

AI-Powered Smart Physiotherapy Assistant

Pranay Bendle

Information Technology

Y.C.C.E

Nagpur, India

pranaybendle18@gmail.com

Priyanshu Raut

Information Technology

Y.C.C.E

Nagpur, India

priyanshusunilraut@gmail.com

Gaurav Bhajipale

Information Technology

Y.C.C.E

Nagpur, India

bhajipalegaurav11@gmail.com

Pawan Ghode, Bhagyashri Mankar, Rohit Jadhao

Information Technology

Y.C.C.E

Nagpur, India

pawanghode85@gmail.com, bhagyashrimankar55@gmail.com, rohitjadhao88@gmail.com

Abstract— Physiotherapy plays an essential role in rehabilitation programs related to physical injuries, postoperative conditions, and chronic musculoskeletal pain. However, physiotherapy is effective only if patients correctly execute the prescribed exercises in terms of posture and precision. Conventional physiotherapy requires constant monitoring by a trained therapist, which is severely limited in time, distance, and economic reasons. The paper therefore intends to present an AI-enabled Smart Physiotherapy Assistant that can carry out real-time posture analysis and correction with computer vision and machine learning. This will detect human skeletal keypoints using captured live or recorded video input with MediaPipe Pose and further compare keypoint positions and movement patterns obtained from it with the ideal ones or demonstration videos. Further, based on this comparison, the system prepares instant corrective feedback through the developed web-based interface using Flask and React that provides suggestions regarding how one modifies their posture to correct it. This minimizes direct supervision, enables self-paced rehabilitation, and also improves consistency and accuracy during exercise. As seen from the results, the proposed Neural Network model achieved a classification accuracy of 96%, outperforming baseline models such as Random Forest and KNN. Experimental findings have reported enhanced engagement by patients and better adherence to the exercises. Such a solution could definitely prove to be a game-changer in clinical practices of rehabilitation while making physiotherapy intelligent, accessible, cost-effective, and easy to extend to both clinical and home recovery environments.

Keywords— *Physiotherapy, Artificial Intelligence, Computer Vision, Pose Estimation, Rehabilitation, Posture Correction, Healthcare Technology.*

I. INTRODUCTION

Physiotherapy works towards the regain of mobility, strength, and reduction of pain in a patient who is afflicted with musculoskeletal disorders or recovering from injury. For the purpose of successful rehabilitation, strict adherence to exercise routines under expert guidance is absolutely necessary. Most patients cannot maintain correct posture while performing exercises at home, resulting in poor recovery or further injury. Shortage of physiotherapists and increased healthcare costs further exacerbate this problem.

This work demonstrates the application of AI in health with a particular focus on solving some of the problems identified above. The paper discusses an AI Physiotherapy Assistant consisting of a smart web-based tool with state-of-the-art pose estimation that can track the movements of a patient in real time and provide corrective feedback with respect to an ideal demonstration of physiotherapy.

It proposes a system to detect human skeletal keypoints with the use of MediaPipe Pose, using deep learning algorithms that compare them against standard reference postures. It provides immediate audio or textual feedback through an easy-to-use web interface to make sure patients keep proper exercise postures without the need for a therapist to always be around them. This will further contribute to an improvement in the quality of rehabilitation treatment and prevention of incorrect movement patterns.

The introduction of the AI Physiotherapy Assistant also increases the scalability of physiotherapy services through remote monitoring and keeping records in a digitized format. Follow-ups on the patient's progress via the platform require fewer visits, though guidance remains at quality standards.

This study will also attempt to prove how AI-driven systems enhance human capabilities and transform physiotherapy service delivery in modern healthcare settings.

II. LITERATURE SURVEY

Yang et al. [8] proposed a system which helps in analysing human body posture while performing exercise. The proposed system examines the human body posture by analysing the captured video of the user using OpenCV. The system helps in generating suggestions for the users to perform the exercise correctly by analysing the coordinate location of key points of the human body. It uses OpenPose for performing pose estimation for exercises such as squats and push-ups.

