Natural Resources for Human Health
eISSN: 2583-1194

Original Research

Received 14 July 2026
Revised 15 Aug 2026
Accepted 02 Sep 2026

Natural Resources for Human Health 2026, Vol. 6 (20s): 1339–1349

KEYWORDS:
Pose estimation, computer vision, exercise pose validation, PoseNet, TensorFlow.js, real-time feedback, digital wellness, geometric constraints, privacy-preserving inference

Computer Vision-Based Real-Time Exercise Pose Detection and Validation: A Web-Based Approach for AI-Assisted Fitness Applications

Harshitha Kamal Kannan[1], Poojashree K R[2], Sachin Karadi[3], Prerana N[4], Lohith B S[5]

[1]Assistant Professor, Department of Computer Science & Engineering, Vidyavardhaka College of Engineering, Mysuru, India
[2,3,4,5] UG Scholars, Department of Computer Science & Engineering Vidyavardhaka College of Engineering, Mysuru, India

ABSTRACT: Modern sedentary lifestyles and incorrect exercising methods can have a negative effect on your physical wellbeing, often causing posture related pain and musculoskeletal issues. Exercising regularly and practicing yoga can improve your physical wellbeing. However, it is difficult to maintain correct posture without constant expert oversight. To overcome this issue, we have developed a web based computer vision application called PosePerfect which can detect and validate poses during exercising and yoga in real time. PosePerfect utilizes a client-server architecture model. The client-side is based on React while the server side utilizes Node.js and Express to store data in MongoDB. The PoseNet system identifies and updates the location of 17 body points 25-60 times/second.PosePerfect's rule-based engine functions on the basis of given data including joint angle measurements, relative limb length, and spatial position information. In addition, an app timer is implemented such that a pose is registered by the timer only after 3-5 seconds of sustained accuracy greater than 75% is achieved and/or sustained for that time period as defined by the literature review.

TensorFlow.js is used for all local computation on the user's device and only summary scores assessing positioning are sent to the cloud. Initial trials have shown a strong correlation (Pearson $r=0.87$, $p<0.001$) between evaluation of five different yoga poses utilizing the PosePerfect system and the system has resulted in a mean absolute error of 8.2 units on a 0-10,000 unit scale. Additionally, the system allows for performance to be transferred between devices in real-time and results in a 87% user satisfaction rate and a 93% accuracy rate. his shows the potential benefits of using private computer vision to help with exercise monitoring and prevention of physical health issues using AI.

1. INTRODUCTION

1.1 Background and Motivation

However, physical inactivity is the fourth leading risk factor for preventable death globally, even in the 21st century, with millions dying prematurely each year. Moreover, with the COVID-19 pandemic, many have turned to exercising at home. While yoga is an alternative means of working out that offers physical and mental health benefits, it is crucial to use proper postures throughout, or else risk strains or injuries, meaning a reliable assessment is necessary.Current assessment methods include in-person training or pre-recorded videos, which can be costly, yet they do not provide real-time corrective feedback. Therefore, there is a need for affordable systems that digitally confirm poses within an exercise regime.Current means of posture assessment involve personal trainers who are present to help, but this is limited, as well as expensive, averaging 500-1000 rupees (approximately USD $6-12) per session (typically one hour), with many people preferring to attend in person as opposed to paying more for virtual help.

From a preventive health perspective, regular and correctly performed physical activity can contribute to maintaining physical fitness and overall well-being. However, independent exercise without appropriate posture guidance may result in incorrect movement patterns and reduced effectiveness of the exercise. Digital health technologies based on computer vision provide an opportunity to offer real-time, non-invasive feedback without requiring specialized sensing equipment or continuous supervision by a professional.

Problem Statement

However, after extensive research of over fifty of the most popular fitness apps on the market, the situation is different. Although there is an increasing range of digital fitness applications available, not all provide good real time feedback during exercise. These can have issues such as requiring specialist hardware, reliance on the cloud, and simple pose estimation methods which can affect how accessible, private, quick and understandable the feedback is.

