IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Posture Intelligence: Next-Gen Human Posture Detection And Correction Using Lstm And Mediapipe

Ankit Gangwar, Sakshi Sharma, Jitendra Kumar

Student, Student, Assistant Professor
Computer Science Engineering,
Galgotias University, Greater Noida, India

Abstract: Now day's many exercises are done at home or by video, it is important to keep the right posture when doing squats, push-ups and pull-ups to avoid injuring yourself and gain the most from each workout. This research features a system that, in real time, utilizes MediaPipe for landmarking body parts and uses LSTM networks to recognize movement changes over time. MediaPipe finds the locations of important joints in each image frame, allowing the user's pose to be logged for all exercises. LSTM is trained to process the important points in the time-series data and separate the posture deviations into three types of movements: correct, suboptimal, and incorrect. The system delivers instant, useful information to help users correct their movements in workouts. Made for use in gyms, this approach allows for simple and inexpensive fitness, online coaching, and rehab services. Accuracy and quick results of the method shown in experiments indicate it could be a valuable resource for AI-assisted fitness training.

Index Terms - LSTM, MediaPipe, real-time posture monitoring, computer vision, artificial intelligence, and machine learning.

I. Introduction

Good posture is important for successful exercise, avoiding injury and protecting your body as you get older. Squats, push-ups and pull-ups are gym exercises where proper form is needed both to develop muscles properly and to prevent most joint, alignment and imbalance injuries. When you don't have an expert there to guide you, those who work out at home are more likely to use wrong techniques, which may cause injury and reduce the benefits of their exercise.

In fitness program, most past posture correction approaches involved sports trainers or trackers focused on general movements. Wearables will measure how your body is moving, but they will not know where or why these movements are made, so feedback is not accurate. They should make contact with the body, but this can sometimes be uncomfortable or troubling during your exercise.

As a way to address these difficulties, this paper describes an intelligent solution using MediaPipe's fast vision technology and LSTM networks. MediaPipe detects important locations on the human body live from a regular camera, which provides a gentle way to get posture data. An LSTM is used next to review the key points one after another in sequences and decide on the correctness of the motion style based on patterns it has learned. MediaPipe and LSTM permit the system to determine how an exercise is performed each time, unlike regular image-based classifiers. Because of this, we can precisely judge small differences in posture such as the position of the knee during squats and the angle of elbows during push-ups.

The purpose of this study is to develop a camera system that would correct posture mistakes right away for those working out at home or in a gym using photos. The system uses MediaPipe and LSTM to help people notice their movements, correct their form and avoid injuries without depending on wearables or instruction from a trainer.

A lot of these overuse injuries come from working out and training, mainly for those who do not exercise correctly. Misdoing squats, push-ups or pull-ups can cause you to suffer from ongoing pain and joint problems, as well as lasting harm to your body [1]. Because many people are working out at home, it's important to guide them in proper posture, because finding a live coach is getting harder. Due to this, people often use unsafe methods during their workouts, which leads to more harm and weaker results [2]. Challenges in form during movement can now be spotted and corrected through recent improvements in computers and machine learning. Compared to other parameters, posture and joint angle analysis in challenging movements is not well done by wearable sensors in real time [3]. While previous solutions needed other equipment, MediaPipe can now quickly and accurately evaluate shots from videos to analyze our training [4]. However, for a more accurate picture and rating of exercise order, additional modeling is needed. The work uses MediaPipe to detect the joints on the body and a Long Short-Term Memory network to decide if each movement is correct. On analyzing movement over time, including gait, sports actions and rehabilitation, LSTM networks have achieved good outcomes [5]. As a result, users are notified instantly if they perform actions incorrectly and can improve their skills.

II. LITERATURE REVIEW

1.LITERATURE SYNTHESIS

Well-formed posture while doing squats, push-ups and deadlifts helps prevent common muscle and bone related problems, as many have suggested. Many injuries in the gym and in sports involving the shoulders, back and knees can result from poor posture while doing squats and similar exercises [6]. So, there is increased research on developing tools to correct and check exercise form. So far, watching your posture up close with a personal trainer was not a choice for most people [7]. At present, technologies like computer vision are being introduced to help analyze and improve how posture is observed during a workout. Researchers have recently found [8] that 2D and 3D pose estimation can observe exercise and pinpoint errors in average movement. It is clear that, due to being able to use any common RGB camera, MediaPipe is a leader in real-time pose detection, which makes it great for low-cost fitness gadgets [9]. Additionally, machine learning, and especially LSTM algorithms, makes it possible to model exercise movements step by step and monitor postures so systems can check the correctness of the exercises [10]. The use of pose estimation and sequence modelling means we can now develop intelligent systems to improve safety and the results of exercise. That's why vision-based systems that highlight poor posture in real time are becoming very popular.