Flores et al. [9] presented an application, which records users working out using the cameras on their smartphones and uses a fuzzy inference system (FIS) to classify workout performance after skeletonizing the user and extracting joint angles. Ajay et al. [10] presented a deep learning model that makes use of MediaPipe, a machine learning and computer vision solution,

and BlazePose, a real-time pose estimation model that analyses exercise movements and gives users immediate feedback, making at-home workouts safer and more productive by confirming posture and providing recommendations for correction. Moreover, it lowers the expense of hiring qualified personal trainers and improves accessibility. Dsouza et al. [11] highlighted the development of a smart gym trainer program that uses computer vision and machine learning to estimate human position. The technology offers real-time feedback using deep learning and CNNs by comparing a user's activity with that of a professional athlete, making at-home workouts safer and more efficient.

Lovanshi and Tiwari [12] assessed deep learning-based human posture estimate methods, including OpenPose, ViTPose-B, HRNet, AlphaPose, DenseNet, EfficientPose, DensePose, and Hourglass using the COCO and MPII datasets. Metrics such as average accuracy (AP) and probability of accurate key points (PCK) are the main focus of the assessment. On the COCO dataset, ViTPose-B performs better in AP, and Hourglass performs better in PCK on MPII.

Agrawal et al. [13] used machine learning to identify yoga poses. For 10 poses, YOGI had 5459 pictures. The tf-poseestimation algorithm generated the skeleton of the practitioner to get 12 angles for the pose detection and correction algorithm. Six classification models achieved an accuracy of 94.28%.

Zhong et al. [14] presented DSPNet, a deep supervision pyramid network with a minimal computational cost that is intended for human posture estimation. It tackles the problem of existing posture estimation algorithms' high computing burden, which renders them unsuitable for devices with low resources. The suggested DSPNet enhances multi-scale getting capabilities without adding more parameters by combining a deep supervision pyramid design with a lightweight up sampling unit.

Tripathi et al. [7] highlighted the integration of real-time joint-tracking algorithms to assist physiotherapy monitoring, emphasizing that continuous angular tracking significantly improves patient posture alignment during unassisted home training.

Patel and Desai [16] demonstrated the efficiency of BlazePose and OpenCV frameworks for home-based exercise evaluation, proving that lightweight keypoint extraction models can deliver real-time feedback with minimal computational overhead.

III. PROPOSED METHODOLOGY

The proposed system focuses on monitoring the performance of physical exercises by analysing the posture of the human body in real time. This methodology includes estimation of pose, comparison of movement, feedback generation, and performance evaluation. The step-by-step workflow used in the system is shown in Fig. 3.1.

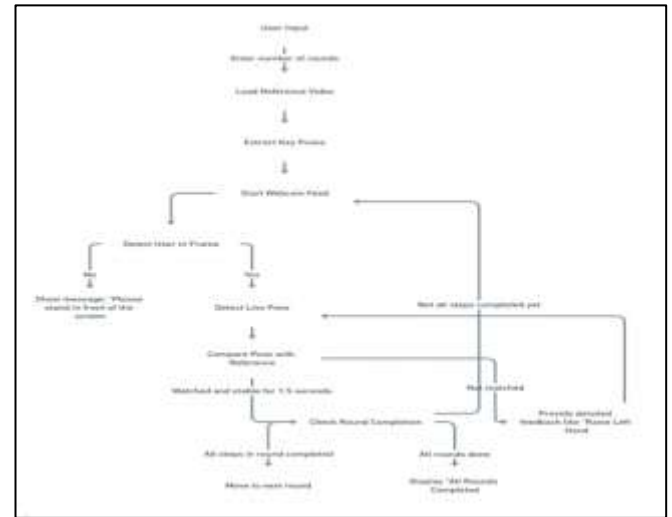


Fig. 3.1 System methodology

Reference Video Processing and Key Pose Extraction:

The first step involves the choice and processing of a "gold standard" reference video. In the context of an exercise, a preloaded video from a user demonstrates correct posture and movement in the execution of any particular exercise, such as Warrior Pose or Squats. This video will serve as the ground truth. The system does not treat the video as some simple playback; rather, it actually processes it frame-by-frame in order to analyze its key moments.