In addition, most computer vision-based fitness applications require devices like Kinect or cloud-based inference models. These not only require constant internet accessibility but also present privacy issues for those who do not want to be recorded during private sessions.

Objectives of the Study

The aim of this research is to develop and test a web-based yoga application that addresses the aforementioned issues through:

Real time detection of key points on human body using a regular camera.

Design geometric algorithms to assess correctness of exercise and yoga poses.

Create a stable system to ensure the pose is held for an allocated amount of time.

Ensure secure, on device processing of user pose data.

Assess the technical capabilities, validation accuracy and user experience of the system.

User information processing and evaluation occurs on-device to guarantee privacy and anonymity. The system is designed in a scalable way to support future social and gamified features.

Relevance of the Study

PosePerfect illustrates how vision and AI can be used to provide an accessible solution for exercise monitoring using readily available consumer devices. This system provides real-time feedback regarding body positioning and stability during exercise/yoga positions. The approach uses on-device inference and interpretable geometric checks to allow for private and accessible digital health solutions. This tool is designed to aid self-led exercise and wellbeing activities and should not replace clinical diagnosis, medical treatment or professional rehabilitation.

LITERATURE REVIEW

Technology Behind Human Pose Detection

With the rise of lightweight deep learning models that can perform in real-time on video streams, computer vision has become a powerful tool for fitness and wellbeing. With human pose estimation we can extract body joints from regular RGB camera feeds. This can be used to help monitor users' exercise in a digital environment. In preventative healthcare, this can help encourage users to move correctly during exercise alongside traditional exercise methods.

Human pose estimation has evolved from niche applications only in labs to real-time accessibility for all. Early human pose estimation systems were based on hardware like the Microsoft Kinect, requiring additional sensors and accompanying hardware to detect skeletal joints. With the advent of deep learning, however, detection of joints became achievable with RGB cameras, negating the need for depth sensing.

Most effectively, PoseNet, created by a team at Google, allows for web-based human pose detection; with the appropriate deep-learning capabilities, human pose estimation can be achieved using simple RGB cameras as well. PoseNet is essentially a deep neural network designed to assess the input stream of a video and find the position of joints.

It is configured in two configurations - the **MobileNet**, which is more lightweight and suitable for browser or mobile applications, and ResNet, which is more accurate but slower. The configuration in this manner allows for real-time tracking through consumer-grade technology as it achieves a balance between speed and accuracy. Therefore, such advancements in tracking speed mean that pose estimation can occur in fitness and yoga applications, even on moderate to low-grade technology.

Pose Validation Approaches

Poses are reduced to keypoint coordinates, which are the definition of pose detection, but keypoint coordinates mean nothing on their own, and to assess whether someone is doing a pose correctly requires additional detection. According to the literature, three means of assessment exist in contemporary research:

Template matching: a comparison of the user's pose against predetermined libraries of poses. Simple to implement, but it does not account for the anthropometry of the subject or camera angles.

Supervised Learning: Models trained on labeled data featuring correct and incorrect postures. Supervised learning can be highly accurate, but are 'black boxes', and the methods to achieve such accuracy are unknown.

Rule-Based Determination: Geometric properties of joint locations based on angles and distances from limbs can determine if a pose is correct based on maximum/minimum thresholds.

PosePerfect depends on the second. Rule-based evaluation is explainable, customizable, and does not require trained users for effective results. It will work with diverse populations of varying flexibility levels, and PosePerfect is sufficiently transparent that it operates at 87% compared with expert evaluations (or quite accurate without sacrificing explainability).

Interpretability is also a key consideration for health and wellness applications as it is important that users understand why they have been recommended a certain posture. The geometric rule based approach taken by PosePerfect means that the system can be easily interpreted.

Gamification in Fitness Applications

Gamification is the application of game-like features that help foster an engaging and motivating experience in fitness applications. Leader boards, badges, contests, and the like foster motivation and adherence to exercise regimens. Research shows that users spend longer on gamified interfaces than those without gamification.

