Smart Workout Companion: Leveraging Deep Learning for Precise Posture Monitoring and YogaPose Refinement

Shubhada Mone[1], Atharva Yanna[2], Pratham Kotkar[3], Kabir Kapse[4], and Om Jagtap[5]
Computer Science Department, Marathwada Mitra Mandals College of Engineering, Pune, India

***Abstract*—This study offers a thorough analysis of the state-of-the-art technology for posture correction and fitness monitoring. These technologies have significantly increased in popularity in a time when there is a growing focus on one's own health and wellbeing. The paper examines a variety of advances, highlighting their benefits and drawbacks, ranging from sensor-based systems to computer vision applications. It combines information from many studies and commercial items to offer light on the merits of posture correction methods and how they affect exercise regimens. Additionally, it covers new developments like AI-powered virtual fitness assistants and the incorporation of mindfulness exercises like yoga and meditation into workout routines. This study seeks to serve as a beneficial resource for academics, practitioners, and individuals looking to make educated decisions in the pursuit of holistic well-being by critically examining the body of current literature and the technological environment.**

***Index Terms*—Neural Network, Deep Learning, MediaPipe, OpenCV**

I. Introduction

In the contemporary world, where digital innovation intersects with health and wellness, the pursuit of an active and balanced lifestyle is increasingly becoming a fundamental aspect of daily life. Individuals across the globe aspire to attain and maintain physical fitness, often with the goal of not only enhancing their physical prowess but also nurturing their mental and emotional well-being. Amid this zeitgeist of well-being, we introduce a groundbreaking technological advancement: the ”Smart Workout Companion.”

Our research embarks on a comprehensive exploration of this transformative application, which promises to reshape the landscape of fitness and well-being. The Smart Workout Companion is envisioned as a real-time, multifaceted solution that intricately monitors exercise postures, accurately counts repetitions, and thoughtfully expands its capabilities to encompass yoga and meditation modules. This visionary application holds the potential to not only redefine fitness routines but also foster holistic well-being, ushering in a new era of health consciousness that empowers individuals to embark on transformative journeys towards optimal physical and mental wellness.

In a world where technology has woven itself into the fabric of everyday life, the Smart Workout Companion represents a convergence of scientific innovation and the age-old pursuit of well-being. Exercise has long been recognized as a cornerstone of physical health, but the efficacy of these endeavors hinges on precise execution and adherence to proper form. The challenge of maintaining correct postures and counting repetitions accurately has been a persistent obstacle on the path to fitness excellence. This application steps in as a dedicated virtual coach, constantly vigilant and responsive, to provide real-time feedback and corrections for exercise postures. By addressing these fundamental challenges, it not only reduces the risk of injury but also optimizes muscle engagement, thereby enhancing the effectiveness of workouts. The Smart Workout Companion's ability to count repetitions with precision ensures that fitness enthusiasts can track their progress with accuracy, offering a motivating element to their exercise routines.

However, this groundbreaking application goes beyond the scope of traditional fitness by embracing a holistic approach to well-being. It extends its capabilities to incorporate yoga and meditation modules, recognizing the interconnectedness of physical and mental health. In a world where stress and anxiety are prevalent, the Smart Workout Companion empowers users to embark on transformative journeys, promoting both physical and mental wellness. Through guided yoga sessions and mindfulness exercises, it endeavors to create a balanced and harmonious lifestyle that transcends the boundaries of physical fitness.

The Smart Workout Companion embodies the fusion of cutting-edge artificial intelligence and real-time monitoring, ushering in a paradigm shift in how individuals engage with their health and fitness routines. In an era where personalization and data-driven insights have become paramount, this application stands as a testament to the potential of technology to facilitate healthier, more mindful living. It not only offers personalized guidance but also collects valuable data that can inform future workout strategies, ensuring that fitness journeys are not only effective but also sustainable.