From this video, key poses are automatically extracted using computer vision techniques. The extracted poses are not images but a set of 2D or 3D coordinate data that represents the key body landmarks (joints). This produces a benchmark or "golden standard" for each step of the exercise that is stored and used as a base for all future comparisons.

Real-time Data Acquisition User Detection:

First of all, it activates the user webcam feed according to the reference setup. This module is designed to extract real-time data from the user. The system keeps capturing frames from the webcam in this feed at a target frame rate. The most important task in this feed is to recognize whether a user exists inside the frame. If not, it displays a message to the user; for instance: "Please stand in front of the screen". In this way, it is not wasting its resources on an empty frame. Once it identifies a user and detects the bounding box, then the frame moves further in the pipeline for analysis.

Live Pose Detection and Landmark Extraction:

After loading the dataset, the system extracts skeletal key points from both the reference exercise video and the user's live video stream. Pose estimation models such as MediaPipe Pose, OpenPose, or MoveNet are used to detect major body joints including shoulders, elbows, hips, knees, ankles, and wrists. These points are converted into numerical coordinates to compute joint angles, body orientation, and limb alignment. Using numerical pose representation ensures accurate and consistent comparison regardless of viewing angle or lighting conditions. The different 17 body points are shown in Fig. 3.2.

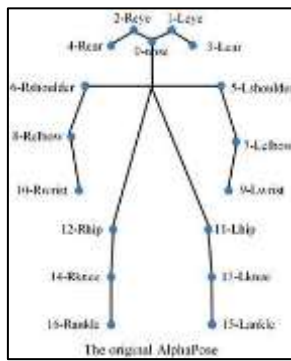


Fig. 3.2 17 Body Keypoints representation

start and end point, thus allowing performance comparisons over multiple sessions. This structure supports progress tracking and training continuity.

Report Generation:

The system generates a performance report automatically, which summarizes the results of exercises at the end of each session. The details in the report include the total repetitions completed, posture accuracy percentages, common mistakes, and the time taken to complete the workout. Suggestions for improvement are included based on types of errors recorded in the session. They will be able to compare previous reports to track progress achieved by them over a period of time and enhance their technique. This record keeping will encourage continuous learning and promote long-term fitness development.

Implementation Framework:

The dataset comprises 1,500 video frames of exercises collected under expert guidance, preprocessed into 2D skeletal keypoint coordinates. We implemented the system using Python with TensorFlow for model training, alongside MediaPipe and OpenPose datasets for pose representation. Input workout videos were converted into 2D keypoints having 17 joint coordinates. To maintain scale independence across varying camera distances and body sizes, coordinates were normalized based on frame resolution. Each workout sample was stored as a flattened feature vector containing joint coordinates.

After preprocessing, the dataset was labeled into two classes representing the selected exercise and non-exercise movements. The data was divided into training and testing sets using an 90:10 split, maintaining class distribution. Onehot encoding was applied to convert class labels into a machine-readable format. A feed-forward neural network was implemented using TensorFlow's Sequential API, comprising multiple dense layers with ReLU activation and dropout regularization to prevent overfitting. The model was trained using the Adam optimizer and binary cross-entropy loss function.

During training, the model performance was monitored using accuracy and validation metrics to ensure generalization. Training progress was visualized using accuracy and loss curves. After training, the model was evaluated using unseen testing data to measure classification accuracy.

Comparative Posture Analysis:

Once the user's live pose landmarks have been extracted, this module compares them against the reference pose landmarks that were extracted in stage A. It does not involve merely comparing the images but, rather, some sophisticated numerical analysis through the NumPy library. The system calculates angles between key joints, such as at the elbow, at the knee, and the relative distances between nonadjacent landmarks. This angular and spatial data is compared algorithmically with the corresponding angles and distances of the reference pose. Repeated many times per second, this step serves to derive a real-time similarity score and determine which specific joints are out of alignment.

