International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 12 Dec. 2024, pp: 120-124 www.ijaem.net ISSN: 2395-5252

A Survey on Ai-Powered Exercise Posture Correction System

Archana Rajendran, Mohammed Anwar Shah, AshnaSiraj, Praveen Kumar S, Dr. Shyama Das
Department of Computer Science
College of Engineering Chengannur
Kerala, India

Date of Submission: 01-12-2024 Date of Acceptance: 10-12-2024

ABSTRACT—Proper exercise posture is essential for preventing injuries and maximizing workout effectiveness, but many individuals lack access to professional trainers. AI-powered exercise posture correction systems provide an affordable and accessible alternative, using tools like Convolutional Neural Networks (CNNs) and pose estimation frameworks such as MediaPipe, OpenPose, PoseNet, and MoveNet. These systems analyze body movements in real time, detect misalignments, and offer immediate feedback to help users improve their form. This paper reviews the technologies, methods, and applications of these systems in fitness, rehabilitation, and elderly care. It highlights their benefits, challenges, and potential for making exercise guidance more widely available and effective.

Keywords— Human Pose Estimation, Health and Fitness, KeyPoint Detection, Pose Tracking, 3D Human Pose Estimation.

I. INTRODUCTION

Daily exercise is vital for maintaining good health, yet many individuals perform their workout routines without guidance from a professional trainer. A trainer typically evaluates key body parts such as the spine, neck, pelvis, shoulders, shoulder blades, hips, knees, and feet to identify potential misalignments. However, exercising with improper posture can cause muscle strain or injury, especially among elderly individuals who face increased risks due to health conditions, excess weight, and age-related issues that make it difficult to maintain balance and posture. While hiring a personal trainer is an option to help correct body alignment and offer guidance on posture during exercises, it can be costly and requires advance scheduling. Therefore, an automated posture evaluation system would be highly beneficial, particularly for the elderly, in identifying misalignments and preventing injuries. As artificial intelligence becomes more integrated across different sectors, posture estimation algorithms are finding numerous applications, from clinical gait analysis to video surveillance. With advances in human-pose estimation using simple web cameras and computer vision technologies, posture evaluation has reached new heights. Human pose estimation has become a key area of research in computer vision. It involves analyzing body posture by identifying skeletal joints and assessing the accuracy of body positions. Most approaches rely on kinematic models that depict the body's structure through its joints and limbs. These techniques are useful in medical settings for evaluating posture during physiotherapy exercises, detecting falls, and supporting in-home rehabilitation. Human body tracking systems can also help prevent injuries during various physical activities, such as aerobic, anaerobic, or agility training. This is especially important for the elderly, where monitoring posture in real-time via a simple webcam can reduce the likelihood of injury. Here, we introduce a machine learning-based posture classifier for real-time exercise posture recognition using a simple web camera. The main goal is to develop an automated system that accurately identifies exercise postures, eliminating the need for a professional trainer. The datasets used for training and testing the system consist of exercise videos contributed and labeled by a physical trainer, ensuring the accuracy of the posture classification process.

A. Role of CNN

Convolutional Neural Networks (CNNs) are highly effective in automatically learning and extracting relevant features from images and videos, particularly the spatial relationships between different body parts. In the context of exercise posture recognition, CNNs analyze video frames to identify body landmarks and skeletal joints, which are essential for assessing whether a posture aligns with the expected form during specific exercises, such as squats or push-ups. By accurately locating and tracking key body joints—like the shoulders, elbows, hips, and knees—CNNs play a crucial role in evaluating posture during physical activities.

After feature extraction, the CNN outputs are directed to classification layers that categorize various postures or exercises. These networks can distinguish between correct and incorrect postures by learning from labeled training datasets, with each posture linked to a specific exercise form. This capability is vital for real-time feedback systems that assist users in correcting their posture. CNNs excel at recognizing complex body movements, making them particularly useful in dynamic exercises. They capture spatial hierarchies in the data, enabling them to track subtle shifts in posture and identify misalignments that traditional algorithms may overlook, especially in multi-joint movements like lunges or burpees.