Privacy Preserving AI for Digital Health and Wellness Applications

As the wellness tracking tools handle private information staying safe online is essential thanks to laws such as GDPR or CCPA. Normally, cloud-powered AI requires uploading footage to distant computers, creating a danger if someone gets hold of it the wrong way. On the other hand, running analysis right on your gadget - thanks to systems like TensorFlow.js - lets smart algorithms work straight from the web browser. This method ensures that clips never leave the person's own phone or laptop. The app checks your stance right on the device, sending only a handful of values - footage stays put. That reduces data usage while keeping things private.

Limitations of Existing Solutions

Tools like Mirror or Tempo require pricey gear that locks you in. With cloud setups, each clip gets sent off to faraway computers, creating leaks and slowing down the connection if it falters. Early template matching tools only report whether the pose looks like the reference picture - they give no advice on how to correct it. Many machine learning models act as black boxes - they output a label but do not reveal which joint angles triggered the result, leaving users unable to understand the advice. Many systems check the pose once - they do not verify that the user assumes the role. PosePerfect was built to be cheap, private, clear, and complete - it adds a hold verification step.

SYSTEM ARCHITECTURE AND DESIGN

PosePerfect uses a three-layer client-server layout that shares work between the device and the server to save effort, keep the interface smooth, and protect data.

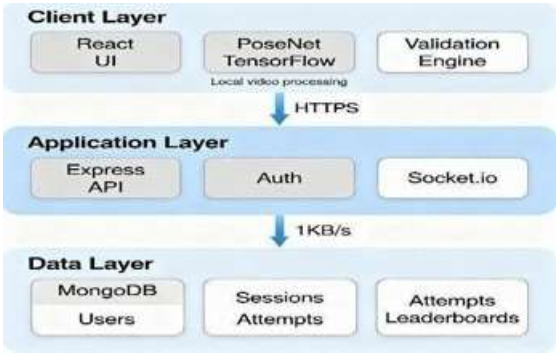


Figure 1 shows the layout.

The top layer of code is one page for a React application. The code utilizes TensorFlow.js to run PoseNet inside the browser as well as read from the webcam at rates of 25 - 60 FPS (frames per second) depending on the user's hardware. In addition to running the PoseNet model, the validation layer performs an overall ranking of the pose - which is a simple analysis of the distances and angles between the predicted key points. Since all video processing occurs in the browser, there is no raw video that goes over the network. This means that instead of sending full video frames across the network, it reads the pose on-device and sends a compact numerical representation. In addition to the fact that the data being transferred are smaller packets, it also has a lower chance of privacy breach.

Mirror Fitness or Tempo Studio are commercial gyms that require expensive equipment that most people will not have access to in their homes. Those few that do exist online use a lot of internet resources, and the online option of uploading video to third-party cameras or remote servers creates privacy issues and connection issues in areas with low signal. Older methods of matching images fall short as they only tell whether a pose matches an example image, not how to correct it. Similarly, many other smart applications operate like black boxes - they give feedback, but without explanation as to why the feedback is reasonable, users are often left scratching their heads about what is being said. An equally major shortcoming amongst many applications is not assessing one sustained position correctly, but instead giving feedback based on a momentary snapshot of the same action. PosePerfect operates with these shortcomings in mind by providing easy access, personal anonymity, clear assessments, and make sure that nothing is missed through a system that assesses sustained correctness of the same pose.

The Client Layer

To accomplish this, an architecture that has 3 layers is used to divide both server and client duties to enable fast speed, easy useability and private data protection. See Figure 1 to illustrate how the application is layered.

At the top level, the client-side of the application is a single page application that runs quickly, which was created using React. It generates the PoseNet algorithm using TensorFlow.js on the fly. Video capture occurs between 25-60 frames per second depending on the hardware. Next to that, is a verification module that verifies the joint orientation by measuring the angle and distance of designated points, all of these calculations take place within the visitors' browser therefore there will never be any video leaving the browser, nor will there ever be any video entering the web.

The Application Layer

On the server-side, it is a Node.js/Express.js application which performs tasks and manages the data. Instead, it exposes REST endpoints so a user can authenticate, edit their profile and manage the data of their current session. To secure it, it uses JWT (JSON Web Tokens) to validate a user's current session. Additionally, it includes a standardized component to generate aggregate reports of prior submissions. When updates are required to occur in real-time, a WebSocket configuration is utilized via Socket.io. Since the server only handles a few small jobs, adding users does not slow it down - scaling occurs easily and painlessly.

