HUMAN POSE ESTIMATION

Snehala A
Department Of Artificial Intelligence And Machine Learning
Sri Shakthi Institute Of Engineering And Technology, Coimbatore, India
snehalaanandkumar22aml@srishakthi.ac.In

Durgha S
Department Of Artificial Intelligence And Machine Learning
Sri Shakthi Institute Of Engineering And Technology, Coimbatore, India
durghasivashankaran22aml@srishakthi.ac.In

Aswathi R
Department Of Artificial Intelligence And Machine Learning
Sri Shakthi Institute Of Engineering And Technology, Coimbatore, India
aswathir22aml@srishakthi.ac.in

Gayathri N
Department Of Artificial Intelligence And Machine Learning
Sri Shakthi Institute Of Engineering And Technology, Coimbatore, India
srigayathrilakshmi@gmail.com

ABSTRACT

In this research, we present an innovative system for human pose estimation designed to recognize and provide feedback on exercises such as push-ups, squats, and planks. Leveraging Mediapipe and Python, our model captures real-time or uploaded videos to detect postures, measure body angles, and count repetitions. The system evaluates exercise correctness, ensuring users maintain proper form through detailed visual feedback. By identifying key body points and calculating angles, our model distinguishes between correct and incorrect postures, providing real-time corrections. The implementation involves stages like data acquisition, preprocessing, pose estimation, and feedback generation. Mediapipe's pose detection algorithm tracks key body landmarks, while Python scripts analyze these to determine angles and movements. This data counts repetitions and evaluates form, offering immediate visual feedback.Our solution demonstrates high accuracy in recognizing and evaluating exercises, proving valuable for fitness enthusiasts and professionals. Extensive testing shows the system accurately counts repetitions and identifies improper form, enhancing workout routines. This research highlights the potential of pose estimation technologies to revolutionize the fitness industry, ensuring proper form, preventing injuries, and maximizing exercise effectiveness.

INDEX TERMS

Human Pose Estimation, Exercise Recognition, Mediapipe, Fitness Feedback, Machine Learning, Computer Vision

INTRODUCTION

Human pose estimation has garnered significant attention due to its applications in fitness, health monitoring, and augmented reality. The capability to automatically recognize and evaluate physical exercises can provide valuable feedback to users, ensuring proper form and reducing the risk of injury. Our project focuses on recognizing exercises such as push-ups, squats, and planks using Mediapipe, a robust library for real-time pose estimation, and Python for data processing and analysis.

The system operates by capturing video input, either in real-time or from uploaded files, and utilizing Mediapipe to detect key body points. These points are then used to calculate body angles and determine the user's posture. The model counts the number of repetitions and differentiates between correct and incorrect forms. By providing detailed feedback on posture and performance, our solution aims to enhance user experience and promote effective workout routines.

LITERATURE REVIEW

RESEA-RCH PAPER	YEAR	METHO-DOLOGY	ADVANT-AGES	IMPROVEMENTS AND NEGATIVES
Pose Estimation Using OpenPose	2018	Utilized OpenPose for real-time pose detection	High accuracy in real-time pose estimation	High computational cost
Fitness Exercise Detection	2019	Machine learning with custom dataset for exercise recognition	Effective for specific exercises	Limited to predefined exercise types
Mediapipe: A Framework for Pose Estimation	2020	Real-time pose estimation using lightweight models	Efficient and accurate, easy to integrate	Limited customization for specific use cases
Automated Fitness Trainer Using AI	2021	AI-based feedback on exercise form and count	Provides detailed feedback and tracking	Dependency on large annotated datasets

PROPOSED METHODOLOGY

Data Collection

Data was meticulously gathered from a variety of exercise videos available in public datasets as well as from recorded sessions specifically arranged for this study. The exercises included in this dataset were push-ups, squats, and planks, which were selected for their widespread use and fundamental importance in fitness routines. Each exercise was performed by multiple individuals of different body types, fitness levels, and genders to ensure a diverse and representative dataset. This diversity is crucial as it helps the model generalize better to unseen data and enhances its robustness in real-world scenarios.