This research paper delves deeply into the architecture, algorithms, and underlying technologies that empower the Smart Workout Companion, offering a comprehensive understanding of its capabilities and its potential impact on the fields of fitness, health, and wellness. As we navigate a world where self-care and well-being are paramount, our work contributes to the evolving dialogue on how technology can be harnessed to foster healthier, more fulfilling lives. The Smart Workout Companion is poised to lead us into a new era of holistic well-being, where technology acts as an enabler for individuals to optimize their physical and mental wellness, ultimately leading to a happier and healthier society.

II. Related Work

Hao-Shu Fang et al [1] proposed the multi-person pose estimation process is described in this part, which also discusses problems with traditional soft-argmax keypoint regression and suggests creative fixes. The gradient asymmetry and size-dependent keypoint scoring issues are dealt with using the symmetric integral keypoints regression approach. In order to isolate joint confidence prediction from integral regression, it employs amplitude symmetric gradients (ASG) to enhance gradient stability. The section also covers part-guided proposal generation for training samples with distributions similar to those of outputs from human detectors and multi-domain knowledge distillation for training the network. Finally, parametric pose non-maximum suppression (NMS), based on data-driven optimization of removal criteria, is provided to reduce redundant pose estimations. Multi-person posture estimate performance is improved, becoming more reliable and precise because to this all-encompassing technique.

Keze Wang et al [2] introduces a 3D human pose machine that can effectively integrate long-range spatio-temporal dependencies and 3D geometry knowledge in a comprehensive manner. The model is enhanced with a novel self-supervised correction mechanism, which involves two dual learning tasks: 2D-to-3D pose transformation and 3D-to-2D pose projection.

Filippos Markolefa et al [3] proposes a new method to improve online training videos by overlaying the trainee's silhouette for easy comparison with the trainer. It combines various techniques for background modeling and foreground extraction, addressing challenges like slow-moving objects and lighting changes.

Venkata Sai P Bhamidipati et al[5] proposed methodology for developing a robust posture estimation system for an AI Gym Trainer using Mediapipe and OpenCV consists of four main stages: data collection, data preprocessing, pipeline training, and pipeline evaluation.

Hustinawaty et al [6] proposed application that uses MediaPipe for pose recognition and OpenCV to capture video from a camera. The procedure begins with the video input being captured by OpenCV, which is then processed by MediaPipe to extract the x, y, and z coordinates of joints, with an emphasis on the left and right knees and the central hip. The left and right legs' respective vectors are calculated using these coordinates, and then the angle between these vectors is calculated. The outcome is overlaid on the video frame. Specifically in a clinical situation with excellent illumination and a distance of 2.5–3 meters from the patient's bed, the application is made to identify movements based on webcam photographs. Users are requested to provide patient information before motion capturing begins. Straight Leg Raise Test (SLRT) postures are identified using image processing software, and range of motion (ROM) angles are computed for analysis. The findings, which include the patient's name, ROM angles, the existence of pain, and the diagnosis of sciatica, are saved in a database for further analysis.

Samhitha G et al [8] proposed Vyayam architectural model utilizes live video input to monitor exercise routines, employing MediaPipe to create a virtual skeleton of the exercising person's body and locate specific keypoints or joints corresponding to the exercise being performed. By calculating the angles between these keypoints, the model tracks and increments the exercise count, displaying it to the user. It also assesses the exercise's correctness based on angle thresholds. The system relies on modules such as cv2, MediaPipe, and numpy, with OpenCV for image processing, MediaPipe for multi-modal ML pipelines, and numpy for angle calculations and point management. This AI-based Vyayam model offers applications in gaming, robotics, animation, movies, and physiotherapy, providing real-time feedback and achieving a high accuracy rate of 91 percent.

Rohit Girdhar et al [16] proposed method presents a two-stage approach for efficient and accurate human pose tracking over time. In the first stage, they employ a 3D human pose predictor based on an extended Mask R-CNN model, which incorporates spatiotemporal operations by converting 2D convolutions into 3D. This model takes short video clips as input and predicts poses for all individuals by integrating temporal information, outperforming its 2D counterpart. To track individuals across frames, a lightweight optimization links these pose predictions, addressing complexity issues. They employ a heuristic to achieve state-of-the-art accuracy on the PoseTrack benchmark while significantly improving computational efficiency compared to recent approaches. The approach combines 3D convolutional networks, tube proposal networks, and anchor-based regression for robust tracking and pose estimation.