B. MediaPipe

Developed by Google, MediaPipe is a versatile open-source solution that employs machine learning models for detecting and tracking key body landmarks, facilitating real-time human pose estimation. It identifies 33 critical body landmarks, including joints like the shoulders, elbows, knees, and hips, which are essential for evaluating a person's posture. This capability allows for the mapping of the body's skeletal structure in real time, making it ideal for monitoring exercise postures. MediaPipe provides 3D coordinates for each key-point, enhancing the understanding of the body's position, depth, and orientation—vital for complex exercises that demand precise movements. Its compatibility with standard webcams makes it accessible for a variety of applications, including home fitness systems, allowing users to monitor their exercises without needing specialized equipment.

C. OpenPose

OpenPose is instrumental in exercise posture recognition, delivering precise, real-time detection of body key points and skeletal movements. Its ability to track multiple individuals and detect full-body movements provides high-accuracy pose estimation, making it an excellent tool for fitness training, physiotherapy, and injury prevention. The system's capability to monitor several individuals simultaneously makes it especially useful in group fitness classes or therapy sessions. OpenPose detects not only body joints, such as shoulders, elbows, and knees, but also key points on the face, hands, and feet, facilitating detailed movement analysis during exercises like yoga and squats. Its high accuracy in detecting body joints and key points is crucial for recognizing small deviations in form that could lead to injury. Additionally, OpenPose allows researchers and developers to integrate it easily into their systems and customize it for specific exercises, making it adaptable for various fitness applications.

D. PoseNet

Developed by Google, PoseNet is a widely used computer vision model for real-time human pose estimation that plays a significant role in exercise posture recognition. It accurately detects human body key points using a simple camera, which can be integrated into posture evaluation systems for fitness, physiotherapy, and other health-related applications. Designed for real-time pose tracking, PoseNet enables immediate feedback on exercise posture and can detect the posture of one or multiple people within a single frame. It identifies 17 key body landmarks, including the shoulders, elbows, hips, knees, and ankles, which are critical for recognizing various exercise postures, such as squats, lunges, and yoga poses, while evaluating alignment and form. Although PoseNet focuses on 2D pose estimation, which is sufficient for many exercise recognition tasks, it is efficient in processing and provides low-latency results, which are essential for live feedback systems. It is also easy to implement across various platforms, including mobile apps, web applications, and desktop software.

E. MoveNet

MoveNet is crucial due to its speed, accuracy, and ease of integration into multiple platforms. Designed for fast, real-time human pose estimation, MoveNet is ideal for applications that require instant feedback in exercise posture recognition. It detects 17 key body landmarks, including essential joints like shoulders, elbows, wrists, hips, knees, and ankles, primarily focusing on single person pose estimation. This focus enhances its performance, making it particularly accurate for personal training or rehabilitation scenarios.

II. LITERATURE REVIEW

Real-Time Human Pose Estimation system contributes to society by enhancing the accessibility and effectiveness of various applications that directly impact health, safety, and overall well-being. In the realm of fitness and health The Real-Time Human Pose Estimation system operates by first receiving video frames from a camera, where each frame is processed independently. These frames undergo preprocessing steps such as resizing and normalization to ensure consistency, with possible augmentations like scaling or rotations applied to enhance the model's robustness.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 12 Dec. 2024, pp: 120-124 www.ijaem.net ISSN: 2395-5252

Using a MobileNet-based convolutional neural network (CNN), the system extracts key human body points (such as shoulders, elbows, and hips) by employing deep learning techniques. To minimize computational complexity while maintaining accuracy, lightweight architectures like Ghost Modules are incorporated. For video input, temporal smoothing techniques, such as Kalman filtering, are applied to reduce jitter and ensure smooth transitions between poses across consecutive frames. Optimized for real-time performance, the model employs tools like TensorFlow Lite to achieve fast and efficient inference on mobile or low-powered devices, ensuring minimal latency. Ultimately, the system outputs the keypoint locations of each human in the frame, enabling it to detect and track human movements in real-time for applications in fitness training, activity recognition, surveillance, and augmented reality [1].