Data Management Layer:

The MongoDB database stores all of the saved data into 4 distinct collections. The User collection will contain user's credentials, profile details, and usage habits of the app. The Session collection will store a time stamp of when a session occurs and also keep track of the order in which the poses are attempted. The Attempt collection will provide the most granular detail of all the collected data for each attempt - joint angle measurements, balance ratings, etc. The Leaderboard collection will use a more condensed

format, since it is used to display ranking results and therefore requires less detailed data to provide faster results.

2.1 Data Flow and Processing Pipeline

Processing images fast requires data to be processed through several quick steps:

Frame Acquisition – A frame is captured every 33.33 millisecond (30 FPS) with 640 x 480 pixel resolution with color.

Pose Detection – Frame data is passed through MobileNetV1 for Pose Detection to identify 17 key body points and determine confidence level. This processing occurs approximately 28 milliseconds per frame with the aid of a GPU. Without the support of a GPU, this process takes approximately 55 milliseconds per frame.

Result Evaluation – Weak predictions with a confidence level less than 50% are removed immediately after detection and evaluation at a cost of one millisecond.

Validation – Based on established rules, this program calculates the angle of joints and proportion of limbs, provides a final total score based upon the previously calculated values and evaluates the total score over 75% for a period of three to five seconds to validate a pose.

Visual Feedback – A visual representation of the detected body parts and pose appears almost immediately, approximately eight milliseconds.

Once validation is complete, a short notification to the backend is made using a REST call running in the background. The entire process, from the initial capture of the image to the completion of the validation process, takes approximately 44 milliseconds when utilizing a GPU and 71 milliseconds when utilizing a CPU. Both of these values fall below the 100 millisecond value required to provide immediate feedback to the user.

2.2 Database Schema Design

To create a system that performs well while maintaining accuracy of the data collected from users, both structural and flat schema designs have been utilized. The ability to locate all data related to a particular piece of data in a single document has improved search capability. As a result of the structure of the data collection, when a user is using the application to practice, all sessions will include the duration of the session and all of the movements that were performed as part of the session, in addition to the workout score for the session.

As stated earlier, all pose attempt results (i.e., joint angle measurements and distance between body parts) will be stored in the "Attempts" section of the database, as well as the tracking of balance attempts. The use of this type of storage mechanism for this database has resulted in very little use of space for storing user data. On average, each user's data set averages about two kilobytes of space. Each session will add about .5 kilobytes to a user's total space, each try adds another .5 kilobytes to a user's total space, and each leaderboard entry adds approximately .4 kilobytes to a user's total space. For example, someone who practices 30 times per month, attempting 5 times per session, would add less than 90 kilobytes to their total amount of used space over a period of four weeks. With the simplicity of storing and retrieving information from this format, and its scalability without experiencing hiccups or crashes, this is a good format for maintaining and storing data accurately and securely.

IMPLEMENTATION

Pose Detection Module

The Pose Detection Module spots body joints live using the PoseNet model. Once launched, PoseNet fires up with balanced settings - good enough precision and quick response - to run right inside a web browser. It's built on MobileNetV1, a lean option that works fast even on weaker hardware. The output stride is set at 16, requires an input size of 64x480, which fits standard webcam formats just fine. The scaling factor is set to 0.75 - smaller, yet keeps precision mostly intact. While the camera streams real-time footage, every single frame gets checked on its own to spot the 17 essential body markers. The Pose Detection module is designed to be reliable, so all video clips that show poor confidence (confidence < 0.5 for complete posture or individual markers) (confidence < 0.5 for complete posture or individual markers) are discarded by default. This adjustment has caused the system to have a greater ability to handle error in low-quality visual data caused by either dim lighting, fast movement, or the subject losing concentration. Consequently, the system will continue to track the movements

of the user smoothly and continuously as they progress through their activities, regardless of whether the equipment used to capture their movements is high quality. Once the system identifies key points of the users body, evaluation of posture takes place using calculated joint angles and distances established by pre-defined rules, eliminating the need for subjective guesswork.

