International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 11 Nov. 2024, pp: 314-320 www.ijaem.net ISSN: 2395-5252

Survey on AI Driven Gym Assistant with Pose Recognition

Devi Krishna R, Devinandana D, Jayalakshmi V, Merin Mary Sabu, Asst Prof. Shiny B
B.Tech Computer Science Engineering
College of Engineering,Chengannur

Date of Submission: 10-11-2024 Date of Acceptance: 20-11-2024

ABSTRACT—In our research, we reviewed studies focusing on pose estimation and correction of forms in exercising using deep learning and machine learning approaches. One such model was the CNN-LSTM-based model that used a combination of spatial and temporal attributes for the recognition action with satisfactory accuracy. Another system utilized OpenPose in real time for pose estimation and improved postures of exercises like bicep curls, shoulder presses, etc. Other researches used 3D pose estimation methods, such as VIBE, for the detection of repeated exercises with a high rate of precision. Other researches utilized wearable sensor data and machine learning techniques to monitor exercise posture. One research focused on arm and shoulder exercises of the user. AI-based models such as Bi-GRU were capable of providing real-time feedback while the user was performing squats so that he did not hurt himself. Advancement on tracking and exercise performance improvement capabilities was also evaluated through humanoid fitness trainers and pose-guided graph convolutional networks. Examples of such studies hereby establish the potentiality of AI-driven systems for health monitoring and personalized exercise feedback.

I. INTRODUCTION

Advanced technologies such as deep learning and machine learning represent a transformative shift in the world of fitness and rehabilitation. Increasingly, it has become evident that the recognition of physical activity should be optimized with regards to workout efficiency, and users should be safe while undergoing it. This is the reason new approaches have emerged, which apply sophisticated algorithms to monitor and improve the performance during exercise. Such classical assessment of movement has been merely based on subjective observations or even basic tracking of motion; therefore, there could be inaccuracies and increased risks of injury. In the current landscape, deep learning techniques, including Convolutional Neural Networks (CNNs) and Long Short-

Term Memory (LSTM) networks, have revolutionized this, as they provide automated real-time feedback related to physical activities.

Recent Advancements of Pose Estimation Technologies Being that recent advancements have been instituted in Open Pose and VIBE, this science has been taken to a different level as it enables body movements to be studied in enriched detail through video input. The program tracks the key landmarks of the body, studies joint angles, and measures the precision of form in exercises in high definition. Continuous improvements allow for machine learning algorithms to be infused for the customization of training procedures that can include correcting committal mistakes of exercise and giving specific feedback based on individualistic needs.

Further, multimodal data utilization in the form of skeleton data merge with pose will significantly improve the recognition capability of such systems so that they would better distinguish actual from wrong movements. Such technology integration would allow a person to exercise safely and efficiently without constant attention but still makes it possible to use completely new approaches in rehabilitation where the right correction at any moment can be generated from a distance. With home-based workouts increasingly being adopted, especially as recently trends all over the world have gone along this direction, these technologies could enormously reshape personal fitness and rehabilitation regimes. Integration of AI-driven methodologies into home-based workouts can lead to improved performance, lower injury risks, and a much more engaging and interactive experience.

A. CNN-LSTM Model for Action Recognition

A combination of Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks was adopted for identifying the complex actions while utilizing the traditional Chinese

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 11 Nov. 2024, pp: 314-320 www.ijaem.net ISSN: 2395-5252

exercises Baduanjin. The CNN extracted spatial features from video

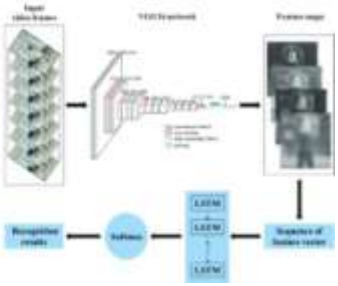

Fig. 1. CNN-LSTM Model

frames, while LSTM processed the temporal dependencies to identify the actions as shown in Fig.1. Such a model may be suitable at an accuracy of 96.43%, surpassing traditional methods, for intelligent rehabilitation and health monitoring.