Bruce Xiaohan Nie et al [17] proposed model that estimates joint depths to anticipate a 3D human position from a 2D input. It makes use of a two-stage LSTM architecture, with the first level combining a patch-LSTM and a skeleton-LSTM to collect both global and local information. Contextual information is effectively propagated via joints by a tree-structured LSTM. By adding a fully-connected layer to the second-level LSTM, depth prediction is made possible. This method gives flexibility for adopting advancements in 2D pose estimate techniques by allowing for independent 2D pose estimation and depth prediction. The model performs better than flat structures, demonstrating the benefits of the tree-structured LSTM for estimating human position.The learning process for the model is divided into three stages: pre-training the skeleton-LSTM on 2D-3D pose pairings from Mocap data, pre-training the patch-LSTM on RGB pictures from Mocap and 2D pose datasets with a 2D pose regression task, and finally integrating both LSTMs for end-to-end training. Absolute and relative depth mistakes are both taken into account by the loss function. In order to prevent overfitting, Mocap data and in-the-wild photos with 2D posture annotations are combined. Multi-task learning is then used to predict 2D joint locations and depth values, improving the generalization of the model.

Jiahua Xu et al [20] proposes system that enables real-time hand gesture recognition and motion tracking, making it suitable for various applications on standard PCs with web cameras. It involves capturing webcam frames, preprocessing them, and utilizing a pre-trained Single Shot MultiBox Detector (SSD) for hand detection. A user-defined threshold validates detections, providing bounding boxes and cropped hand images. Gesture recognition relies on a Convolutional Neural Network (CNN), trained with a custom dataset, to classify gestures based on cropped hand images. The motion tracking component calculates hand centroid trajectories to determine directional gestures. This versatile system offers practical applications in real-time gesture recognition and motion tracking, with potential for integration into different domains.

Kang Wang et al [21] proposed system introduces advanced human-computer interaction using head pose, eye gaze, and body gestures. It's designed for large screens and offers versatile applications, including PowerPoint presentations and a balloon shooting game. The system builds on a previous immersive system, emphasizing vision-based interaction while omitting speech. Notably, it allows smooth PowerPoint presentations via body gestures and estimates human attention through head pose and eye gaze.

III. METHODOLOGY

A. Data Collection and Preparation:

Collect Exercise, Yoga, and Meditation Data: Gather a diverse dataset of exercise, yoga, and meditation video clips demonstrating correct and incorrect postures. This dataset should encompass various individuals with different body types, skill levels, and lighting conditions.

Labeling and Annotation: Annotate the collected data with labels for correct and incorrect postures and repetition counts. Annotate key body joints and points of interest in yoga poses and meditation postures. Data Augmentation: Augment the dataset by introducing variations in lighting conditions, camera angles, and user demographics to ensure robustness.

B. Preprocessing:

Video Preprocessing: Apply techniques like resizing, cropping, and frame extraction to prepare video data for model training and inference.

Pose Estimation with MediaPipe: Utilize the MediaPipe library to estimate key body joints and points in yoga poses, meditation postures, and exercise routines. Extract relevant keypoints for posture analysis.

C. Model Development:

Gym Exercise Module: Train a neural network model (e.g., Convolutional Neural Network) to analyze real-time webcam video data, monitoring users' exercise postures. Implement corrective guidance using pose estimation data, providing real-time feedback on alignment and balance during exercises. Develop repetition counting algorithms to track and tally exercise repetitions.

Meditation Module: Create a model to guide users through meditation postures and techniques. Implement computer vision techniques to track body movements during meditation and provide feedback on posture and alignment. Develop visualization techniques for mindfulness and relaxation guidance.

Yoga Module: Train models to recognize yoga poses from key body joints and points obtained through pose estimation. Develop real-time posture correction algorithms, offering step-by-step guidance for various yoga poses.

D. System Integration:

Combine the Gym Exercise, Meditation, and Yoga modules into a unified application that allows users to switch between modules seamlessly. Implement personalized user profiles and progress tracking to provide tailored recommendations and guidance based on individual preferences and performance.

