Pose-Triggered Alarm System using Computer Vision and Machine Learning

of

Project Report submitted in the partial fulfilment of

Bachelor of Technology Artificial Intelligence & Data Science

by

Neeraj Pawar (A004)
Bindi Gondalia (A005)
Palak Singh (A033)

Under the supervision of

Shailendra Aote

(Associate Professor, STME NMIMS Navi Mumbai)

SVKM's NMIMS University

(Deemed-to-be University)

SCHOOL OF TECHNOLOGY MANAGEMENT & ENGINEERING (STME)
Kharghar, Navi Mumbai - 410210

(2021-25)

CERTIFICATE

This is to certify that the project entitled **"Pose-Triggered Alarm System using Computer Vision and Machine Learning"**, has been done by **Mr. Neeraj Pawar, Ms. Bindi Gondalia** and **Ms. Palak Singh** under my guidance and supervision & has been submitted in partial fulfilment of the degree of **Bachelor of Technology** in **Artificial Intelligence and Data Science** of STME, SVKM's NMIMS (Deemed-to-be University), Navi Mumbai, India.

Project mentor (name and Signature) Examiner (name and Signature) (Intenal Guide)

Date: 09-11-2024

Place: Navi Mumbai

(HoD) (name and Signature)

ACKNOWLEDGEMENT

We would like to sincerely thank to the mentor, Prof. Shailendra Aote, for indispensable direction, motivation, and assistance in the progression of our project. His knowledge and skills helped remarkably in elucidating complex concerns and improvement of our methodology. His support and commitment motivated us to stretch our limits and look for excellence.

We would like to express our heartfelt gratitude to NMIMS Navi Mumbai for creating an environment that facilitates learning, discovery, and personal growth. The resources, facilities, and congenial structure of this place have significantly helped us to successfully complete our project. NMIMS is not just a supporting environment, but also it provides opportunities and motivation to conclude this project, and thus, it is a great opportunity to work on this project in such an inspiring atmosphere.

We are grateful to the mentor and the institution for their consistent and steady support. This project epitomizes the commitment the institution was willing to make to facilitate student learning and achievement.

Student Names

NAME	**ROLL NO.**	**SAP ID**
Neeraj Pawar	A004	70562100018
Bindi Gondalia	A005	70562100022
Palak Singh	A033	70562100133

ABSTRACT

Inspired by yoga-inspired movement to energize the mind and body, this idea of developing an innovative posetriggered alarm system arises. This research introduces a new application which advances computer vision and machine learning for an interactive wake-up solution. As per the above, the system is designed to prompt users into appropriate specific physical activity routines, mostly yoga, which must be done exactly to turn off the alarm. Advanced pose estimation algorithms and dependency on MoveNet model are utilized for detecting and analyzing users' posture, against which verification of correct alignment is determined through pre-defined pose metrics. The application guides the user through poses- Warrior, Plank, and Tree with real-time accuracy and count completion. This integrates the MediaPipe library with core components of key pose-tracking, which allows for better recognition of keypoints and accurate skeletonization, thereby smoothening the interface between the physical activity and the corresponding application response. Preliminary tests on a diverse sample set show a high reliability of pose detection with accuracy rates of over 98% in various lighting and environmental conditions. This system, other than its potential as an alarm, motivates physical wellness through exercising regularly in ways that foster more user involvement and wellness in general. The integration of the pose-triggered alarm system utilizes a friendly user interface for guiding the users in practicing yoga poses. On the activation of the alarm, it forces users to execute movements with instant feedback concerning the correctness of poses executed. The pilot study had 50 participants, and results showed that 85% of the participants felt the system motivated better morning routines and higher physical activity. The average deactivation time for the alarm was 12 seconds, and retention rate remained above 90%. This work shows that the PoseStop Alarm wakes up the user without interrupting the healthy nature of daily routines by actively involving the user in movement within daily routines. Future avenues may lie in targeted interventions and adaptive learning that can effectively maximize engagement and well- being.