2.IDENTIFIED GAPS

It shows how keeping the appropriate posture while strength training reduces your risk of musculoskeletal disorders (MSDs). If you exercise with bad technique, your risk of knee issues, shoulder pain and lower back problems greatly increases [11]. Therefore, scientists are interested in finding strategies to check and adjust how people move during training. Until now, people would learn posture by themselves or with trainers, but it was tricky to make this process easy for many people [12]. Image processing, computer vision, and machine learning are speeding up the process of analyzing posture in gyms. In new findings, we see that 2D and 3D pose estimation models can tell when motions in exercise techniques do not line up with the recommended routines [13]. In addition, MediaPipe is now able to estimate poses in real time, using RGB cameras to identify main body points, which benefits low-cost exercise systems [14]. Experts have attempted to model the changing nature of human motion by using machine learning and, in many cases, LSTM recurrent neural networks. Due to posture flow in machine networks, the systems can identify whether exercise movements are correct or incorrect [15]. Because of better links between pose estimation and sequence modeling, we can now develop more adaptable and intelligent assistive technology solutions during exercise. Because of this, using vision to check posture as a person exercise is now common, improving safety and making it easier to stay aware.

III. SYSTEM DESIGN AND METHODOLOGY

1.System Architecture

Our system uses an LSTM-based model and MediaPipe to watch and determine how well people handle manual lifts. Being linked to webcams, computer vision systems process posture information all the time, even at night. Having examined many types of lifts and poses, the LSTM decides whether specific postures are correct for the lifting motion. The main aim of an effective strategy is to harmonize all these elements within the operations of the system. The approach contained in the model can be seen in Figure 1.

Several key advantages make this system suitable for helping gyms, so the MediaPipe-LSTM choice is made for recognizing postures there. MediaPipe is a computer vision software that allows very accurate monitoring of the body joints during exercise. Since making mistakes with the way you lift heavy objects in a gym may lead to strained muscles or MSDs, it’s very important to correct your form. It recognizes poor movement and stops the user before an injury occurs. If your training includes shoulder presses, squats or deadlifts, proper alignment is very important. Monitoring by your coach spots your mistakes and helps. Since heart rate measurements are immediate, athletes don’t have to wait until after their workout to cheque their results and take action. Analysis of user posture data is easy because it happens immediately and is clearly displayed online. With this tool, people in the gym gain better focus on how they work out. In addition, trainers can watch and instruct their clients remotely using different technologies.

The system is built around simple webcams and is set up to be purchased in big numbers and easily scaled up. Its accessibility is made possible by the lack of expensive motion capture tools, so gyms, personal trainers and individual users can all use it. Moreover, the system can be changed to match what users require and the introduction of improved technology.

The system is built around simple webcams and is set up to be purchased in big numbers and easily scaled up. Its accessibility is made possible by the lack of expensive motion capture tools, so gyms, personal trainers and individual users can all use it. Moreover, the system can be changed to match what users require and the introduction of improved technology.

Thanks to their sequence management, LSTM networks are the best fit for this job in machine learning. Rather than observing each standing pose alone, posture recognition during fitness gym classes looks closely at the full range of movements. Because of the LSTM memory cells, the system matches its previous observations to the new instances which improves its ability to recognize motion. When moving through different parts of a squat, each movement helps you learn something about your body, so make sure to pay attention to time.