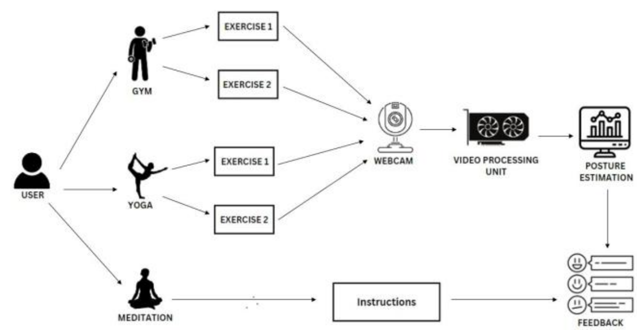

Fig. 1. **System Architecture**

E. Software and Hardware Integration:

Develop the application on platforms like Jupyter notebook, Anaconda Navigator, Spyder, or other suitable development environments. Integrate computer vision libraries such as OpenCV and MediaPipe for video analysis and pose estimation. Ensure compatibility with both Windows and Linux operating systems. The application requires a camera or webcam as a hardware component.

F. Testing and Evaluation:

Conduct extensive testing using real users and a variety of exercises, yoga poses, and meditation postures. Evaluate the application's ability to monitor postures, count repetitions, and provide guidance in real-time. Measure the application's performance in terms of exercise efficiency, safety, and well-being enhancement.

G. Expected Outcomes:

Real-time Posture Monitoring for Gym Exercises, Yoga, and Meditation. Accurate Repetition Counting. Improved Exercise Performance with Corrective Guidance. Enhanced Health and Safety during workouts.

Algorithm Analysis

Time Complexity: The overall time complexity is dominated by the continuous monitoring of angles, making it **O(n)**,

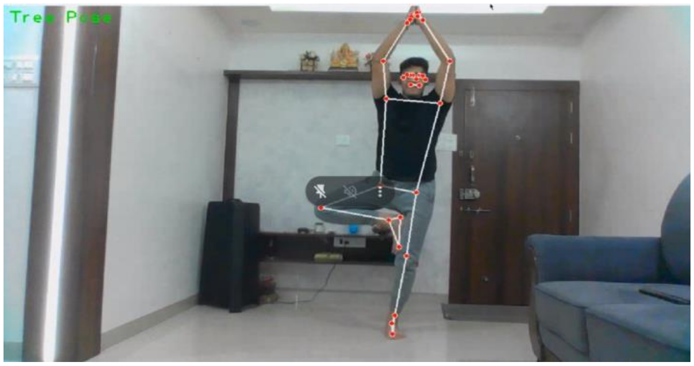

Fig. 2. **Yoga Pose Detection**

Fig. 3. **Bicep workout with Repetition counting**

where n is the number of frames processed during the exercise.

Space Complexity: The space complexity is influenced by the storage requirements of the pose estimation model, landmarks, calculated angles, and additional data related to angle changes over time. Assuming the dominant factor is the continuous monitoring of angles, the space complexity is **O(1)**, indicating constant space requirements.

Conclusion and Future Scope

A. Conclusion:

The Smart Workout Companion, based on artificial intelligence and machine learning, offers an innovative and powerful solution for real-time posture monitoring and yoga position optimization. To provide consumers with a holistic fitness experience, the project makes use of modern technologies like as deep learning and computer vision, notably the OpenCV and MediaPipe libraries. The research's algorithmic architecture offers a systematic approach to frame preprocessing, landmark recognition, angle computation, and repetition counting. The use of neural networks adds a level of sophistication, helping to increase precision in exercise monitoring.

The flexibility to various workouts and yoga postures, together with its user-friendly interface and progress monitoring features, presents it as a viable tool for anyone looking for a more holistic approach to fitness. As fitness technology advances, the Smart Workout Companion demonstrates the power of AI-driven solutions in improving workout routines, supporting good form, and boosting general well-being. Future initiatives may investigate other modalities, improve algorithms for better accuracy, and broaden the system's interoperability with upcoming technologies, therefore adding to the continuing breakthroughs in the convergence of AI and fitness research.

