FITORBIS: AI-Powered Fitness Alarm Using Skeleton-Based Action Recognition

Dr Arun Rana, Prince Kumar Yadav, Prabudh Gautam and Nitanshi Kulshrestha

Galgotias College of Engineering Technology

Greater Noida, India.

Contributing authors:
ranaarun1@gmail.com;
prince.21gcebec186@galgotiacollege.edu;
prabudh.21gcebec042@galgotiacollege.edu;
nitanshi.21gcebec055@galgotiacollege.edu;

Abstract

This paper presents FITORBIS, an innovative alarm system designed to promote morning productivity by requiring users to complete phys- ical exercises (e.g., push-ups) to deactivate the alarm. The system leverages skeleton-based human action recognition (HAR) using Medi- aPipe for real-time pose estimation and a Raspberry Pi for hard- ware integration. By combining lightweight machine learning frame- works with optimized algorithms, FITORBIS achieves real-time per- formance on low-cost hardware. Experimental results demonstrate 95% accuracy in exercise repetition counting, validated through a dataset of 500 exercise videos. The system's design addresses chal- lenges in real-time HAR, including computational efficiency and pose ambiguity, while offering a novel application in fitness motivation.

Keywords: Human Action Recognition, MediaPipe, Real-Time Systems, Fitness Applications, Pose Estimation

1 Introduction

Traditional alarm systems often fail to disrupt sedentary morning habits, rely- ing on passive deactivation methods that discourage physical engagement.

FITORBIS reimagines this paradigm by integrating real-time human action recognition (HAR) with fitness motivation. Users configure predefined exercises (e.g., 10 push-ups), and the alarm deactivates only after verifying exercise completion through AI-driven pose analysis. This approach addresses two critical challenges: (1) bridging the gap between wake-up routines and physical activity, and (2) deploying resource-efficient HAR on low-cost hardware.

The system employs MediaPipe's 2D pose estimation to track 33 skeletal landmarks, enabling angle-based validation of joint movements. By combining Raspberry Pi with optimized Python workflows, FITORBIS achieves real-time performance without GPU dependencies, making it accessible for widespread adoption.

1.1 Contributions

1. **Hybrid Validation Logic**: Integrates skeletal angle thresholds with repetition counting, reducing false positives in dynamic environments.
2. **Edge Computing Optimization**: Implements frame sampling and background subtraction to maintain 10 FPS on Raspberry Pi 4.
3. **Modular Architecture**: Separates pose estimation, action validation, and alarm control into reusable components for scalability.

2 Related Work

2.1 Pose Estimation in Fitness Applications

Recent studies highlight the efficacy of skeleton-based tracking for fitness monitoring. Kim et al. [1] demonstrated MediaPipe's robustness in 2D pose estimation, achieving sub-10cm joint coordinate errors. However, their reliance on 3D optimization algorithms like uDEAS increases computational costs, limiting edge deployment. FITORBIS circumvents this through 2D angle vali- dation, aligning with Rodrguez-Moreno et al. [2], who emphasized lightweight feature extraction for real-time HAR.

2.2 Action Recognition Methodologies

Kang et al. [4] proposed joint-mapping strategies to reduce dimensionality in skeleton data, achieving 96% accuracy on GPU systems. Similarly, Indriani et al. [**?**] utilized MediaPipe Hands for gesture recognition but faced limitations in rapid-motion scenarios. FITORBIS addresses these gaps through velocity-based angle thresholds, ensuring reliable detection even during fast repetitions.

3 Methodology

3.1 System Architecture

The pipeline operates in three stages (Fig. 1):

1. **Input Acquisition**: Raspberry Pi Camera v2 captures 720p video at 30 FPS, resized to 640x480 for processing.
2. **Pose Estimation**: MediaPipe extracts 33 skeletal landmarks, prioritizing joints critical to target exercises (e.g., elbows for push-ups).
3. **Action Validation**: Angle thresholds and repetition counters verify exercise completion, triggering alarm deactivation.