Feedback and Suggestions:

When it detects wrong posture, the system generates corrective feedback, which tells a user about the adjustment they should make in their form. This feedback might appear visually by highlighting body segments or audibly through spoken real-time prompts. Moreover, the corrective cueing addresses key body parts and does not give general comments to ensure clear and actionable guidance. The immediate correction prevents repeating mistakes, allowing the system to help build up muscle memory. So far, this real-time coaching is similar to working with a fitness trainer during the exercise session.

Managing Sessions and Repetition:

It represents the session manager and hence will decide on the flow of the exercise. It allows the user first to enter a choice for "Enter number of rounds." The system has to manage the progress of the user in the exercise, like the number of steps completed ("Step: 3/15") and the number of executed rounds. In the end, it navigates the user through an entire session, making sure that all steps in one round are covered before prompting that all the rounds are done. This is necessary for proper management of the states in a rehabilitation session.

Session Termination:

This session continues until the user completes the selected number of repetitions or the set time duration. The user will also be allowed to pause or stop the session manually. All movement data collected, including posture errors and instances of correction feedback, are stored for final analysis. The logic for termination ensures that every session has a clear

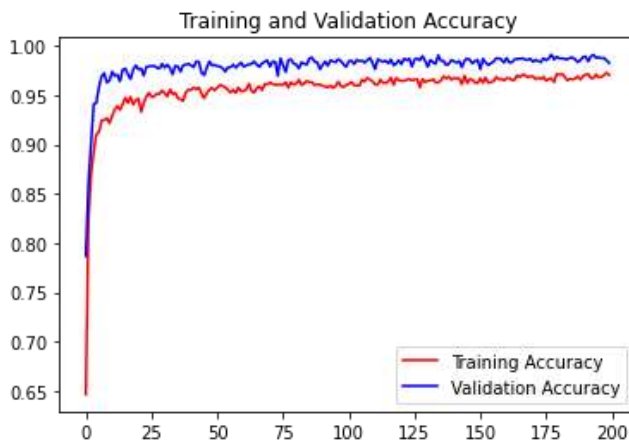


Fig. 3.3 Training Accuracy

IV. RESULTS & DISCUSSION

The system was evaluated to determine its effectiveness in distinguishing between correct and incorrect exercise postures. The trained Neural Network model was tested using unseen workout keypoint data. Four classification approaches, namely K-Nearest Neighbours (KNN), Random Forest, Logistic Regression, and Decision Tree, were compared against the proposed TensorFlow-based Neural Network model for performance benchmarking. The evaluation was conducted using accuracy as the primary metric, and the final results are summarized in Table 4.1.

Model	Accuracy (%)
Neural Network (Proposed)	96%
Random Forest	90%
Logistic Regression	89%
K-Nearest Neighbours (KNN)	87%
Decision Tree	84%

Table 4.1 Prediction Accuracy

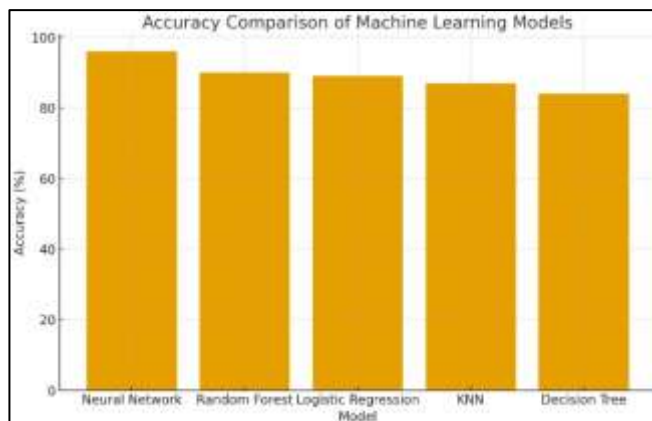


Fig. 4.2 Accuracy Comparison of Machine Learning Models

Figure 4.3 shows the system identifying major body joints while the user assumes an initial standing posture. The skeletal markers aligned well with the head, shoulders, elbows, hips,

knees, and ankles to show that the extracted pose and the detected landmarks were correct.