Table of Contents

List of the figures

List of tables

Abbreviations

Abbreviation	**Name of the Table**
COCO	Common Objects in Context
MPII	Max Planck Institute for Informatics
LDY	Low-Down Yoga dataset
XgBoost	Extreme Gradient Boosting
OpenCV	Open Source Computer Vision Library

Chapter 1

INTRODUCTION

Sleep is an integral component of human health and influential in the body for maintaining physical health, cognitive functions, and emotional balance. Though it is essential for all, many people fail to sleep well, hence restricting their ability to wake up and start their day. [10] Traditional alarm devices do not motivate their users enough, and hence, they develop dependence on the snooze button and get out of bed unwillingly. This challenge offers a chance for creative solutions that enhance not only alertness but also general well-being.

Embracing yoga as a comprehensive practice has experienced great rises during the past few centuries, for the purposes it serves to refresh intellects as well as physical fitness. Yoga encompasses heightened vigilance and bodily movements making it an activity ideal to be adapted for all practices. With a hybrid mix of yoga philosophy with recent advanced technology, one may then conceive the possibility of a developing engaging system that stimulates all of its users to increase engagement from the start of day. [5]

This paper introduces an alarm system that is activated by pose, using advanced computer vision and machine learning techniques to create an interactive wake-up experience. It uses the MoveNet framework for instant pose estimation, thus allowing the system to accurately identify and assess user positions. In this case, the system requires users to perform certain yoga poses such as Warrior, Plank, and Tree to turn off the alarm. This integration helps to readily switch from sleep to a physical activity, thus providing alertness and overall wellbeing.

The central objective of this research is to evaluate the effectiveness of the pose-activated alarm system in engaging users while promoting a healthy lifestyle. Using the MediaPipe library for accurate keypoint detection and skeletal mapping, the system ensures reliable pose tracking, providing instant feedback to users about their actions. [12] Apart from the role it plays as an alarm, this system also holds its potential in its ability to influence one's daily habit positively. In that it encourages people to practice yoga in the morning, it is more of a tool for enhanced physical and mental clarity. Essentially, the new strategy transforms the waking experience by fusing technology and wellness into a more cohesive lifestyle.

1.1 Background of the project

The PoseStop Alarm System integrates movement-based wake-up routines with machine learning and computer vision technologies. Drawing from the ancient practice of yoga, which combines mental and physical activation, the system provides a solution that motivates users to engage physically at the start of their day. The increasing accessibility of pose-estimation tools, such as MediaPipe and MoveNet, allows for accurate, real-time tracking of body positions. This technology enables the system to detect specific yoga poses and ensure user alignment, making it suitable for both fitness and daily wellness applications.

1.2 Motivation and scope of the report

This report is motivated by the need for more engaging and health-promoting alarm systems that prevent users from simply snoozing and encourage movement. Existing alarms lack the interactivity necessary to stimulate physical and mental activity effectively. The scope of this research is to explore the design and efficacy of an interactive alarm system that relies on yoga poses for deactivation, thus merging the goals of wakefulness and wellness. The report further investigates technical aspects, including pose estimation and real-time feedback, which are essential to creating an effective and user-friendly experience.

1.3 Problem statement

Traditional alarms fail to sufficiently motivate users to wake up and remain active. Many users rely on the snooze function, contributing to grogginess and poor morning routines. This project addresses the challenge by developing a pose-triggered alarm system that not only wakes users but also encourages them to begin their day with physical activity. The system’s primary objective is to evaluate the feasibility of using pose estimation to engage users in a simple, effective morning routine, which has the added benefit of promoting wellness.

1.4 Salient contribution

The PoseStop Alarm System contributes to the fields of computer vision and wellness technology by introducing a novel application of pose estimation for morning routines. Key contributions include:

1. Implementation of the MediaPipe Pose model for real-time pose detection.
2. Use of the MoveNet model for high-accuracy skeletal tracking.
3. Integration of Kalman filtering for smooth, stable pose estimation.
4. A user feedback system that guides users through pose completion. This system provides an innovative approach to improving user alertness and fostering physical activity, which can be extended to various health and fitness applications.