B. Pose Estimation for Exercise Posture Correction

A real-time body keypoints tracker using OpenPose for a posture correction system is proposed. The system assesses the poses with geometric heuristics and machine learning to offer personal advice on exercise, including the bicep curl and shoulder press. With more than 100 videos as its training set, it tracks posture flaws and helps avoid injury. It allows running both on Windows and Linux, offering enhanced home-based training with access to a GPU.

C. Repetitive Action Segmentation and Counting

The system identifies and counts recurring physical activities by segmenting them, based on video data. It uses VIBE 3D pose estimation models to track humans' skeleton movements and gives a real-time feedback over joint angles and repetition counts. While it was tested on the NOL-18 dataset, it was able to yield an accuracy of 96.27% action segmentation and 99.06% activity recognition, making it preferred for use in fitness tracking and rehabilitation.

D. Time-Series Data for Gym Exercise Recognition

A recognition system for gym exercises with feature engineering from the time-series data of the Zephyr BioHarness 3 device. Researchers convert accelerometer data into 106 features and classify using these features. The best classifier is K-SVC (Kernel Support Vector Classifier), that obtained over 91% accuracy in classifying core-body, chest, and back exercises.Extraction of Features were done as shown in Fig.2.

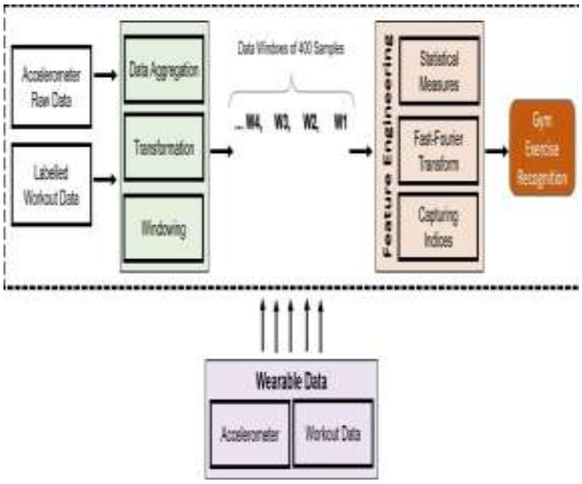

Fig. 2. Time Series Data Architecture

E. Squat Movement Classification Using AI

A system of analyzing and correcting squat movements was developed using stereo cameras and the MediaPipe framework to generate 3D key points. An autocoder, based on a Bi-GRU model, following an attention policy attains 94% for determining good and bad form for squats. The system offers real-time feedback help in injury avoidance for users

F. Pose-Guided Graph Convolutional Network (PGGCN)

With skeleton-based action recognition, the PG-GCN model fuses skeleton and pose data together in a dynamically designed attention module that improves feature representation and reduces redundancy. The model does well in recognizing actions with apparently good results, showing that incorporating pose data is effective.

II. LITERATURE REVIEW

J. Chen et al. [1] introduces a CNN-LSTM model for action recognition from video recordings with particular focus on the conventional Baduanjin characterized by complicated and coordinated movement. As an attempt to capture spatial and temporal features within the action, this model was based on the combination of CNNs and LSTM networks. CNNs are used to extract spatial features from single video frames. The extracted features are further processed by LSTMs that try to find the relationship between consecutive frames in order to decide on their dependency. This will enable the model to appropriately identify the intricate actions involved in different activities with continuous fluid movements.

This model showed an excellent test accuracy of about 96.43%, which is largely surpassed by the hand-crafted geometric motion-based action recognition techniques that rely on such features,

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 11 Nov. 2024, pp: 314-320 www.ijaem.net ISSN: 2395-5252

achieving an accuracy of only 66.07%. Here, the difference in performance clearly validates that deep learning techniques are better suited for recognizing the complex actions as this one has automatic feature learning without any kind of human intervention.

This system may be greatly worth utilization for applications in intelligent rehabilitation and health monitoring by accurately being able to recognize movements in exercises such as Baduanjin. Besides, it will be useful for the evaluation of skills in any kind of physical activity. In addition, the model is very fitting into the context of self-guided training programs offered to the users so that they can practice exercise sessions without waiting for professionals. By allowing extraction of spatial and temporal features, one is able to understand motion better, which it can be a perfect solution for improvements in self-training and rehabilitation systems.