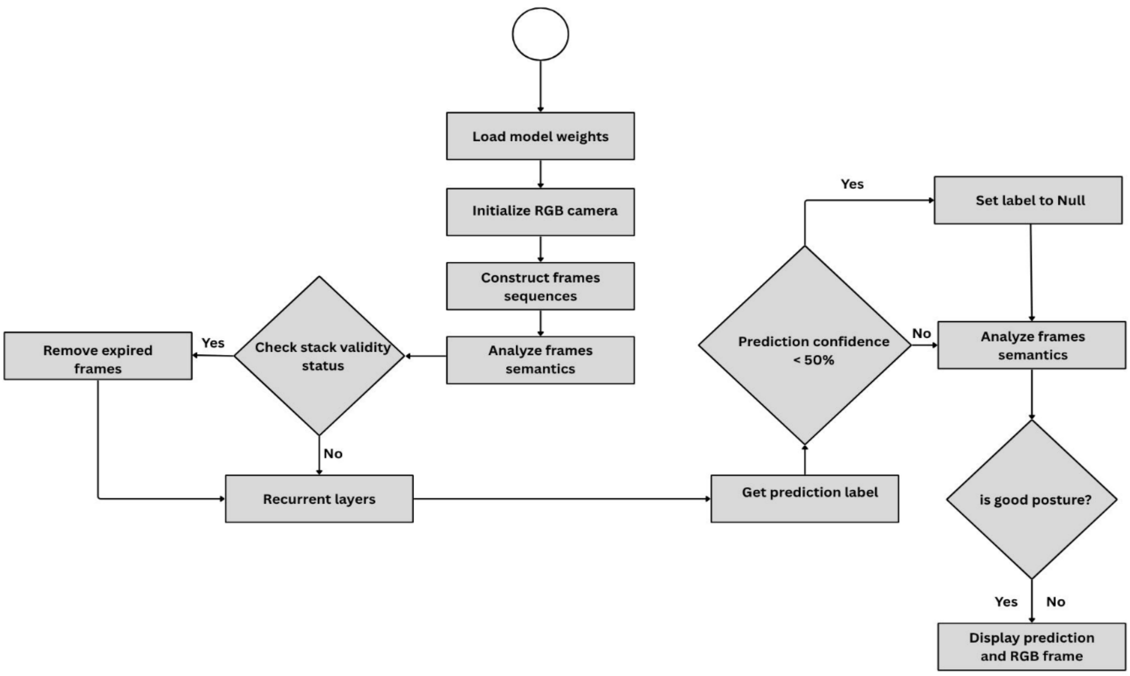

Figure 1. Workflow diagram showing the information flow from the webcam to the AI model for a worker's lifting posture state.

2. Dataset Details

The dataset for this project had 64 videos containing both properly done postures and improper ones. To complete the study, videos were taken by different cameras, from different positions and with people of all heights. MediaPipe is used by the model to identify the key points and put each list of values into specific folders for later use. A wide variety of actions are covered in the dataset by sequences and movements. The detailed breakdown of the dataset is shown in Table 1.

Table 1. Detailed Breakdown of Dataset.

Posture Type	Number of Videos	Frames per Video	Landmarks per Frame	Total Key point Values per Frame
Good Postures	32	30	33	132
Bad Posture	30	30	33	132

3. Implementation

In the first part of the implementation, real-time pose estimation is implemented with the MediaPipe Holistic model. Basic implementation for pose estimation uses the designated components of the MediaPipe library, allowing video frames to be captured, processed through the MediaPipe Holistic model, and rendered with face, pose, and hand landmark visualizations.

The next part of the implementation is the LSTM based neural network. The LSTM based model combines key point values taken from each pose frame (temporal patterns) and process them through a series of LSTM layers followed by dense layers. The model is compiled using the Adam optimizer with categorical cross-entropy loss.