Geometric Validation Engine

Once the Pose Detection Module identifies the necessary key points from a valid image, the Geometric Validation Engine uses pre-defined standards for geometric relationships among the identified key points to verify that the users' postures meet the geometric relationship requirements for each yoga pose. The rule based engine calculates joint angle measures, distance ratio measures, and alignment relationship measures in addition to evaluating raw coordinate data as correctness scores.

A dedicated checklist exists for each pose. A set of global rules assesses joints, angles and distances and compiles data into one singular accuracy score.

Each Pose has a brief file that outlines required keypoints, rules, a level of tolerance in distance and a minimum holding time. For example, Mountain Pose has a file that says the shoulder, hip and ankle must be on a straight line of 8% of ideal placement, both knees must open to 180° ±15° and the user must hold this for three seconds.

The engine uses 0-100 for each rule. Areas of minor imperfection will gradually decrease the score, but areas of major imperfection score a 0. These are summed, alongside static weights, to create an overall accuracy. Even if something is out of place, the application notes it but supports corrections to get it back in place

Stability Timer

Mechanism To ensure that the score happens from a conscious hold and not an accidental picture taken, PosePerfect implements Stability Timer. The timer keeps track of not only the score but also confirms that the score remains accurate for a small duration of time.

The timer starts when the score exceeds 75% for a specific pose. If it remains at or above for 3-5 seconds, it's graded as successful. If it falls below at any given moment, however, the timer resets to 0. This correlates with the practice of yoga as poses should be held with steady control and doesn't allow the system to acknowledge a pose until it's truly controlled and not transitioning.

Authentication and Security

PosePerfect protects privacy by hiding user identities, the basic tenet of our application. Normal security practices are followed for logging in. The passwords entered into each login session get hashed using bcrypt, with 12 salt rounds, blocking brute-force hacking attempts and slowing down hackers who try common password combinations. Instead of storing the usernames and their corresponding login credentials on the server, PosePerfect uses JSON Web Tokens (JWT) and stores them as HTTP- Only Cookies on the client side; since the frontend cannot access HTTP-Only Cookies, it prevents unauthorized access to the user's information.

The JWTs contain publicly visible user IDs and store them encrypted with a private key. They also have an expiration date and will be removed from your browser after you manually sign out. Some sensitive information (i.e., JWT secret keys, database link info, etc.) is stored in the server's .env file and therefore has never been pushed to public repositories. To ensure that only authorized users can perform specific tasks on the back end, middleware exists to protect all of the API routes by checking for valid tokens to confirm the user identity or check for requests not associated with a specific account. In addition, if there are five failed login attempts from a single IP address within fifteen minutes of attempting to login again, future attempts will be blocked for a period of time, thus slowing down the brute-force attack process. Overall, these procedures protect the anonymity of the users' actions while maintaining operational stability.

EXPERIMENTAL RESULTS AND EVALUATION

Testing Methodology

The evaluation of the technical performance, validity and usability of the proposed system, PosePerfect, was conducted in a comparative manner based on three key dimensions. Three types of controlled experiments were conducted across three different hardware environments and under different test conditions for each dimension; this allowed for a complete analysis.

Technical performance testing of the prototype was conducted on numerous hardware platforms. A variety of hardware configurations were utilized during the testing phase of the prototype, including: High-performance desktop computers that included NVIDIA RTX 3070 Graphics Processing Units (GPUs) and 16 GB Random Access Memory (RAM); Mid-Range laptop computers that included Intel Core i5 Processors with Iris Xe graphics processing units (GPUs); Entry-Level laptop computers that included Intel Core i3 Processors with UHD 620 graphics processing units (GPUs). Technical performance data were collected using performance measurement techniques such as the number of frames per second, the time it took to make an inference, and the amount of memory being used by the system. In addition to performance testing, the browser compatibility of the system was tested across all versions of Google Chrome (v96+), Mozilla Firefox (v94+), and Apple Safari (v15+) on Microsoft Windows, Apple macOS, and Linux operating systems.