R. R. Yang et al. [2] developed a system to improve exercise form leveraging pose estimation and machine learning techniques. The system utilizes the latest Open Pose, a state-of-the-art pose estimation model that can track all key points of a user's body in real time using video input. These poses are further analyzed through both geometric heuristics and algorithms using machine learning and generate customized feedback for the improvement of posture during bicep curls, shoulder presses, and shrugs.

The system was trained based on a large number of videos-more than 100 video streams that depict different exercises performed both correctly and incorrectly. Such a large dataset allows the system to recognize the most common mistakes, for example, an incorrect angle between the user's joints or too much movement, and then give timely, according to the situation, feedback in order to correct these mistakes to the user. In turn, this enables users to avoid accidental injuries and deliver better overall performance. Being a machine learning-based model, it will constantly update and refine feedback about itself. This makes it well-suited to address the individual's needs as well as common exercise errors.

Pose Trainer is designed to be relatively easily deployable on systems running Windows or Linux, assuming they have access to a GPU. Access through this service enables the use of the full potential of the system on wide-spectrum platforms, thereby accommodating and offering a flexible solution to maximize workout effectiveness. This service is utilized for improving personal fitness routines or implemented as part of more extensive systems within the confines of the gym. The Pose Trainer offers innovative support and assistance in the practice of self-guided training, void of constant professional supervision. It's a great system for anyone to have an improved exercise form and to work on their reducing risk of injury.

M. A. Sarwar et al. [3] proposed a system to segment, count, and recognize repetitive physical activities from video data. In the use of deep learning together with a 3D pose estimation model, known as VIBE, the tracking of human skeletons during exercise makes the system count repetitions. Therefore, the system gives detailed feedback on various parameters, such as joint angles, time intervals, and sets, thus giving an accurate way for tracking and evaluating exercise routines with considerable depth about their movements. The system was tested using the NOL-18 dataset, which has a variety of captured physical activities from multiple camera angles. Results: Overall performance is good as reflected by 96.27 % action segmentation accuracy and 99.06 % in activity recognition accuracy. These measures indicate how the system can strongly determine and evaluate physical activities under complex conditions.

This system advances a way of monitoring the performance of workouts and also provides real-time feedback that involves movement accuracy and repetition counts for users to attain a better form, thus tracking change over time. This solution tracks activities using videobased data with more refined and accurate analysis as well as sophisticated techniques of pose estimation, compared to traditional fitness devices or wearable gadgets.

This system is invaluable for users desiring fitness optimization, minimizing injury, and better tracking of improvement. This system works to address the problem of complex physical activity recognition and delivers insights beyond what existing fitness gadgets or wearable devices can deliver. This, with its advanced tracking and monitoring capabilities, will probably make an enormous difference to the approach users take when following a fitness or rehabilitation regime. It provides much more detail in feedback than has been available previously.

M. F. Trujillo-Guerrero et al. [4] described a comparison of the ability to correctly classify Human Activity Recognition on sensor data from Inertial Measurement

Units and Movement Analysis Systems of Convolutional Neural Networks, Long Short-Term Memory models, and the Transformer is made into study. In the course of this research, the study will center on the recognition of arm exercise by gathering data from 10 participants who performed six specific upper limb movements. Key preprocessing steps include normalization and feature selection using random forests to reduce sensor input.

This work evaluated eight state-of-the-art deep learning models across different architectures for performance comparison. The results indicate that the highest accuracy was achieved by CNN and LSTM hybrid architectures with test accuracies ranging from 89% to 99% . A study of feature reduction was conducted, which indeed showed that even at four sensor features, the high accuracy was maintained. IMUs are a better option compared with MAS regarding flexibility and overcome the occlusion problem as well; however, MAS worked well in scenarios that have large movements.

The contribution of the work is to note the importance of sampling rates and window sizes towards real activity recognition. Perhaps one of the most important contributions of the work is its demonstration that one IMU at the wrist is enough for accurate exercise recognition, thus allowing more efficient HAR systems that are portable.