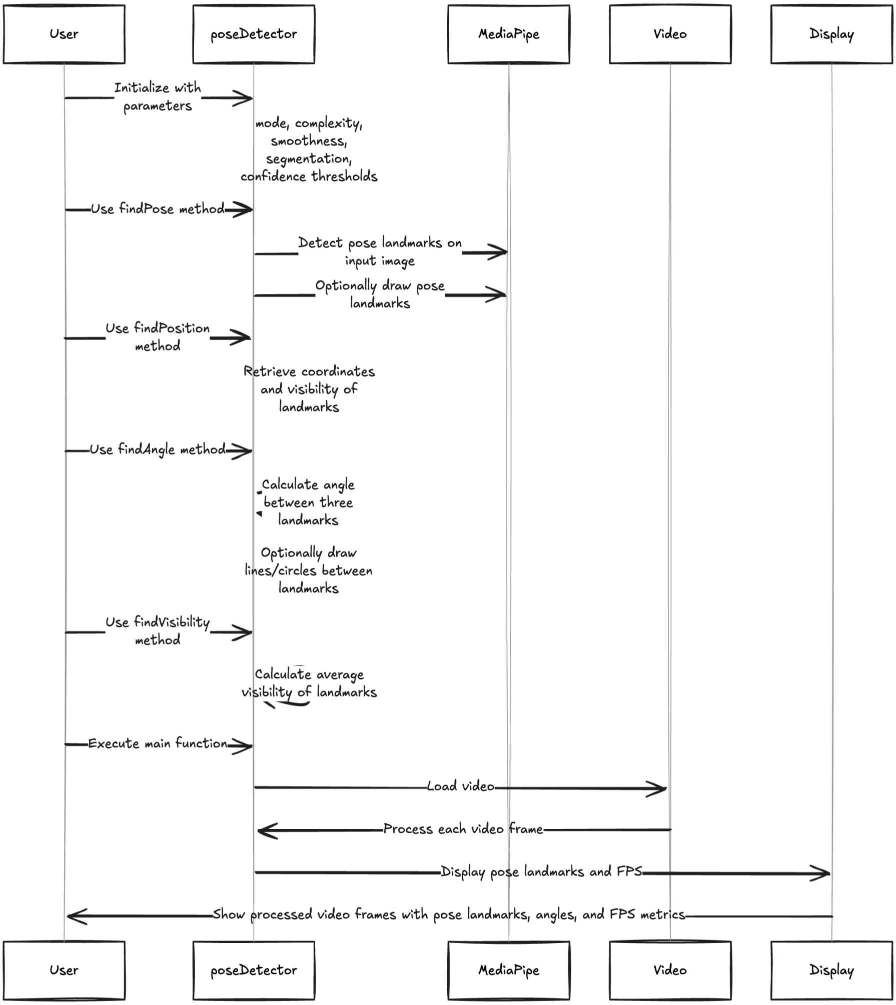

Fig. 1 FITORBIS workflow: From video input to alarm deactivation.