1.5 Organization of report

This report is organized as follows:

Chapter 1	Introduces the project background, motivation, problem statement, contributions, and report structure.
Chapter 2	Reviews relevant literature on pose estimation, yoga-based applications, and advancements in real-time computer vision.
Chapter 3	Details the methodology, including system design, pose detection techniques, and user feedback mechanisms.
Chapter 4	Discusses system limitations, including challenges with real-time video processing and multi-user tracking.
Chapter 5	Presents the results of the system's testing, including performance metrics for pose detection and user feedback.
Chapter 6	Concludes the report with insights on the system's effectiveness and potential future improvements in wellness technology.

Chapter 2

LITERATURE SURVEY

2.1 Introduction to overall topic

Human pose estimation has made significant strides over the past decade, primarily driven by advancements in deep learning and computer vision technologies. This field enables computers to detect, analyze, and interpret human body positions in real-time, allowing the development of interactive systems with applications spanning health, fitness, and rehabilitation. High-resolution representations, as proposed by Sun et al. (2019) [1], are foundational for accurate pose detection, allowing the identification of key body points essential in applications that require precise spatial localization, such as yoga pose monitoring. Pose estimation models now leverage networks that retain high-resolution representations throughout the processing pipeline, ensuring effective tracking across various challenging datasets, including COCO and MPII [1].

Yoga pose monitoring, particularly, has benefited from these advancements. Specialized deep learning models, like the Thunder variant of MoveNet combined with MediaPipe, achieve remarkable accuracy in recognizing complex yoga postures. A study on yoga pose detection demonstrated an accuracy of 99.50% in identifying poses such as Downdog, Goddess, Plank, Tree, and Warrior using the LDY dataset [2].

Pose comparison systems play a pivotal role in providing corrective feedback based on angular deviations at joint alignments between a user’s pose and a reference expert pose [3]. This approach ensures a safer, more accurate practice environment, particularly beneficial for a yoga pose-triggered alarm system that requires consistent accuracy and user-friendly feedback.

Google’s MediaPipe framework has emerged as a powerful tool for cross-platform, real-time pose estimation, particularly useful for applications requiring low latency and mobile compatibility [4]. With models like Blazepose, MediaPipe has made it feasible to deploy pose estimation in low-power devices with machine learning classifiers, including XgBoost, to achieve both accuracy and efficiency [5].

Tele-rehabilitation is another field that benefits from real-time human pose tracking, providing insights that are adaptable to yoga pose monitoring. Vision-based 3D data algorithms and stereo cameras allow the creation of 3D avatars for pose similarity scoring, offering real-time feedback between patients and therapists in remote settings [6].

In dynamic and continuous applications, the Kalman Filter is frequently used for its ability to estimate states recursively while minimizing error [7]. This filter has applications in navigation and linear filtering and is well-suited for yoga pose tracking due to its continuous state estimation capabilities. By reducing noise in pose estimation, Kalman filtering can improve accuracy in real-time yoga tracking systems, making it an invaluable tool for systems like the yoga pose-triggered alarm.

MediaPipe's support for iterative system performance improvement makes it an optimal choice for real-time applications requiring high precision and user accessibility across multiple device types [8].

In sum, the integration of human pose estimation technologies with computer vision and machine learning has paved the way for advanced applications in health and fitness. The yoga pose-triggered alarm system builds on these foundational technologies to deliver a practical solution that encourages physical engagement and wellness by using accurate, feedback-oriented pose estimation frameworks. The application of computer vision and deep learning to yoga pose estimation is part of a broader trend in developing personalized health and wellness solutions. In addition to pose detection, systems that offer real-time feedback based on body alignment have gained popularity. For instance, techniques such as angle-based feedback mechanisms allow users to self-correct their posture by comparing their joint alignments to reference models, as shown in studies that leverage landmark-based tracking to guide physical activities [9].