A. B. Asghar et al. [5] introduced a comparative performance analysis of machine learning algorithms for arm and shoulder exercises using the wrist-worn band aims at concerning the issue of posture maintenance during exercise performances in home settings, which picked up pace during the COVID-19 pandemic. The study is toward developing a solution based on data provided by a wrist-worn band equipped with an MPU6050 sensor, which contains acceleration data. After some preprocessing of the collected data, it is applied to four machine learning algorithms, namely Weighted KNN, Bagged Trees, Decision Trees, and Fine Gaussian SVM.

Out of these, the highest accuracy was achieved by the Weighted KNN algorithm with an accuracy of 92%.

The method here involved building an exercise app that would give immediate feedback in regards to exercising performance, so one could exercise from home without a personal trainer as the risks of injury were also kept at bay.It is possible to notice that the results of this study demonstrate the efficiency of machine learning algorithms concerning exercise data analysis from a wrist-worn band and demonstrates their potentiality in creating personalized exercise programs. In comparison, it could be noticed how machine learning techniques greatly improve the monitoring ability of arm and shoulder exercises, as well as defining which algorithm is the most effective in correct classification.

Other related work in wearable technology and its application on exercise monitoring has also been discussed in this context. Future work would include usage of more comprehensive datasets as well as development of personalized exercise plans using more advanced machine learning techniques, ending with the view that home-based exercise regimens could be transformed by wrist-worn devices and the models provide quite accurate performance and posture feedback in real time.

A. Hussain et al. [6] introduced a Time-Series Data to Refined Insights-A Feature Engineering-Driven Approach to Gym Exercise Recognition that derives a machine learning system aimed at gym exercise recognition using feature engineering techniques applied on timeseries data. Data was collected using a device known as the Zephyr BioHarness 3, which features a 3-axis accelerometer, over six weeks of workout routines that targeted different muscle groups. The focus of the research would be on raw accelerometer data of vertical, lateral, and sagittal axes transformed into 106 features through statistical measures, Fast Fourier Transform, and more. This helps in better modeling and classification of the gym activities.

In all, the researchers performed four different experiments to ascertain the classification performance of six classifiers, built from GridSearchCV, which include Random Forest, Decision Tree, Support Vector Machine, and so on. The experiment has been performed in such a way so as to test not only the models specific to muscle-groups and athlete-specific models but also generalized independent models. The classification models were trained on the transformed dataset and then tested for accuracy in their abilities to recognize 44 gym exercises. Out of all these classifiers, K-SVC (Kernel Support Vector Classifier) developed the best results with very high accuracies developed in the identification of exercises at the gym; muscle group again remained the most important factor in this case.

The study shows that feature engineering has high importance in terms of predictive accuracy, as enhanced from only using raw time-series data. In fact, the proposed system was able to achieve over 91% accuracy for exercises dealing with core-body, chest, and back muscles. Therefore, feature transformation plays a significant role in improving the performance and the application of machine learning models regarding the analysis of time-series data beyond gym activity recognition.

M. Chariar et al. [7] introduced an approach that relies on deep learning and computer vision that analyze and correct squat exercises with particular attention to common defects possibly causing injury. Information was gathered from over 50 volunteers who performed squats in various conditions and obtained 1332 video samples shot using stereo cameras calibrated on the MediaPipe framework to generate 3D key points. These videos were also annotated by the experts. The movements are labeled as ”good” or ”bad.” The dataset was preprocessed and split into training and test set by padding for consistency.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 11 Nov. 2024, pp: 314-320 www.ijaem.net ISSN: 2395-5252

The best model for squat movement classification was developed as a custom auto encoder constructed on Bi-GRU architecture and further extended with the addition of an attention mechanism. This model achieved validation accuracy at 94 % and was found to be quite effective at distinguishing between good and poor squat form. To analyze the quality of the squat, the system applied polynomial curve fitting to key points, taking into account the size variability among people. Confusion matrices were also presented to clearly present details regarding the performance of the model.

This trainer, AI-driven, offers real-time feed backs, so that it is possible for the users to learn proper squatting form and avoid injury. Such a methodology has proved to be effective in identifying form deviation and thus providing corrective guidance on the matter. Future improvement plan could include using multi-view setups with more advanced analysis capabilities, as well as exploring alternative architectures of neural networks to increase the machine's accuracy and performance. It can, therefore, really enhance the way users perform squats through the provision of accurate, data-driven feedback for an even more efficient training experience.