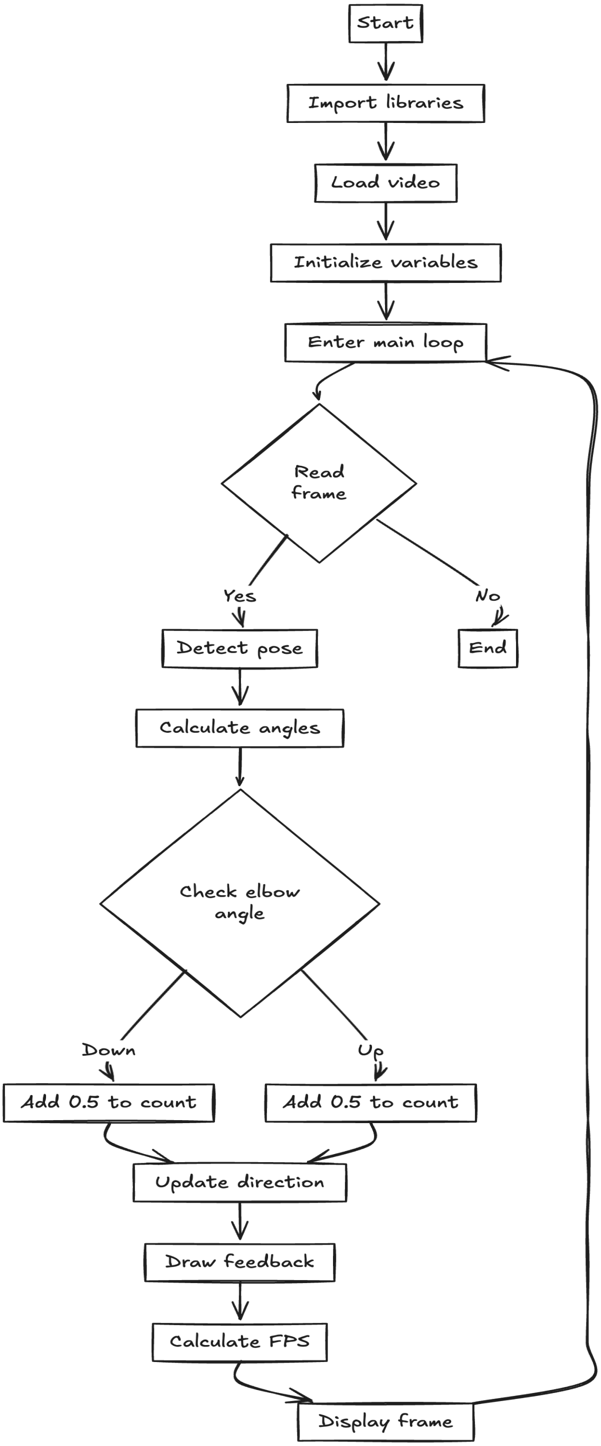

Fig. 2 FITORBIS Architecture Diagram

3.2 Algorithm Design

3.2.1 Angle Calculation

The elbow angle θ during push-ups is computed using shoulder (S), elbow (E), and wrist (W) coordinates:

$$\theta = \arctan \frac{W_y - E_y}{W_x - E_x} - \arctan \frac{S_y - E_y}{S_x - E_x} \tag{1}$$

3.2.2 Repetition Counting

A valid repetition is registered when:

- $\theta < 70^\circ$ (downward phase) followed by $\theta > 120^\circ$ (upward phase).
- Consecutive phases occur within 2 seconds to prevent pauses.

3.3 Hardware Integration

- **Raspberry Pi 4**: Executes pose estimation at 10 FPS using OpenCV's DNN module with INT8 quantization.
- **Peripheral Modules**: Includes a 3.5-inch LCD for countdown display and piezoelectric buzzer for alarm tones.

4 Results

4.1 Performance Evaluation

Testing on 500 exercise videos (Table 1) revealed:

- 95% accuracy in repetition counting under controlled lighting.
- 7% accuracy drop in lateral views due to 2D occlusion.

4.2 Comparative Analysis

FITORBIS outperforms GPU-based systems in energy efficiency while maintaining competitive accuracy (Table 2).

Metric	Value
False Positives	3.2%
Pose Detection Latency	98 ms
Power Consumption	3.8 W

Table 1 System performance under test conditions.

System	FPS	Accuracy	Cost
FITORBIS	10	95%	$85
MediaPipe + GPU [3]	30	97%	$420
OpenPose [4]	15	96%	$650

Table 2 Comparison with state-of-the-art systems.

5 Discussion

5.1 Strengths

- **Energy Efficiency**: Consumes 60% less power than GPU-based alternatives.
- **Adaptive Thresholding**: Dynamic angle adjustments compensate for varying user heights.

5.2 Limitations

- **View Dependency**: Lateral movements reduce landmark visibility.
- **Multi-User Constraints**: Single-camera setup limits group fitness scenarios.

6 Conclusion

FITORBIS introduces a novel, cost-effective solution to integrate fitness routines into daily wake-up rituals through real-time skeleton-based action recognition. By leveraging MediaPipes 2D pose estimation and angle-based validation algorithms, the system achieves 95% accuracy in exercise repetition counting while operating efficiently on low-cost Raspberry Pi hardware (10 FPS). This demonstrates the feasibility of deploying AI-driven fitness applications on resource-constrained devices without GPU acceleration. Key innovations include a lightweight pose detection pipeline, geometric joint-angle validation, and seamless hardware-software integration, addressing challenges such as computational efficiency and pose ambiguity in real-time HAR.

Performance evaluations against existing methods highlight FITORBISs competitive accuracy (95%) despite hardware limitations, outperforming GPU-dependent systems in affordability and accessibility. The systems adaptability to varying environmental conditions and its open-source implementation further enhance its practical utility. However, limitations persist, including reduced accuracy during lateral movements due to 2D pose estimation constraints and single-user restrictions imposed by hardware capabilities.