Algorithm 1 Algorithm for Posture Monitoring and Yoga Pose Recognition

Require: Video frames from a webcam feed.
Ensure: Repetition count and visual feedback displayed on the video frame.

Step 1: Preprocess each frame by resizing and converting it to RGB.
Step 2: Use the pose estimation model to identify pose landmarks.
Step 3: Extract relevant landmarks for the specific exercise or yoga pose.
Step 4: Calculate angles using these landmarks.
Step 5: Map angles to a percentage scale.
Step 6: Monitor angle changes for repetition counting.
Step 7: Define criteria for recognizing repetition start and end.
Step 8: Update the repetition count based on these criteria.
Step 9: Draw pose landmarks on the frame for visual feedback.
Step 10: Display the repetition count on the video frame.
Step 11: Allow the user to exit the application (e.g., by pressing 'q').
Step 12: Release video capture and close OpenCV windows when done.

B. Future Work:

1. Integration of Multiple Modes: Future Smart Workout Companion versions may investigate integration with other sensors or modalities, such as inertial measurement units (IMUs) or depth sensors, to collect more thorough data. This multi-modal method may give more detailed insights on user movements, allowing for more sophisticated research of workout postures and yoga poses.

2. Personalization and Adaptive Learning: The system might grow to include daptive learning techniques, which would use user input and performance data to tailor the instruction delivered. Individual fitness levels, preferences, and develop-

ment may need modifying workout recommendations, form corrections, and repetition targets.

3. Collaboration and Social elements: Adding collaborative or social elements to the project might boost user engagement. Incorporating features such as virtual training groups, shared progress monitoring, and interactive challenges may help users feel more connected and motivated.

4. Integration of Health data: Future versions may incorporate tracking health data like as heart rate and oxygen saturation levels to offer users with a more holistic knowledge of their fitness and well-being. This might be especially useful for people who have specific health concerns or are undergoing rehabilitation.

REFERENCES

[1] Hao-Shu Fang, Jiefeng Li, Hongyang Tang, Chao Xu, Haoyi Zhu, Yuliang Xiu, Yong-Lu Li, and Cewu Lu, ”AlphaPose: Whole-Body Regional Multi-Person Pose Estimation and Tracking in Real-Time” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 6, 2023, Publisher: IEEE.

[2] Keze Wang, Liang Lin, Pengxu Wei, Chen Qian, and Chenhan Jiang, ”3D Human Pose Machines with Self-supervised Learning” IEEE Transactions on Industrial Informatics, vol. 15, no. 4, 2019, Publisher: IEEE.

[3] Filippos Markolefa, Konstantia Moirogiorgou, George Giakos, Michalis Zervakis, ”Virtual Video Synthesis for Personalized Training” 2018 IEEE International Conference on Imaging Systems and Techniques (IST), 2018, Publisher: IEEE.

[4] Tong Zhang, Jingxiang Lian, Jingtao Wen, and C. L. Philip Chen, ”Multi-Person Pose Estimation in the Wild: Using Adversarial Method to Train a Top-Down Pose Estimation Network” IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 53, no. 7, 2023, Publisher: IEEE.

[5] Venkata Sai P. Bhamidipati, Ishi Saxena, Mrs. D. Saisanthiya, Dr. Mervin Retnadhas, ”Robust Intelligent Posture Estimation for an AI Gym Trainer using Mediapipe and OpenCV” 2023 International Conference on Networking and Communications (ICNWC), 2023, Publisher: IEEE.

[6] Hustinawaty, Tavipia Rumambi, Matrissya Hermita, ”Motion Detection Application to Measure Straight Leg Raise ROM Using MediaPipe Pose” 2022 4th International Conference on Cybernetics and Intelligent System (ICORIS), 2022, Publisher: IEEE.

[7] Ashish Gupta and Hari Prabhat Gupt, ”YogaHelp: Leveraging Motion Sensors for Learning Correct Execution of Yoga With Feedback” IEEE Transactions on Artificial Intelligence, vol. 2, no. 4, 2021, Publisher: IEEE.