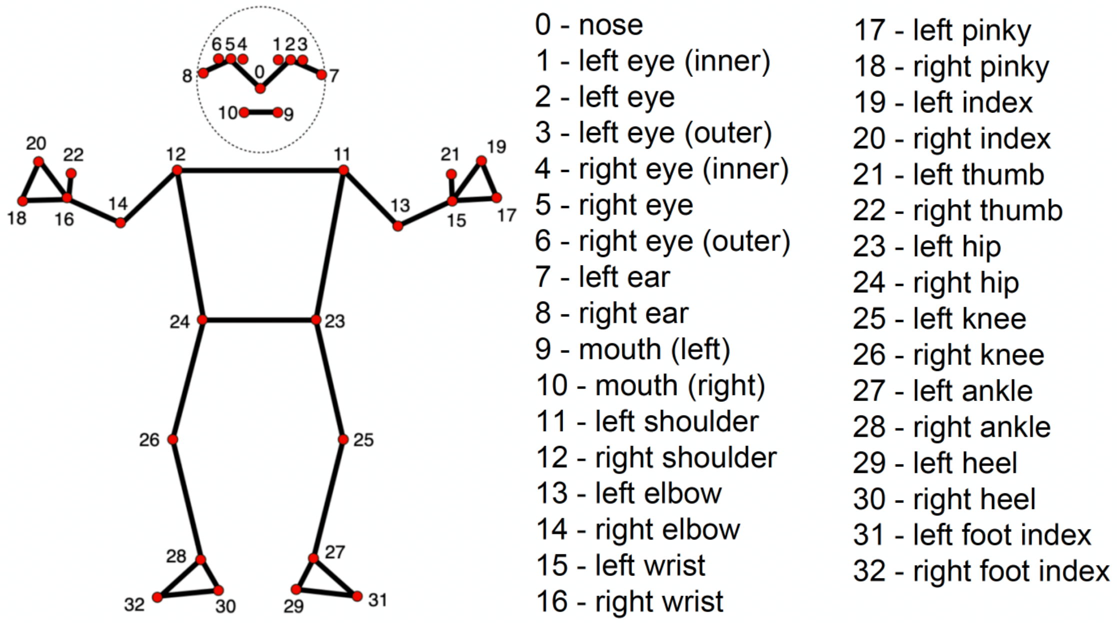

Image credit to Pose landmarker model from Google AI at https://ai.google.dev/edge/mediapipe/solutions/vision/pose_landmarker

Figure 2: Shows how to use MediaPipe to identify 33 landmarks on the human body.

The LSTM (Long Short Term Memory) layer is a kind of recurrent neural network (RNN) layer that can learn long term dependencies. The governing equations for an LSTM layer are given as follows:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

Eq.1

Input gate (i_t): Determines the amount of new information to be added to the cell state.

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right)$$

Eq.2

Forget gate (f_t): Decides what information from the cell state to discard.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$

Eq.3

Output gate (o_t): Controls the output from the LSTM cell.

$$C_t = f_t * C_{t-1} + i_t * \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

Eq.4

Cell state (C_t): Updated cell state incorporating the input gate and forget gate decisions.

$$h_t = o_t * \tanh(C_t)$$

Eq.5

Hidden state (h_t): The output of the LSTM cell.

The dense layer at the end of the network uses the softmax activation function to produce the final classification output:

$$y = \text{softmax}(W.h + b)$$

Eq.6

Where:

- σ is the sigmoid function
- tanh is the hyperbolic tangent function
- $*$ denotes element-wise multiplication
- W and b are weights and biases, respectively
- x_t is the input at time step t
- h_{t-1} is the hidden state from the previous time step
- C_{t-1} is the cell state from the previous time step

Figure 3. Testing the MediaPipe model for activity.

Each time through the training process, the dataset is used 151 times. We set up an early stopping callback to frequently check the categorical accuracy and stop training if it goes past the set number so that it doesn't learn too much from the training data.

For evaluation, only 25% of the total data is used from the test data set. A confusion matrix and an accuracy score show, when evaluated, how the model will perform on fresh, unseen data. Using MediaPipe, the pose land marker model recognizes 33 landmark points on the human body, as illustrated in Figure 2. Eleven of the sites are located on the tops of the peaks and, therefore, do not influence our study. That's why we decided not to incorporate them into our model.

The last part of the project shows how the trained model is put into action. Using the live video, the system guesses the body position and feeds the LSTM with the series of detected key points to help derive the action. You get to watch the model's prediction and also see a graph that represents how much it believes the prediction. You can see this idea depicted in Figure 3.