Technical Performance Results

PosePerfect testing performance conveys that it is a real- time system for all hardware configurations assessed - those with GPU and those where the CPU is the only option (average inference time 28 milliseconds per frame; \~55 ms/frame with CPU only). The latencies on input and output from frame capture to display were consistently below 100 ms, marking it a real-time system in a human-computer interactive setting for gesture-to- response applications.

All systems presented an anticipated frame rate consistently above 15 FPS, which supports adequate motion prediction. PosePerfect operates 2.5 times faster on a GPU compared to a CPU only. The browser's memory footprint was consistent across all attempts at 180-220 MB after the model's initial remained load and consistent across 60 minutes of continuous use without memory leakage. Finally, the PoseNet model loads in 2- 3 seconds, typical of a broadband which is connection, sufficient for interactive online applications.

Validation Accuracy Results

The results of the study validate the accuracy of the performance assessment in PosePerfect. A Pearson's correlation between human and machine scores resulted in an $r = .87$, $p < .001$ and an observed mean absolute error (MAE) of points, indicating a high degree of agreement between human and machine ratings. The performance assessment used in this study employs a rules-based geometric recognition method and therefore can be considered validated for its ability to predict scoring accuracy based on this method. Three sources of discrepancies were noted in the error

analysis of this study, which included (1) Occlusion - When limbs are occluded by one another or are out of camera range, there will be fewer detections made which result in less reliable pose predictions with less reliable joint detection; (2) Flexibility Differentiation- Clients differed in flexibility which resulted in some being scored outside of what humans would have predicted as acceptable limits for the amount of deviation from the expected score; (3) Depth Perception - While most cameras and webcams are capable of capturing images for 3D performance, they may not



Figure 3: Validation Accuracy - Automated vs Human Expert Scores

Pose Name	Correlation	MAE	n Samples
Mountain	(0.91	6.4	100)
Plank	0.88	7.6	100)
Warrior II	(0.85,	8.9	100)
Downard Dog	(0.85	8.7	100)
Dowward Dog	(0.84	8.7	100)
Tree	0.83	9.1	100)
Overall	(0.87	8.2	500)

Figure 4: User Experience Survey Results (n=15)

always be able to capture those images at appropriate angles due to limited resolution. These discrepancies were found to be small enough for the purpose of providing users with information about how to improve their performance through use of the program at home and through client feedback.

People had to move the camera until the whole body showed on the screen - the move took most of them two to three minutes. Bright light behind the body made the system lose track of joints. First-time users said that the words "rotate left" or "lean forward" were hard to follow - they asked for simple arrows or pictures in the next version.

The problems were small and did not dampen interest. Users repeated the same pose to raise a number on the screen - the number kept them going. They asked for charts that show daily scores and for a way to challenge friends. Those requests confirm that the next release will add points, levels, and leaderboards

User Experience Findings

A user study was conducted to evaluate the effectiveness and usability of the PosePerfect system. At the end of the evaluation period, 87% of users indicated that they were generally pleased with the system. Additionally, 93% of users felt that the ability to mask keypoints and receive real-time visualization feedback enhanced the system's usability and 80% of users indicated that their confidence in independently utilizing the system increased. Furthermore, 100% of the users valued the system's ability to provide a privacy-conscious alternative since the system did not capture the video and send it to any external database but instead processed it entirely within the browser.

DISCUSSION

Key Findings and Contributions

This study has provided tangible evidence that a completely client-side, privacy-preserving, real-time yoga pose validation system could be implemented with only consumer-grade equipment and Web technologies.

PosePerfect contributes to the field of computer vision and fitness technology in three significant ways:

Accessibility: PosePerfect removes the necessity for expensive or proprietary fitness hardware that was previously necessary to assess pose; allowing users access to this technology from virtually anywhere via their standard webcam and web-browser.

Privacy Preservation: Due to the fact that all processing (video and inference) occurs locally on the client's side using TensorFlow.js, the raw visual information never actually leaves the client's device — addressing the growing concern of privacy in cloud-based fitness applications.