The Pose Trainer application designed to help users perform fitness exercises correctly and safely [2]. It detects the user's exercise pose in real-time and provides personalized recommendations on how to improve their form, preventing potential injury. The app uses advanced pose estimation technology to evaluate the angles and positions of body parts during exercises.Pose Trainer works by analyzing exercise videos to assess a user's posture and provide real-time feedback. The process begins with the user recording a video of themselves performing a specific exercise, which is then processed using a pose estimation model like OpenPose. This model detects key body points (such as the shoulders, elbows, hips, and knees) in each frame, creating a skeletal representation of the user's movements. To account for variations in body size and camera angles, the key-points are normalized based on the user's torso length, ensuring consistent analysis. The system identifies the side of the body performing the movement and evaluates the angles between different body parts to determine whether the exercise is being performed with proper form. If the user's movements align with the correct posture, they receive positive feedback. If deviations are detected, such as excessive joint movement or incomplete range of motion, the system provides corrective feedback. Additionally, machine learning models can be used to classify the user's form by comparing their movements with pre-recorded examples of correct and incorrect techniques [2] .

AI-Powered Pose Analysis using MoveNet for Yoga Refinement This study explores the use of AI to create a virtual yoga trainer that mimics real-life instruction. It addresses current challenges in AI yoga models and seeks to improve their effectiveness. The research envisions future AI trainers providing a more realistic and helpful yoga experience for users.The implementation of the AI-powered YogaSiddhi system involves two main phases: the training pipeline and the web application. The training pipeline begins with the collection of a labeled dataset containing images of various yoga poses. Over 2300 images are gathered and divided into training and testing subsets. These images are then preprocessed by extracting key body points using the MoveNet model, which generates key point coordinates for 17 landmarks like the nose, shoulders, elbows, and hips. The images are resized and normalized to match the input format of MoveNet, with important regions cropped to focus on the person performing the pose. These preprocessed key points are saved in CSV files for use in model training.The yoga pose classification model is then trained using the key-point data, with a network architecture that includes dense layers, dropout layers, and a softmax output layer. The model is optimized using the Adam optimizer with a categorical cross-entropy loss function, suitable for multi-class classification. The training process takes place over 200 epochs, with validation after each epoch to prevent overfitting. After training, the model is saved in a TensorFlow.js format for integration into the web application. The web application is developed using React.js for the front-end and Node.js for the back-end. It provides real-time pose detection and feedback using the MoveNet Thunder model, allowing users to either manually select a yoga pose or let the system automatically detect their pose. Real-time feedback is offered both visually, by displaying the detected keypoints, and audibly, through spoken instructions to correct the user's pose. The system operates at 30 frames per second and can be accessed from most modern devices. With the MoveNet Thunder model integrated into the web application, the YogaSiddhi system successfully provides users with over 97% accuracy in pose detection, real-time feedback, and performance tracking, making it an effective tool for yoga practice refinement [3].

A new approach to improve the efficiency and speed of single-person pose estimation, with a focus on fitness tracking and mobility-related applications. Human pose estimation, which involves identifying key body joints and tracking their movements. However, traditional methods struggle with challenges like the variability of human movements, joint occlusions, and the computational load needed for real-time processing. To address these issues, the authors introduce the Intensive Feature Consistency (IFC) network, which refines initial pose estimates through two key stages: global body intensity correction and local body part adjustment. The former smooths the overall body structure across frames, while the latter fine-tunes smaller and more variable body parts like the hands and feet [4] .