A. G. Abhinand et al. [8] developed a humanoid fitness trainer that incorporates AI, which can be used as an assistant in home-based exercise performance, using human pose estimation. The emergence of remote working from home brought about by COVID-19 led to the challenge of home-based workouts and maintaining proper form while performing exercises without professional trainers. The AI trainer is based on a robust model by MediaPipe known as Blaze Pose, combining machine learning and computer vision. It captures the user through their webcam while identifying key body landmarks and their joint angles for comparison to predefined ideal postures for a variety of exercises.

The system offers immediate feedback, pointing out the appropriate form whenever deviated from. This is bound to help users embrace better posture so as to reduce risks of getting hurt. Such intelligent fitness devices are not like general models such as smartwatches whose very basis would be lost if such comparisons were made with detailed versions. It's designed to be accessible with a very basic camera, but it does particularly well for those looking for a low-cost and flexible alternative to professional trainers.

This however, focuses more on how the model is easy, flexible, and accurate in improving the workout routines while not requiring more hardware. It also improves the safety and effectiveness of the home workout by analyzing the user's form and providing instant feedback. In this respect, it has a great potential to revolutionize how people approach fitness with its automated and accessible, more interactive means for guiding exercise, especially in remote or home settings.

H. Chen et al. [9] introduced a novel Pose-Guided Graph Convolutional Network PG-GCN model that aims to enhance skeleton-based action recognition. Traditional GCNs for this task rely heavily on skeleton data only, and as an outcome, fail to preserve the complexity of sophisticated relationships between joints and integrate other useful data such as pose information. The authors address the shortcomings in the proposed multi-stream network processing skeleton and pose information. The proposed model has an attention module, which dynamically fuses the two data streams at the early stages; thus, this approach produces a more solid feature representation. The dynamically adjusted skeleton graph, once combined with the pose information, further amplifies the flexibility and competence of the model.

Extensive experiments on the NTU RGB+D 60 and NTU RGB+D 120 datasets demonstrate that PG-GCN outperforms existing state-of-the-art methods in the tasks of action recognition. We could see that the use of pose data greatly improved the generalization ability across the classes and environments. Most especially, the dynamic attention mechanism is more effective in reducing feature redundancy and extracting more discriminative features, thus improving the model performance.

Altogether,It demonstrates that the pose data combined with skeleton data is of prime importance for action recognition with greater precision. The proposed PG-GCN offers a more powerful and flexible approach to human action recognition than previous works by the use of multi-modal data along with a dynamic learnable attention mechanism. The model sets a new benchmark for integrating pose-guided information into GCN-based frameworks. It provides significant improvement in comparison to the previously used state-of-the-art methods.

A. Raza et al. [10] introduced a design that involves prevention of injuries as well as optimizing the performance of exercise by correcting the physiotherapy exercises using human pose estimation. Advanced Machine Learning and deep learning techniques that are used in advancing the accuracy of pose estimation on over a multi-class human skeleton movements dataset. The approach brings in a novel feature selection method called LogRF that reduces the dimensionality, and selects the top twenty features for training the model. Of all models tested, Random Forest showed the best efficacy at an accuracy of 0.998.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 11 Nov. 2024, pp: 314-320 www.ijaem.net ISSN: 2395-5252

Besides Random Forest, the system also tried other models like GRU, pointing out how AI technology is going to advance physiotherapy. The instant feedback the system provides helps in a remote monitoring and guiding process conducted by physiotherapists regarding proper exercise form and reduced chances of injury. This feature is crucial for distance-based rehabilitation, where patients could exercise from their homes and receive real-time correction by the system.

Ongoing developments focus on including an interactive interface with real-time camera input. This gives the system the ability to correct exercises in real-time, thereby improving its usability and effectiveness. The use of AI-driven methodology gives a more precise and efficient manner of monitoring physiotherapy exercises as opposed to the traditional ones.