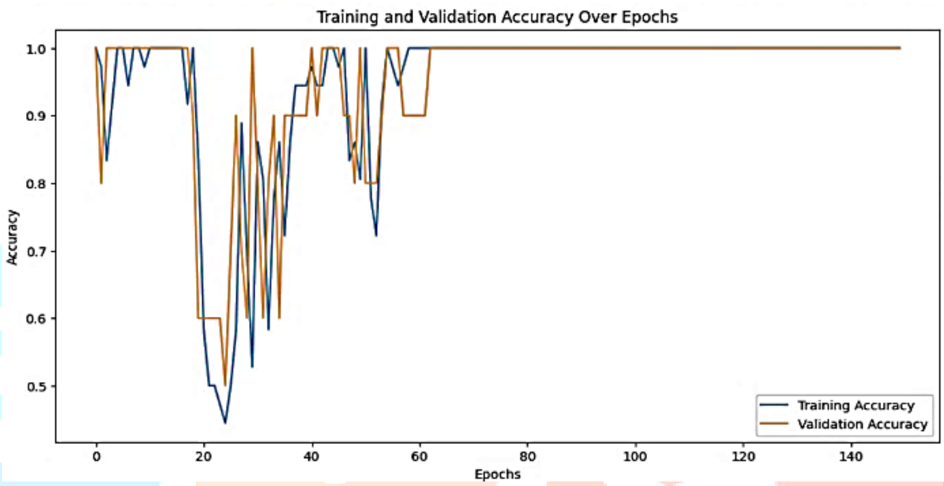

Figure 4. Training and Validation Accuracy Performance Visualization across Epochs.

IV. RESULT

1.Visualizations

Table 2 shows the primary performance data from the experiment by comparing against the benchmarks utilized by the model under investigation. The model achieved a 95.65 percent accuracy as it only made one error. The accuracy of the model improves at each step of the training process as seen in the accuracy curve in Figure 4. The confusion matrix in Figure 5 indicates the categories of true positives, true negatives, false positives, and false negatives to depict the model's interpretation of original data. Fig. 6 illustrates the connection between accuracy and strictness for the classifier.

2.Comparative Analysis

We use an architecture with three LSTM layers and three dense layers that is specially designed to recognize when different postures happen, helping our model classify postures effectively. It blends a model that works on sequences, computer vision for uninterrupted pose tracking, and an understandable web interface that shows feedback instantly. The addition of a framework that ranks postures and gives immediate guidance on lifting safely is a major new feature in our solution.
Our model is better than previous approaches such as YOLOv7 with LSTM and LSTM Pose Machines. This feature is especially valuable since other products do not manage so many changing postures, as this one does. Additionally, including risk analysis makes the model easier to apply, something that other existing models typically ignore. Because of our model's improvement over others, real-time response, and ability to calculate risk, it is the top choice for posture monitoring.

Our model offers several useful benefits. Using ergonomic 3D software greatly cuts down the risk of musculoskeletal disorders by correcting posture promptly and accurately. The use of improved temporal data in analysis helps create a safer workplace by leading to more precise management of posture. As a result, workplace productivity increases, and the number of MSD-related healthcare costs goes down. When operators know their lift is safe and work in a safe environment, the usage of real-time feedback and a simple interface becomes even more beneficial.

Table 2. Model Training and Evaluation Metrics

Metric	Value	Comments
Epochs	200	No. of training epochs
Categorical Accuracy	0.9567	Accuracy on training data
False Positives	1	No. of false positive predictions
True Negatives	1	No. of true negative predictions
True Positives	1	No. of true positive predictions
Final Accuracy Score	0.9489	Accuracy on evaluation dataset

The results from the comparison prove that our model can effectively monitor people's posture in the moment. The proposed LSTM model was the top performer, scoring 0.9375 in accuracy and 0.9565 for categorical accuracy. Categorical accuracy and accuracy score were 0.8348 and 0.8085, respectively, for komLSTM Pose Machines and 0.8616 and 0.8537 for the tennis shot identification model using YOLOv7 and LSTM.

Table 3: Comparative experiments with different models.

Model	Categorical Accuracy	Accuracy Score
Yoga LSTM Pose Machines	0.8348	0.8085
Tennis Shot Identification (YOLOv7 + LSTM)	0.8616	0.8537
Proposed LSTM-Based Model	0.9565	0.9375

LSTM-based LSTM has the highest accuracy and is best at generalizing since the results clearly show this in Table 3. One of the reasons for the model's great results is its use of many datasets, repetitive data-gathering steps, and a straightforward architecture. Although they do well in certain domains, existing models still face problems with high computational demands and limited ability to be used in many places. Our method works well for real-time posture monitoring, as it is both easy to use and efficient and accurate in a wide range of contexts.

These comparative studies demonstrate the value of our suggested approach in posture monitoring and ergonomic evaluations while also validating its efficacy. The model's significance as a useful and effective tool for improving workplace safety and reducing the risk of musculoskeletal disorders is highlighted by its capacity to outperform current solutions.