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 12 Dec. 2024, pp: 120-124 www.ijaem.net ISSN: 2395-5252

The IFC network is designed to maintain long-term consistency in pose estimation, achieving a real-time processing speed of 31 frames per second (FPS), which is crucial for applications like fitness monitoring. By integrating both global and local refinement, the network enhances the stability of poses, even during complex movements or occlusions. Evaluated on benchmark datasets like Penn Action and Sub-JHMDB, the framework outperforms existing methods by providing more reliable tracking while reducing computational demands. This makes the IFC network a promising solution for real-time applications requiring precise and rapid pose estimation, such as augmented reality, sports analysis, and robotics. The paper emphasizes that this framework effectively balances performance and speed, opening up new possibilities for future developments in pose estimation technology. Dependency on video quality, lack of training data are the limitations.

The performance of four state-of-the-art human pose estimation libraries – OpenPose, PoseNet, MoveNet, and MediaPipe Pose which are widely used for applications like video surveillance, medical assistance, and sports motion analysis were reviewed [5]. The study focuses on comparing these libraries based on image and video datasets, using percentage of detected joints (PDJ) as the primary evaluation metric. MoveNet emerged as the top performer for both image and video datasets, showing the highest PDJ in most experiments, while MediaPipe Pose also performed well, especially in video-based actions. OpenPose was the weakest performer, particularly in video datasets where it struggled with continuous frame detection and self-occlusion issues. The paper highlights that different libraries excel in various scenarios and emphasizes the importance of selecting the right library for specific applications.

Aimed at recognizing actions that involve gathering 3D skeletal joint data across the X, Y, and Z axes, which is not dependent on the viewpoint [6]. In the preprocessing stage, changes in joint positions over time are identified by computing variances between consecutive frames. This leads to a matrix that illustrates the motion data for every joint as time progresses. Next, these matrices are utilized in the formation of training and testing datasets, where clip lengths vary from 20 to 32 frames. The clips are designed to have a half overlap to effectively capture temporal information.The custom-made deep convolutional neural network (DCNN) includes convolutional layers, batch normalization, ReLU activation, max-pooling, fully connected layers, and a softmax output layer for processing the data. The structure of the network is created to effectively categorize the skeleton data into various types of actions. The system was tested using UTD-MHAD and MSR-Action3d datasets, which include different types of human actions. The proposed model's performance was pitted against ResNet18 and MobileNetV2, demonstrating comparable accuracy with decreased computational complexity [6].

Human Pose Estimation Using MediaPipe Pose and Optimization Method Based on a Humanoid Model aimed to develop a lightweight, efficient system for 3D human pose estimation using a combination of MediaPipe Pose for 2D joint detection and an optimization method (uDEAS) for estimating 3D joint angles based on a humanoid model. The goal was to provide an accessible solution for monitoring human poses, particularly for applications like elderly care, without relying on expensive motion capture systems.The system combines MediaPipe Pose for 2D human pose estimation with the uDEAS optimization technique for precise 3D joint angle estimation using a detailed humanoid model. MediaPipe Pose recognizes 33 2D landmarks, with 12 key landmarks specifically utilized for pose analysis. The optimization reduces the gap between the identified 2D joint positions and the 3D model's projection by modifying joint angles and clarifying depth uncertainties. This is accomplished by using a loss function that considers deviations in the center of mass and penalty functions for joint rotation limits in order to guarantee authentic and organic poses. Efficiently using RGB camera data, the system offers a lightweight, real-time 3D pose estimation solution, eliminating the requirement for costly motion capture technology [7].

BlazePose is a lightweight CNN designed for real-time human pose estimation, optimized for mobile devices. It uses an efficient architecture that combines heatmaps and regression to estimate 33 body keypoints, running at over 30 FPS on devices like the Google Pixel 2. The model eliminates heatmap generation during inference, making it lightweight, and avoids complex NMS algorithms. BlazePose is ideal for applications like fitness tracking, sign language recognition, and augmented reality, offering keypoint topology consistent with other models. Trained on 60,000 images with occlusion simulations, it outperforms OpenPose in tasks like yoga and fitness while being up to 75 times faster on mobile CPUs. The paper suggests extending BlazePose to 3D pose estimation and other complex tasks due to its flexible, scalable architecture [8].