III. FUTURE SCOPE

- Integration of Multi-View and 3D Pose Estimation Advanced applications may include the combining of views to capture pose from different angles, an increase in precision of pose estimation and recognition. Further advancements in 3D pose estimation algorithms may further increase precision in complex movement identification, with or without partial occlusion in the scene or challenging camera angles.
- AI-Driven Personalized Exercise Programs The system can be further extended to provide fully personalized workout plans based on the personal performance metrics and tracking of progress. Further refinements can consider adaptation to the users' capabilities and give relevant recommendations that allow for more tailored fitness or rehabilitation routines. Users would also automatically get their difficulty levels adjusted based on improvements.
- Interactive Real-Time Correction with Augmented Reality It can use AR for real-time visual feedback for making user experience more exciting and immediate correction of exercises. Because users will see virtual overlays of their posture in comparison with ideal form, the system will be more interactive and upgrade user engagement in fitness training and rehabilitation.

IV. CONCLUSION

Deep learning and machine learning technology has greatly improved the accuracy and efficiency of recognizing and correcting the physical activity that occurs as a result of improved monitoring and rehabilitation techniques in exercise. When complex movements are considered, systems that consist of CNNs and LSTMs, such as Open Pose and VIBE pose estimation, show impressive identification ability compared with classical geometric techniques. Some of these technologies provide real-time feedback about posture and movement, which is critical to avoiding risk of injury and optimizing performance in trainings. Algorithms like Random Forest and several hybrid architectures have scored quite well on personalized exercise monitoring, so they could be used for working out at home, independent of constant professional supervision. The integration of pose-guided information further improves action recognition capabilities, though actual analysis would rely on multimodal data. Therefore, the applications are positioned to revolutionize fitness and rehabilitation by giving users tools for correct and safe exercise while allowing self-guided training and effective rehabilitation from home.

REFERENCES

[1] J. Chen, J. Wang, Q. Yuan, and Z. Yang, "CNN-LSTM Model for Recognizing Video-Recorded Actions Performed in a Traditional Chinese Exercise," IEEE Journal of Translational Engineering in Health and Medicine, vol. 11, pp. 351-359, June 2023.

[2] R. R. Yang and S. Chen , "Pose Trainer: Correcting Exercise Posture using Pose Estimation," 2020.

[3] M. A. Sarwar,S. H. Cheng, Y. A. Daraghmi, T. U. lk, and Y. L. Li, "Periodic Physical Activity Information Segmentation, Counting and Recognition From Video," in IEEE Access, vol. 11, 2023.

[4] M. F. Trujillo-Guerrero, S. Roman-Niemes, M. Ja´ en-Vargas, A.´ Cadiz, R. Fonseca and J. J. Serrano-Olmedo, "Accuracy Comparison of CNN, LSTM, and Transformer for Activity Recognition Using IMU and Visual Markers," in IEEE Access, vol. 11,2023.

[5] A. B. Asghar, "Comparative Performance Analysis of Machine Learning Algorithms for Arm and Shoulder Exercises Using Wrist-Worn Band," in IEEE Access, vol. 11,2023.

[6] A. Hussain, M. A. Zahid, U. Ahmed, S. Nazeer, K. Zafar, and A. R. Baig, "Time-Series Data to Refined Insights: A Feature Engineering-Driven Approach to Gym Exercise Recognition," in IEEE Access, vol. 12,2024.

[7] M. Chariar, S. Rao, A. Irani, S. Suresh, and C. S. Asha, "AI Trainer: Autoencoder-Based Approach for Squat Analysis and Correction," in IEEE Access, vol. 11, 2023.

[8] A. G. Abhinand, M. Anas, N. K. B, R. G, and V. Jituri, "AI Fitness Trainer Using Human Pose Estimation," 2023.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 11 Nov. 2024, pp: 314-320 www.ijaem.net ISSN: 2395-5252

[9] H. Chen, Y. Jiang, and H. Ko, ”Pose-Guided Graph Convolutional Networks for Skeleton-Based Action Recognition,” in IEEE Access, vol. 10, 2022.
[10] A. Raza, A. M. Qadri, I. Akhtar, N. A. Samee, and M. Alabdulhafith, ”LogRF: An Approach to Human Pose Estimation Using Skeleton Landmarks for Physiotherapy Fitness Exercise Correction,” in IEEE Access, vol. 11, 2023.