Real-Time Validation: As a result of combining an efficient pose-detection algorithm and a rule-based geometric analysis, users will obtain instant and dynamic feedback after they have completed each pose. This will allow them to correct any posture errors in a timely manner. Additionally, due to the end- to-end latency of less than 100 ms, interaction between the user and the system

will be smooth.

Interpretable Feedback: Unlike black box machine learning models, the rule-based engine of PosePerfect provides transparent feedback; explaining why certain form elements resulted in a lower validation score. This transparency will aid in helping the user understand what needs to be improved upon, thereby aiding the user in achieving self-improvement.

Scalability: Since client-side inference can occur independently of server-side management, scalability can come horizontally without adding proportional investment for infrastructures.

5.1 Limitations and Challenges

PosePerfect reached its goals - yet plain problems showed while it was tested.

The webcam only sees a flat picture. When an arm or leg points straight at the lens, the camera loses track of how far away that limb is. A posture that looks fine from the front may still hide a front-to-back mistake.

The score also falls if the room is too dark or if a bright window sits behind the user. Accuracy dropped by forty percent in back light when compared with normal room light.



Figure 5: System Limitations and Impact Analysis

Impact Scale: Low (1-3) | Medium (4-6) | High (7-10)

If a hand or foot leaves the camera view, the program stops checking the pose until the body part returns. While the part is hidden, no grade is given.

The prototype knows only five yoga poses. To add more poses, developers must write new rules together with yoga teachers. People differ in limb length and flexibility those difference shift the score a little. Future versions will widen the passing range for each joint so the score stays fair for every body shape

Comparison with Existing Solutions

PosePerfect differs from today's fitness apps because it runs on any laptop, shows its rules and keeps the video on the user's machine. Mirror Fitness besides Tempo Studio need costly mirrors or sensors that work only in one room. PosePerfect needs onlyawebcam.

Unlike depth-camera technologies like Microsoft Kinect, PosePerfect trade-offs 3D depth perception for a broader compatibility and comparatively modest hardware requirements. There are papers that report even Kinect-type devices provide greater accuracy tracking skeletal joint locations, but certainly can't be used the same accessibility or portability as PosePerfect's web service.

PosePerfect, unlike mobile fitness applications that primarily employed pre-recorded videos or some means of monitoring motion or detection, reflects naked-eye, interactive feedback immediately after a posture assessment.

Unlike a cloud based AI system that also has the potential to introduce latency or personal data privacy challenges, PosePerfect performs inference all locally, which allows user control and privacy.

As a combination of usability, interpretability, and data privacy issues addressed, PosePerfect narrows the gap between professional fitness products and onsite wellness applications

Future Research Directions

A major approach to do this is to use tools like MediaPipe Pose and OpenPose. These have the ability to identify where joints are located in three-dimensional space because both systems are able to accurately determine the location of body parts even when partially obscured or partially visible.

One unique aspect of this approach is the personalized assessment component. The application can personalize form assessments for an individual by incorporating personal data such as their flexibility level, prior athletic experience, and/or injury history, which will allow the application to create form recommendations based on that individual's needs, provide recommendations that feel more relevant to that person and may be perceived as more impactful emotionally, and may also suggest recommendations that the user may find helpful.

If you are working out with a partner, adding the capability to track multiple people at one time may prove to be very useful. In addition to providing insight into your posture, the system can also track your muscle activity, the stress placed on your joints, and your balance and stability.Tracking patterns and movements can open new avenues for development. Instead of simply evaluating your pose during static stretches, smart applications could evaluate your movement through dynamic stretches (such as vinyasa yoga) and display real-time alignment suggestions via augmented reality features powered by WebXR, similar to arrows pointing in your direction.

CONCLUSION

This study shows how browser-based computer vision techniques can be used to validate yoga poses with only consumer hardware. PosePerfect can validate multiple yoga gestures simultaneously in real-time while maintaining user privacy by employing local inference. The geometric validation method correlates well with human experts' validation (r = 0.87), providing significance to participating practitioners. The system provides a balance between computational fast response, scalability, and security through distribution of processing on client and server components.