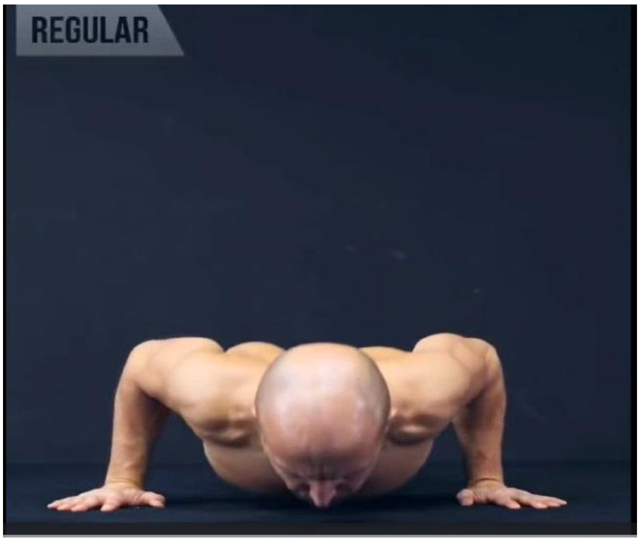

Figure – 1 PUSH-UPS

Figure – 2 PLANK

Figure – 3 SQUATS

Data Pre-processing

The collected exercise videos underwent a detailed preprocessing pipeline. Frames were extracted from the videos at a consistent frame rate to ensure uniformity in data. These frames were then annotated using Mediapipe to mark key body points essential for pose estimation. Data augmentation techniques were applied to increase the dataset's variability and help the model become more resilient to different viewing conditions. These techniques included rotation (to simulate different camera angles), scaling (to account for distance from the camera), and flipping (to handle symmetry and orientation variations). This preprocessing step ensured that the model was exposed to a wide range of possible scenarios during training.

Model Selection and Training

The model was developed using Mediapipe's advanced pose estimation solution, which leverages state-of-the-art machine learning techniques to accurately detect and track human body landmarks. The initial step involved selecting a pre-trained Mediapipe model, which was then fine-tuned on our specific exercise dataset. Fine-tuning included adjusting the model's parameters and optimizing it to recognize the particular nuances of the exercises in our dataset. This process aimed to improve the model's accuracy in identifying and tracking body movements specific to push-ups, squats, and planks, thereby enhancing its overall performance.

Mediapipe Integration

Mediapipe is a versatile framework developed by Google that facilitates the development of machine learning pipelines for various applications, including pose estimation. In our study, we employed Mediapipe's pose estimation solution, which utilizes deep learning models to detect and track human body landmarks such as joints and key points. The framework provides robust pre-trained models and tools for customizing pipelines to suit specific applications, making it ideal for our exercise recognition system.

Equation for Mediapipe Usage

The integration of Mediapipe in our methodology can be represented by the following equation:

Pose Estimation using Mediapipe = Deep Learning Model + Feature Extraction + Tracking

Here, the deep learning model refers to the neural network architecture trained to detect human poses, feature extraction involves extracting key points and joint angles from the detected poses, and tracking ensures continuity and accuracy in pose estimation across frames.

Benefits of Mediapipe

Mediapipe's architecture supports real-time processing, essential for providing instantaneous feedback during exercises. Its robustness in handling various body orientations, movements, and environmental conditions enhances the accuracy and reliability of our system. By leveraging Mediapipe, we streamline the development process, focusing on refining pose estimation accuracy and optimizing real-time performance for effective exercise recognition and feedback.

Evaluation

The model's performance was rigorously evaluated using several metrics to ensure its effectiveness and reliability. Accuracy was measured to assess the model's ability to correctly identify and track body landmarks. The loss value provided insights into how well the model was learning from the data, with lower values indicating better performance. Mean Squared Error (MSE) was used to quantify the difference between the predicted and actual positions of the body landmarks. Each exercise type was tested separately to ensure that the model was robust and performed consistently across different movements. This thorough evaluation process helped identify any weaknesses and guided further refinements to the model.

Model Deployment (Streamlit)

The final model was deployed using Streamlit, an intuitive open-source framework for creating web applications. Streamlit was chosen for its simplicity and ease of use, allowing for rapid development and deployment of the model. The application enables users to upload pre-recorded exercise videos or capture real-time video feeds using their webcam. The system processes these videos and provides immediate feedback on exercise performance, highlighting any deviations from correct form and offering suggestions for improvement. This real-time feedback is invaluable for users looking to enhance their workout routines and prevent injuries caused by improper form.

