor-based representation learning with mixup triplets. IEEE Access. 2022;10:1-12. [DOI]
20. Ketsoi V, Raza M, Chen H, Yang X, editors. Secured human-computer interaction via hand action recognition. Proc IEEE Int Conf Syst Man Cybern (SMC); 2022: IEEE. 3_https://doi.org/DOI}
21. Patel HR, Doshi JT, editors. Recognizing human activities in dark videos using deep models. Proc Int Conf Artif Intell Mach Vis (AIMV); 2021: IEEE. 3_https://doi.org/DOI}
22. Sánchez-Caballero Á, Fuentes-Jimenez D, Losada-Gutierrez C, editors. Depth video-based recurrent networks for real-time human action recognition2022: Springer. 3_https://doi.org/DOI}
23. Ryu J, Patil AK, Chakravarthi B, Balasubramanyam A, Park S, Chai Y. Real-world action recognition using angular features for fine-grained unit actions. IEEE Access. 2022;10:1-12. [DOI]