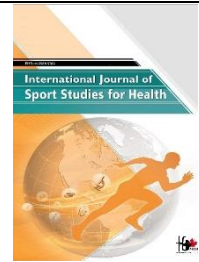


International Journal of Sport Studies for Health

Journal Homepage






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

Vanita. Babanne¹ , 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

E d i t o r	R e v i e w e r s
Karim Chamari ¹ 	Reviewer 1: Yaghob Badriazarin ¹ 
Tunisian Research Laboratory "Sports Performance Optimization" National Center of Medicine and Sciences in Sports (CNMSS), Tunis, Tunisia pademil@hotmail.com	Associate Professor of Sport Sciences, Tabriz University, Tabriz, Iran. Email: badriazarin@tbzmed.ac.ir
	Reviewer 2: Masoud Mirmoezi ¹ 
	Department of Physical Education and Sport Sciences, Islamic Azad University, Central Tehran Branch, Tehran, Iran. Email: massoudmirmoezi@live.com

1. Round 1

1.1 Reviewer 1

Reviewer:

The paragraph beginning with "Proper execution of physical exercises is essential for children's physical and mental well-being" repeats concepts already discussed in the preceding paragraphs. Consider integrating this material earlier to avoid redundancy and improve cohesion.

The statement "Research focuses on the to enhance the movement analysis and correcting posture with advanced techniques" is ungrammatical and unclear. Rewrite this to articulate the explicit objectives of the research in a single precise sentence.

The introductory problem framing spans multiple paragraphs without a clear narrative structure. Consider reorganizing into: (1) problem context, (2) limitations of existing systems, (3) research gap, and (4) statement of contribution.

The description of how the PoseHeatMap is generated from MediaPipe landmarks is vague. Specify whether Gaussian heatmaps, binary masks, or channel-encoded spatial maps were used.

The study mentions that each participant performed "10 exercise classes," but there is no description of class distribution, number of frames per class, or whether the dataset is balanced. This information is essential for evaluating the reported accuracy.

Data were collected in a highly controlled environment (plain wall, uniform lighting). Discuss how these controlled conditions may limit generalization to real-world classrooms with clutter, uneven lighting, and multiple children.

The Results section reports accuracies such as “97–99%,” “above 90%,” and “0.99 accuracy,” but does not provide confusion matrices, per-class accuracy, or variance across participants. These metrics are necessary for assessing robustness.

The statement “the system evaluated pose correctness with an accuracy between 97% and 99% across all trained classes” is extremely strong but unsupported by detailed evidence. Include per-pose precision, recall, and examples of misclassifications.

The Discussion mentions that children benefited from real-time feedback but does not explain how this aligns with known motor learning theories. Incorporating such theoretical grounding would strengthen the argument.

Terms such as “PoseHeatMap,” “PoseHeatNet,” and “lightweight CNN” appear inconsistently. Ensure terminology is standardized throughout the manuscript to avoid confusion.

Author revised the manuscript and uploaded the updated document.

1.2 Reviewer 2

Reviewer:

Although the Introduction mentions children’s unique anatomical characteristics, the manuscript does not sufficiently justify why existing adult-focused models cannot be adapted. Provide more technical reasons for child-specific pose variability.

The Literature Review is written as a long list of studies without thematic grouping. For example, sensor-based studies, CNN-based pose estimation, and skeleton-based methods are mixed together. Grouping them into subsections would significantly enhance readability.

Sentences such as “It is deep learning based computer vision technologies. The cameras are required for the capturing the video.” contain grammatical errors and do not clearly contribute to the argument. They need rewriting for clarity and academic tone.

The review merely reports what each paper did but does not critically analyze limitations of prior methods or how they motivate the current study. Add a paragraph synthesizing the gaps in existing literature.

Although the proposed system uses Logistic Regression at the decision layer, the Literature Review does not discuss why alternative classifiers (e.g., SVM, Random Forest, Gradient Boosting) were not considered. This weakens the justification of the model choice.

The description of the CNN functioning “primarily as a feature extractor” is repeated twice in the same section. Consider merging these sentences to improve conciseness.

In the Methods section, the formulas for centering and scaling are presented without proper notation or explanation. For instance, the meaning of each variable (x_i , y_i , c_i) should be clearly defined before being used in equations.

The manuscript states that “a few convolutional layers” were used but provides no architectural details (number of layers, kernel sizes, activation functions, pooling strategies, or dropout). Include a full architecture table to support reproducibility.

Author revised the manuscript and uploaded the updated document.

2. Revised

Editor’s decision after revisions: Accepted.

Editor in Chief’s decision: Accepted.