However, BlazePose is optimized for single-person tracking and doesn't perform well in multi-person scenarios, limiting its use when multiple individuals need to be tracked simultaneously. It also assumes that the head is always visible in the frame, reducing its effectiveness in cases where the head is occluded or out of view.

A survey on yogic posture recognition focuses heavily on the importance of correct posture in yoga practice. Proper body alignment is critical in yoga [9]. The paper reviews the key technological approaches used in human pose estimation (HPE), which is a central challenge in yogic posture recognition. Human pose estimation involves detecting and evaluating the position of body parts or key points in images or video. Several types of models and sensors used in yoga posture recognition are reviewed. These include skeleton-based models, contour-based models, and volume-based models, all of which play a role in accurately modeling the human body f or pose estimation. Sensor-based systems, such as accelerometers, gyroscopes, and magnetometers, are also used for posture recognition, often in wearable devices or smart mats that monitor a practitioner’s movements and provide feedback. The paper concludes by discussing the potential for further research and development in yogic posture recognition. It emphasizes the need for more advanced systems that can handle complex poses and provide accurate feedback across a wider range of yoga postures. Cost and accessibility, individual differences are the disadvantages.

III. CONCLUSION

In summary, the use of cutting-edge technologies like Convolutional Neural Networks (CNNs), along with tools like MediaPipe, OpenPose, PoseNet, and MoveNet, is revolutionizing how we monitor and correct exercise posture. These systems provide accurate posture analysis and instant feedback, making them incredibly useful for fitness, rehabilitation, and preventing injuries. By leveraging AI and computer vision, these solutions reduce the need for personal trainers, offering a more accessible and cost-effective way to ensure proper posture, especially for older adults. As these technologies continue to evolve, their impact on health and fitness will only grow, making them even more valuable for everyday use.

REFERENCES

[1] Real-Time Human Pose Estimation Using Deep Learning COMP6982/ENGI9805 PROJECT REPORT, APRIL 2024
[2] Pose Trainer: Correcting Exercise Posture using Pose Estimation arXiv:submit/3236829 [cs.CV] 21 Jun 2020
[3] YogaSiddhi: AI-Powered Pose Analysis using MoveNet for Yoga Refinement International Journal of Computer Applications (0975 – 8887) Volume 186 – No.8, February 2024
[4] Sheu, Ming-Hwa, SM SalahuddinMorsalin, Chung-Chian Hsu, Shin-Chi Lai, Szu-Hong Wang, and Chuan-Yu Chang. “Improvement of human pose estimation and processing with the intensive feature consistency network.” IEEE Access 11 (2023): 28045-28059.
[5] Chung, Jen-Li, Lee-Yeng Ong, and Meng-Chew Leow. “Comparative analysis of skeleton-based human pose estimation.” Future Internet 14, no. 12 (2022): 380
[6] Tasnim, Nusrat, Md Mahbubul Islam, and Joong-Hwan Baek. “Deep learning-based action recognition using 3D skeleton joints information.” Inventions 5, no. 3 (2020): 49.
[7] Kim, Jong-Wook, Jin-Young Choi, Eun-Ju Ha, and Jae-Ho Choi. “Human Pose Estimation Using MediaPipe Pose and Optimization Method Based on a Humanoid Model.” Applied Sciences (2076-3417) 13, no. 4 (2023).
[8] Bazarevsky, V. "BlazePose: On-device Real-time Body Pose tracking." arXiv preprint arXiv:2006.10204 (2020).
[9] Rajendran, Arun Kumar, and SibiChakkaravarthySethuraman. "A survey on yogic posture recognition." IEEE Access 11 (2023): 11183-11223.
[10] Kolla, Morarjee, Phani Varma Gadiraju, and DhruvrajTondanoorthy. "Form Check: Exercise Posture Correction Application." Journal of Electrical Systems 20, no. 4s (2024): 25-33.