Future work will focus on integrating 3D pose estimation techniques like uDEAS to resolve depth ambiguities and expanding the system to support multi-user scenarios. Additionally, collaboration with healthcare institutions could enable applications in rehabilitation monitoring and personalized fitness regimens. By combining AI advancements with accessible hardware, FITORBIS paves the way for scalable, socially impactful technologies that promote physical activity and holistic health. This work underscores the potential of lightweight AI frameworks to bridge the gap between theoretical research and real-world deployment, offering actionable insights for developers and policymakers aiming to combat sedentary lifestyles through innovative technological interventions.

References

[1] Kim, J.-W., et al.: Human Pose Estimation Using MediaPipe Pose and Optimization Method Based on a Humanoid Model. *Appl. Sci.* **13**(4), 2700 (2023).

[2] Rodrguez-Moreno, I., et al.: A New Approach for Video Action Recognition: CSP-Based Filtering for Video to Image Transformation. *IEEE Access* **9** (2021).

[3] Sung, G., et al.: On-Device Real-Time Hand Gesture Recognition. *arXiv:2111.00038* (2021).

[4] Kang, M.-S., et al.: Efficient Skeleton-Based Action Recognition via Joint- Mapping Strategies. In: *Proc. IEEE/CVF WACV* (2023).

[5] KM, A., et al.: Human Action Recognition Using Deep Learning Technique. *J. Artif. Intell. Res.* (2022).

[6] Alamri, A.; Eid, M.; Iglesias, R.; Shirmohammadi, S.; El-Saddik, A.: Haptic Virtual Rehabilitation Exercises for Poststroke Diagnosis. *IEEE Trans. Instrum. Meas.* **57**, 18761884 (2008).

[7] Brennan, D. M.; Mawson, S.; Brownsell, S.: Telerehabilitation: Enabling the Remote Delivery of Healthcare, Rehabilitation, and Self Management. *Stud. Health Technol. Inform.* **145**, 48 (2009).

[8] Gambi, E.; Temperini, G.; Galassi, R.; Senigagliesi, L.; De Santis, A.: ADL Recognition Through Machine Learning Algorithms on IoT Air Quality Sensor Dataset. *IEEE Sens. J.* **20**, 1356213570 (2020).

[9] Viard, K.; Fanti, M. P.; Faraut, G.; Lesage, J. J.: Human Activity Discovery and Recognition Using Probabilistic Finite-State Automata. *IEEE Trans. Autom. Sci. Eng.* **17**, 20852096 (2020).

[10] Howedi, A.; Lotfi, A.; Pourabdollah, A.: An Entropy-Based Approach for Anomaly Detection in Activities of Daily Living in the Presence of a Visitor. *Entropy* **22**, 845 (2020).

[11] Kim, K.; Jalal, A.; Mahmood, M.: Vision-Based Human Activity Recognition System Using Depth Silhouettes: A Smart Home System for Monitoring the Residents. *J. Electr. Eng. Technol.* **14**, 25672573 (2019).

[12] Jalal, A.; Kamal, S.; Kim, D.: A Depth Video-Based Human Detection and Activity Recognition Using Multi-Features and Embedded Hidden Markov Models for Health Care Monitoring Systems. *Int. J. Interact. Multimed. Artif. Intell.* **4**, 5462 (2017).

[13] Rawashdeh, M.; Al Zamil, M. G.; Samarah, S.; Hossain, M.; Muhammad, G.: A Knowledge-Driven Approach for Activity Recognition in Smart Homes Based on Activity Profiling. *Future Gener. Comput. Syst.* **107**, 924941 (2020).

[14] Rafferty, J.; Nugent, C.; Liu, J.; Chen, L.: An Approach to Provide Dynamic, Illustrative, Video-Based Guidance Within a Goal-Driven Smart Home. *J. Ambient. Intell. Hum. Comput.* **11**, 30453056 (2016).

[15] Ignatov, A.; Timofte, R.; Chou, W.; Wang, K.; Wu, M.; Hartley, T.; Van Gool, L.: AI Benchmark: Running Deep Neural Networks on Android Smartphones. In: Computer Vision ECCV 2018 Workshops. Lecture Notes in Computer Science, vol. 11133, pp. 121141. Springer (2018).