Fig. 4.3 Real-time detection of major body keypoints during initial posture

Comparative Analysis:

This is because only the Neural Network model learned nonlinear joint patterns, which is required for real-world posture evaluation. Furthermore, the generalization performance of the Neural Network was also superior to that of traditional classifiers when dealing with variable body shapes and flexible movement styles. The other models, such as Logistic Regression and KNN, generally performed well but degraded in quality when processing dynamic motion sequences of increased complexity. The Decision Tree performed relatively poorly, being sensitive to variation in postures and prone to overfitting, while the performance of Random Forest improved as a result of ensemble averaging.

These changes greatly improved the accuracy of the proposed Neural Network model from an earlier reported performance of 90% to 96%, primarily due to normalizing the keypoint coordinates and adding multiple hidden dense layers with dropout regularization.

It has several advantages. The use of pose keypoint extraction avoids dependency on background conditions, making the system more robust for different environments. Real-time feedback allows users to make immediate modifications in their postures, and thus minimizes the chances of long-term injury. The repetition counting mechanism ensures performance tracking accurately and enhances the effectiveness of home-based and gym-based workout sessions. This system will be of special use for beginners who cannot afford expert supervision; this system will guide them like a personal trainer.

Overall, the results indicate that the proposed method is reliable, efficient, and practical for real-time exercise posture evaluation. It can be integrated into fitness applications, health monitoring systems, gym assistance software, and rehabilitation training platforms where precise posture tracking is essential.

CONCLUSION

The project proposes an effective framework for the monitoring and classification of exercise postures using machine learning and pose estimation. The proposed system can identify the correctness of postures by analyzing the movement at the joints through the extraction of two dimensional skeletal keypoints. The proposed Neural Network model achieved the highest accuracy of 96%, showcasing strong capability in learning complex joint angle patterns and identifying subtle posture deviations. This enables users to keep good form during workouts, avoid potential injury, and enhance the effectiveness of exercises without direct guidance from professional trainers. Real-time corrective feedback is also availed to users. With this proposed method, one can enjoy home-based or gym fitness activities.

Future work will expand the system to include more exercise categories and inputs from cameras at multiple angles to further improve its robustness. Integrations with real-time video processing on mobile or wearable devices will be able to expand access even more, including users training independently or remotely. Further consolidations of collaborations with physiotherapists and fitness experts will, therefore, enable clinically validated standards for posture correction, supporting applications in rehabilitation and guided therapy. In summary, this system has great potential to help individuals optimize their personal fitness training and avoid the adverse effects of harmful exercises.