V. CONCLUSION

Overall, this research introduces a reliable and expandable system for boosting safety and results in gym workouts thanks to real-time posture feedback. By using precise measurements from MediaPipe and LSTM networks to understand the pace of the exercises, the system provides feedback through bodyweight moves such as squats, push-ups and pull-ups. As a result of this approach, proper technique during exercise can be reviewed and the possibility of musculoskeletal disorders from poor technique can be reduced.

Future efforts will include enriching the system by using a wider dataset covering many different users, for example by varying body types, fitness and exercise experience. Work will be done to ensure the data annotation is of high quality which will improve the model's generalization across various movement techniques and gyms.

To improve the usefulness of the system, work will be done to reduce latency, so feedback is always real-time during fast, continuous motions. For exercises such as push-ups and squats, fixing your posture matters right away. Letting the AI detect key areas more precisely and working with new neural networks—including attention-based recurrent networks or transformers—will help it work in any gym, adaptive to adjustments in lighting, angles and backgrounds.

Collaborative trials lasting months and years involving physiotherapists, personal trainers and health workers will be conducted to assess how the system continues to help users avoid injuries and be satisfied with their exercise. Using this research, we plan to make it easier for all gym users to benefit from intelligent, just-in-time posture adjustments in their workouts.

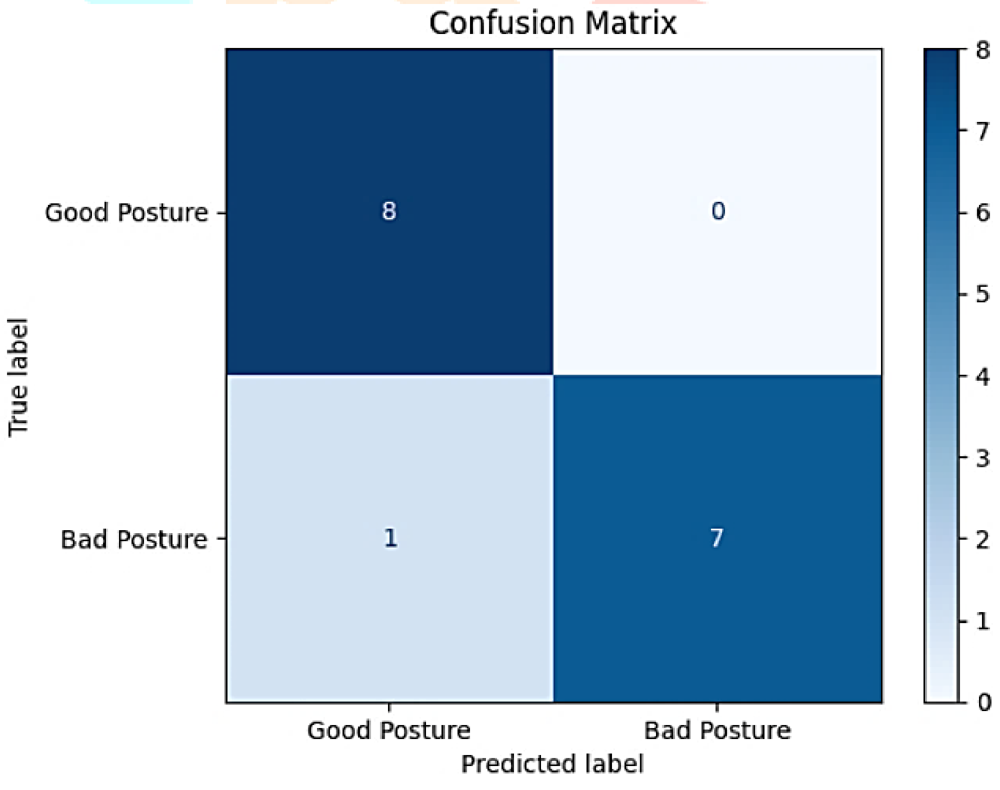

Figure 5. Training and Validation Accuracy Performance Visualization across Epochs.

Going forward, user experience will be of utmost importance. An upgraded, interactive user interface will show live feedback, make posture analysis and create dashboards that track how your form develops over the process. As a result, users can cheque their progress and follow actionable tips in every session.