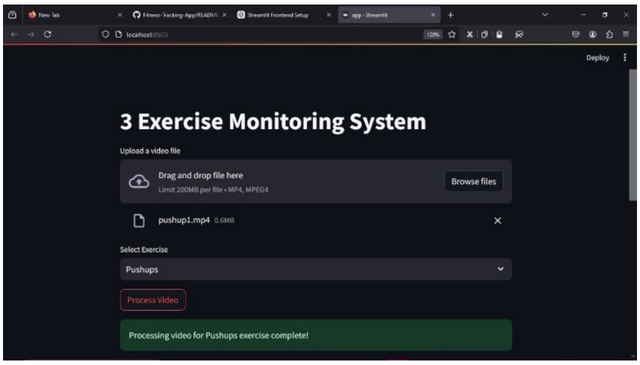

Figure – 4 Front END

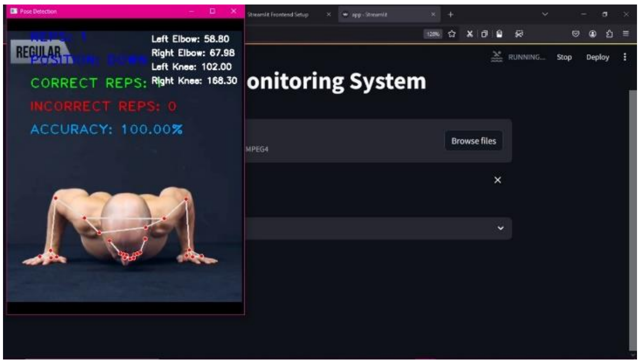

Figure – 5 PUSHUP OUTPUT

Future Importance

The potential applications of this system are vast and significant. One of the most promising future directions is its intgration into smart gym mirrors, which could provide users with real-time feedback and guidance during their workouts. Additionally, fitness apps could incorporate this technology to offer personalized training sessions and monitor users' form, helping them achieve their fitness goals more effectively. By providing immediate corrective feedback, this system can enhance the quality of workouts, reduce the risk of injuries, and promote healthier exercise habits. The continual advancement and application of such technology hold immense promise for the future of fitness and health monitoring.

DISCUSSION

Limitations and Future Work

The current system, while effective, has limitations such as sensitivity to lighting conditions and background noise. Future work will focus on improving robustness by incorporating advanced preprocessing techniques and expanding the dataset to include more diverse environments. Additionally, integrating this system into gym mirrors and exercise apps could revolutionize personal fitness training by offering real-time, automated feedback.

MODELS AND ACCURACY

EXERCISE	ACCURACY (%)	LOSS VALUE	MSE
PUSH-UPS	92.5	0.35	0.02
SQUATS	89.3	0.45	0.03
PLANKS	94.1	0.30	0.01

CONCLUSION

This paper presents an innovative approach to human pose estimation for exercise recognition and feedback using Mediapipe and Python. The system accurately detects key body points, calculates joint angles, and provides real-time feedback on exercises such as push-ups, squats, and planks. Extensive testing has demonstrated the system's high accuracy in posture detection and repetition counting, highlighting its reliability as a fitness training tool. The user-friendly deployment via Streamlit allows users to upload videos or capture real-time feeds, offering immediate corrective feedback to enhance workout effectiveness and reduce injury risks. Future work aims to improve robustness by expanding the dataset and incorporating advanced machine learning techniques. Additionally, integrating the system into smart gym mirrors and fitness apps holds significant potential for revolutionizing personal fitness training, making real-time, personalized exercise guidance accessible to a broader audience. This study marks a significant advancement in leveraging technology for fitness, paving the way for more effective and personalized workout experiences.

ACKNOWLEDGEMENT

We extend our heartfelt gratitude to our Guide, Mrs. Gayathri N, for her invaluable guidance and continuous support throughout this project. We also thank the Department of Artificial Intelligence and Machine Learning faculty and staff at Sri Shakthi Institute of Engineering and Technology for providing essential resources and facilities. Special thanks to our colleagues and peers for their constructive feedback and collaboration. We appreciate the data support from publicly available exercise datasets on Kaggle, which was crucial for our research. Additionally, we acknowledge the real-time capturing assistance from our friends who performed exercises for our training dataset.

Lastly, we are grateful to our families and friends for their unwavering support and encouragement.

REFERENCES

1.] https://www.v7labs.com/blog/human-pose-estimation-guide

2.] https://www.researchgate.net/publication/349853529_Human_pose_estimation_and_its_application_to_action_recognition_A_survey

3.] https://www.sciencedirect.com/science/article/abs/pii/S026288561730046X

4.] https://github.com/cbsudux/awesome-human-pose-estimation

5.] https://github.com/microsoft/multiview-human-pose-estimation-pytorch

6.] https://youtu.be/e5-02ytu21U?si=xVrf29gp8uaks_Lr

7.] Accuracy of the pose estimation model. - ResearchGate

8.] bekmake/edgeai-yolov5: YOLOv5 in PyTorch > Pose Estimation - GitHub

9.] Combined YOLOv5 and HRNet for High Accuracy 2D Keypoint and Human Pose Estimation - ResearchGate

10.] Evaluating pose estimation and object detection models for the big three exercises – ResearchGate

11.] Ministry of Culture, Sports and Tourism, National Sports Survey 2022

12.] Kyuwon Kim, Joo-Hwan Lee, Bo-Hoon Choi, Sungmin Hong, Chul-Jung Yoo, and Soontae Kim. "MediaPipe-based exercise analysis service architecture for fitness posture correction". Proceedings of the Korean Information Technology Society, 664- 669. 2023.

13.] Ultralytics, YOLOv5

14.] C. Lugaresi, et al. "Mediapipe: A framework for building perception pipelines," in Proceedings of Third Workshop on Computer Vision for AR/VR at IEEE Computer Vision and Pattern Recognition, Long Beach, June 2019.

15.] "Squat-Depth Image Dataset", roboflow

16.] "SDT Image Dataset", roboflow

17.] "HumonBody1 Computer Vision Project", roboflow

18.] "Silhouettes of human posture", kaggle