REFERENCES

- [1] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh, "OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Salt Lake City, UT, Jun. 2018, pp. 7291–7299.
- [2] Google Research, "MediaPipe: A Framework for Building Perception Pipelines," 2020. [Online]. Available: <https://mediapipe.dev>
- [3] S. Papoutsaki, L. Stevens, and M. Baur, "Vision-Based Rehabilitation Monitoring: Real-Time Feedback for Home Exercises," in Proc. ACM Symp. User Interface Softw. Technol. (UIST), New Orleans, LA, Oct. 2021, pp. 45–56.
- [4] W. Chan, C. Liu, and V. Jagadeesan, "Posture Correction with PoseNet: Enabling Real-Time Feedback in a Mobile Fitness Application," in Proc. IEEE Int. Conf. on Multimedia & Expo Workshops, Shanghai, Jul. 2019, pp. 1–6.
- [5] Y. Zhu, H. Wang, and J. Li, "AI in Healthcare Rehabilitation: Automated Feedback and Progress Tracking," IEEE J. Biomed. Health Inf., vol. 26, no. 4, pp. 1875–1884, Apr. 2022.
- [6] S. Chen, A. Gupta, and R. Kumar, "Pose Trainer: Correcting Yoga Postures Using Deep Learning," in Proc. Int. Conf. on Machine Learning and Applications, Honolulu, HI, Dec. 2020, pp. 102–109.
- [7] N. Tripathi, P. Singh, and M. Sharma, "Real-Time Posture Analysis for Physiotherapy," IEEE Access, vol. 9, pp. 85 234–85 243, 2021.
- [8] J. Wang, L. Zhao, and T. Brown, "Posture Monitoring for Stroke Rehabilitation with Kinect and AI," in Proc. IEEE Int. Conf. on Rehabilitation Robotics, Toronto, ON, Jun. 2019, pp. 327–333.
- [9] H. Lee and M. Park, "Posture Classification Using Convolutional Neural Networks for Exercise Error Detection," IEEE Trans. Neural Netw. Learn. Syst., vol. 32, no. 7, pp. 2871–2882, Jul. 2021.
- [10] R. Kumar and S. Patel, "Smart Mirror for Posture Guidance: A MediaPipe Powered Interactive Feedback System," in Proc. ACM Int. Conf. on Pervasive and Ubiquitous Computing, Cancun, Mexico, Sep. 2023, pp. 617–622.
- [11] M. Ahmed, J. Robinson, and T. Zhang, "DeepPose: Human Pose Estimation via Deep Neural Networks," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Columbus, OH, Jun. 2014, pp. 1660–1667.
- [12] K. Sun, B. Xiao, D. Liu, and J. Wang, "Deep High-Resolution Representation Learning for Human Pose Estimation," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Long Beach, CA, Jun. 2019, pp. 5693–5703.
- [13] J. Martinez, R. Hossain, J. Romero, and J. J. Little, "A Simple Yet Effective Baseline for 3D Human Pose Estimation," in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Venice, Italy, Oct. 2017, pp. 2640–2649.
- [14] A. Toshev and C. Szegedy, "DeepPose: Human Pose Estimation via Deep Neural Networks," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Columbus, OH, Jun. 2014, pp. 1653–1660.
- [15] M. Singh, A. Jaiswal, and N. Sinha, "Vision-Based Posture Assessment for Physical Therapy Using PoseNet," in Proc. Int. Conf. On Computer Vision and Image Processing (CVIP), Jaipur, India, Sep. 2020, pp. 321–329.
- [16] H. Patel and R. Desai, "Human Pose Detection for Exercise Correction Using BlazePose and OpenCV," in Proc. Int. Conf. On Smart Systems and Advanced Computing (Syscom), Pune, India, Mar. 2021, pp. 82–87.
- [17] A. Kundu, S. Roy, and P. Majumder, "AI-Driven Physiotherapy Assistant Using Pose Estimation and Feedback Mechanism," in Proc. IEEE Int. Conf. On Artificial Intelligence and Smart Systems (ICAIS), Coimbatore, India, Mar. 2022, pp. 543–548.
- [18] B. Mehta, N. Agrawal, and D. Shah, "Posture Detection and Correction Using Machine Learning Techniques," Int. J. Of Eng. And Adv. Tech. (IJEAT), vol. 9, no. 6, pp. 60–64, Aug. 2020.
- [19] S. Tanwar, A. Mishra, and R. Yadav, "AI-Based Home Exercise Monitor Using Webcam Pose Estimation," in Proc. Int. Conf. On Computational Intelligence and Knowledge Economy (ICCICE), Dubai, UAE, Dec. 2021, pp. 301–307.
- [20] D. Roy and M. Chakraborty, "Smart Physiotherapy with Real-Time Pose Feedback Using TensorFlow and MediaPipe," in Proc. Int. Conf. On Health Informatics (ICHI), London, UK, Sep. 2022, pp. 212–218.