In the future, using IoT devices such as smart mirrors, gym kiosks, or mobile phones could offer gym-goers a more complete view and let several users be tracked in group settings. Consequently, the apparatus will become suitable for both private exercise at home and for sale in commercial gyms.

VI. Acknowledgment

References

[1] T. Abraham, S.A. Binoo's, and K.R. Remes Babu. 2022.Virtual modelling and analysis of manual material handling activities among warehouse workers in the construction industry.Work 73, 3 (2022), 977–990.https://doi.org/10.3233/WOR-210742

[2] A. Singh, S. Agarwal, P. Nagrath, A. Saxena and N. Thakur, "Human Pose Estimation Using Convolutional Neural Networks," 2019 Amity International Conference on Artificial Intelligence (AICAI), Dubai, United Arab Emirates, 2019, pp. 946-952, doi: 10.1109/AICAI.2019.8701267.

[3] M.F. Antwi-Afari et al. 2022.Deep learning-based networks for automated recognition and classification of awkward working postures in construction using wearable insole sensor data.Automation in Construction 136 (2022).https://doi.org/10.1016/j.autcon.2022.104181

[4] V.C.H. Chan et al. 2022.The role of machine learning in the primary prevention of work-related musculoskeletal disorders: A scoping review.Applied Ergonomics 98 (2022).https://doi.org/10.1016/j.apergo.2021.103574

[5] X. Li, M. Zhang, J. Gu and Z. Zhang, "Fitness Action Counting Based on Media Pipe," 2022 15th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Beijing, China, 2022, pp. 1-7, doi: 10.1109/CISP- BMEI56279.2022.9980337.

[6] Chuen-Lung Chen, David B Kaber, and Patrick G Dempsey. 2000.A new approach to applying feedforward neural networks to the prediction of musculoskeletal disorder risk.Applied ergonomics 31, 3 (2000), 269–282.

[7] M. Ciccarelli et al. 2023.SPECTRE: a deep learning network for posture recognition in manufacturing.Journal of Intelligent Manufacturing 34, 8 (2023), 3469–3481.https://doi.org/10.1007/s10845-022-02014-y

[8] E. Fisher et al. 2023.Occupational Safety and Health Equity Impacts of Artificial Intelligence: A Scoping Review.International journal of environmental research and public health 20, 13 (2023).https://doi.org/10.3390/ijerph20136221

[9] M.M. Hussain et al. 2019.Digital Human Modeling in Ergonomic Risk Assessment of Working Postures using RULA. In Proceedings of the International Conference on Industrial Engineering and Operations Management. 2714–2725.https://research.ebsco.com/linkprocessor/plink?id=6db22889-9f94-3985-b9e6-bead12460591

[10] Irfan, I., Muthalib, M. A. (2022). Implementation of Human Pose Estimation Using Angle Calculation Logic on The Elder of The Hands as A Fitness Repetition. International Journal of Engineering, Science and Information Technology, 2(4), 101-110.

[11] Chen KY, Shin J, Hasan MAM, Liaw JJ, Yuichi O, Tomioka Y. Fitness Movement Types and Completeness Detection Using a Transfer-Learning-Based Deep Neural Network. Sensors (Basel).

[12] Pauzi, A.S.B. et al. (2021). Movement Estimation Using Mediapipe Blaze Pose. In: Badioze Zaman, H., et al. Advances in Visual Informatics. IVIC 2021. Lecture Notes in Computer Science (), vol 13051. Springer, Cham. https://doi.org/10.1007/978-3-030-90235- 3_49

[13] Ashish Ohri, Shashank Agrawal, Garima S. Chaudhary, "On-device Realtime Pose Estimation & Correction", in International Journal of Advances in Engineering and Management (IJAEM), Volume 3, Issue 7 July 2021.

[14] Long, C., Jo, E. \& Nam, Y. Development of a yoga posture coaching system using an interactive display based on transfer learning. J Supercomput 78, 5269–5284 (2022). https://doi.org/10.1007/s11227- 021-04076-w

[15] R. Gadhiya and N. Kalani, "Analysis of Deep Learning Based Pose Estimation Techniques for Locating Landmarks on Human Body Parts," 2021 International Conference on Circuits, Controls and Communications (CCUBE), Bangalore, India, 2021, pp. 1-4, doi: 10.1109/CCUBE53681.2021.9702726