[8] Samhitha G, Srinivasa Rao D, Ch. Rupa, Y. Ekshita, and R Jaswanti, ”Vyayam: Artificial Intelligence based Bicep Curl Workout Tracking System” 2021 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), 2021, Publisher: IEEE.

[9] Bugra Tekin, Pablo Marquez-Neila, Mathieu Salzmann, Pascal Fua, ”Learning to Fuse 2D to 3D Image Cues for Monocular Body Pose Estimation” 2017 IEEE International Conference on Computer Vision, 2017, Publisher: IEEE.

[10] Nachaat Mohamed, ”Importance of Artificial Intelligence in Neural Network through using MediaPipe” Proceedings of the Sixth International Conference on Electronics, Communication and Aerospace Technology (ICECA 2022), 2022, Publisher: IEEE.

[11] Siddarth Sreeni, Hari S R, Harikrishnan R, and Sreejith V, ”Multi-Modal Posture Recognition System for Healthcare Applications” Proceedings of TENCON 2018 - 2018 IEEE Region 10 Conference, 2018, Publisher: IEEE.

[12] Tong Zhang, Jingxiang Lian, Jingtao Wen, and C.L. Philip Chen, ”Multi-Person Pose Estimation in the Wild: Using Adversarial Method to Train a Top-Down Pose Estimation Network” IEEE Transactions on Systems, Man and Cybernetics, Systems, vol. 53, no. 7, July 2023, Publisher: IEEE.

[13] Linsen Dong, Yuanlong Li, Xin Zhou, Yonggang Wen, and Kyle Guan, ”Intelligent Trainer for Dyna-Style Model-Based Deep Reinforcement Learning,” IEEE Transactions on Neural Networks and Learning Systems, 2020, Publisher: IEEE.

[14] Pratool Bharti, Debraj De, Sriram Chellappan, and Sajal K. Das, ”HuMAn: Complex Activity Recognition with Multi-modal Multi-positional Body Sensing” IEEE Transactions on Mobile Computing, vol. 18, 2019, Publisher: IEEE.

[15] Jianyuan Sun, Hui Yu, Guoqiang Zhong, Junyu Dong, Shu Zhang, and Hongchuan Yu, ”Random Shapley Forests: Cooperative Game-Based Random Forests With Consistency” IEEE Transactions on Cybernetics, 2020, Publisher: IEEE.

[16] Rohit Girdhar, Georgia Gkioxari, Lorenzo Torresani, Manohar Paluri, and Du Tran, ”Detect-and-Track: Efficient Pose Estimation in Videos” IEEE Conference on Computer Vision and Pattern Recognition, 2018.

[17] Bruce Xiaohan Nie, Ping Wei, and Song-Chun Zhu, ”Monocular 3D Human Pose Estimation by Predicting Depth on Joints” IEEE International Conference on Computer Vision, 2017, Publisher: IEEE.

[18] Ze Wu, Jiwen Zhang, Ken Chen, and Chenglong Fu, ”Yoga Posture Recognition and Quantitative Evaluation with Wearable Sensors Based on Two-Stage Classifier and Prior Bayesian Network” Sensors 2019, 19, 5129, Publisher: Sensors.

[19] Tewodros Legesse Munea, Yalew Zelalem Jembre, Halefom Tekle Weldegebriel, Longbiao Chen, Chenxi Huang, and Chenhui Yang, ”The Progress of Human Pose Estimation: A Survey and Taxonomy of Models Applied in 2D Human Pose Estimation” Preparation of Papers for IEEE Access, 2017, Publisher: IEEE.

[20] Jiahua Xu, Priyanka Mohan, Faxing Chen, and Andreas Nürnberger, ”A Real-time Hand Motion Detection System for Unsupervised Home Training” IEEE International Conference on Systems Man and Cybernetics, 2020, Publisher: IEEE.

[21] Kang Wang, Rui Zhao and Qiang Ji ”Human Computer Interaction with Head Pose, Eye Gaze and Body Gestures” 13th IEEE International Conference on Automatic Face and Gesture Recognition 2018, Publisher: IEEE.