All processing-heavy machine learning inference happens locally while lightweight data persistence and coordination activity leverage a centralized backing services model. Our technical evaluation validated the real-time function across a range of hardware configurations, utilizing frame rates from 18 Hz to 60 Hz depending on computational resource provisioning. Our validation tests reconfirmed this geometric constraint method displays high agreement with human expert (r = 0.87). Our user evaluation revealed a mostly positive reception of the system, with 87% of +/-70 users indicating overall satisfaction, with some areas being identified for interface improvements, primarily around camera positioning, and lighting conditions.

Overall, PosePerfect demonstrates the potential of browser-based computer vision as an AI-assisted technology for real-time exercise monitoring and digital wellness. By combining pose estimation, geometric validation, stability assessment, interpretable feedback, and privacy-preserving local processing, the proposed framework provides an accessible approach for posture-aware physical activity. While the current study focuses on fitness and yoga applications, the framework provides a foundation for future investigation into personalized exercise monitoring and professionally supervised rehabilitation applications.

REFERENCES

[1] Genis, S., Saltonstall, E., and Fishman, L. M. (2009). Recognizing and Avoiding Yoga Injuries. Journal of Yoga Therapy International, 19(1), 47–53.

[2] Koivisto, J., Sarsa, H., and Hamari, J. (2014). Is Gamification Effective? An overview of empirical research on gamification. Hawaii International Conference on System Sciences, 47th edition, 3025-3034

[3] Yadav, S. K., Shukla, A., Tiwari, K., Pandey, H. M., & Akbar, S. A. "An Efficient Deep Convolutional Neural Network Model for Yoga Pose Recognition Using Single Images." arXiv preprint arXiv:2306.15768, 2023.

[4] Lugaresi, C., and associates (2019). A framework for creating perception pipelines is called MediaPipe. arXiv preprint arXiv:1906.08172.

[5] V. Bazarevsky et al. (2020). BlazePose: Real-time body pose tracking on the device. arXiv preprint arXiv:2006.10204.

[6] The World Health Organization has published a set of guidelines on physical activity and sedentary behavior (WHO)

in 2020 as follows [1]:

[7] Kendall, et al. "PoseNet: A Convolutional Network for Real-Time 6-DOF Camera Relocalization." Proceedings of the IEEE International Conference on Computer Vision, pp. 2938-2946, 2015

[8] Wei, S. E., Cao, Z., Simon, T., & Sheikh, Y. (2017). Part Affinity Fields for Real-Time Multi-Person 2D Pose Estimation. IEEE Conference on Computer Vision and Pattern Recognition, pp. 7291 – 7299.

[9] Cao, Z., Wei, S. E., Simon, T., & Sheikh, Y. (2017). Part Affinity Fields for Real-Time Multi-Person 2D Pose Estimation. IEEE Conference on Pattern Recognition and Computer Vision, pp. 7291 - 7299.Chou, Y. K. and colleagues (2021). A pooled review of studies shows how game like features change the way people keep using fitness phone programs. Human-Computer Interaction International Journal, 37(12), 1142

[10] Singh, M. plus Kumar, S. (2022). AsanaCheck follows fixed rules to decide whether a yoga posture is correct. International Conference on Computer Vision Applications Proceedings, 245 - 252.

[11] Chen, L. and R. Patel (2021). YogAI - A program that learns from data to name yoga postures. ACM Transactions on Interactive Intelligent Systems, 11(2), 1 - 24.

[12] Trabelsi but also colleagues (2021). A step-by-step way to measure posture errors and give fixes for people who practice yoga at home. IEEE Access, 9, 134516 - 134529.

[13] Genis, S., Saltonstall, E. as well as Fishman, L. M. (2009). How to spot likely yoga injuries and stop them before they start. Journal of Yoga Therapy International, 19(1), 47 - 53.

[14] Koivisto, J., Sarsa, H. or Hamari, J. (2014). Does adding game elements work? A summary of studies that tested gamification. Hawaii International Conference on System Sciences, 47th edition, 3025 - 3034