Research by Cao et al. (2021) highlights the significance of accurate real-time body pose tracking, especially in applications like fitness and physiotherapy, where precise alignment feedback supports a more effective self-guided experience [10]. This capability is further amplified by incorporating high-resolution pose detection models and machine learning frameworks like XgBoost and neural networks, which classify and correct poses with high accuracy [11].

One of the challenges in building such systems is achieving both high accuracy and low latency, especially on mobile or low-power devices. Recent advancements in lightweight models and optimized algorithms, such as those used in MediaPipe, address this challenge by enabling real-time processing with minimal computational load [12]. By focusing on resource management and cross-platform support, MediaPipe's design makes it possible to integrate pose estimation systems into mobile and wearable devices, broadening the accessibility of applications in health monitoring and interactive wellness systems.

Another advancement that strengthens the utility of yoga pose-triggered systems is multi-person detection, which allows a single system to differentiate between multiple individuals in the frame. Techniques like DeepCut, developed by Insafutdinov et al. (2016), segment poses in multi-person scenarios, offering the capacity for group applications, such as fitness classes or team-based exercise monitoring [13].

Pose estimation accuracy can be further enhanced with human feedback and supervised learning techniques, which refine models through continuous interaction and corrections. A feedback loop system not only supports alignment accuracy but also contributes to user engagement, as users receive personalized advice that adapts to their progress [14].

Integration with real-time monitoring tools like webcams or smartphone cameras allows pose-tracking systems to gather data unobtrusively, providing a smoother user experience. For example, a study by Park et al. (2020) demonstrated how depth-sensing cameras and machine learning classifiers could support accurate pose identification without requiring wearable sensors, making the system more convenient for users [15]. This hands-free approach is particularly advantageous for yoga, where practitioners often seek a more immersive experience without the distraction of wearable technology.

Furthermore, implementing a yoga pose-triggered alarm system has potential applications beyond personal fitness. Systems designed for tele-rehabilitation, for example, often employ pose estimation to compare patient movements with therapist-set standards, ensuring that exercises are performed correctly [16]. By adapting these techniques, a yoga alarm system can function not only as a personal motivator but also as a tool for structured physical wellness programs.

Moreover, pose estimation frameworks that combine computer vision with reinforcement learning, as explored in recent work by Zhang et al. (2022), have opened doors for adaptive feedback mechanisms that learn user preferences over time [17].

In the context of improving overall wellness, real-time feedback and pose correction present meaningful opportunities for injury prevention. Research by Akash et al. (2021) points out that incorrect postures can lead to cumulative strain over time, particularly in practices like yoga, where each posture has specific alignment criteria [18]. Providing immediate corrective feedback helps users avoid common alignment mistakes, creating a safer, more sustainable practice environment.

Finally, advancements in low-latency, high-accuracy pose estimation frameworks, combined with cross-platform adaptability, position yoga pose-triggered systems as promising tools for promoting wellness and engagement in an increasingly health-conscious society. The ability to offer real-time, accessible, and reliable feedback aligns with the growing demand for personalized fitness solutions that can be integrated into daily routines. By bridging the gap between personal fitness and technological convenience, these systems represent a forward-thinking approach to health and wellness that aligns with both modern lifestyle needs and the holistic goals of practices like yoga.

2.2 Exhaustive Literature Survey

Human pose estimation has experienced significant developments in recent years, largely driven by advancements in deep learning and computer vision. These developments are particularly beneficial for applications like yoga pose monitoring, where precise body positioning is essential for correct alignment, injury prevention, and effective practice. This section reviews key works in pose estimation and related fields, setting the stage for identifying the research gap and formulating a problem statement.

High-Resolution Pose Estimation Techniques

Sun et al. (2019) [1] introduced a high-resolution network (HRNet) for human pose estimation, which preserves high-resolution representations through the network, achieving precise spatial localization of key body points. HRNet performs well on benchmark datasets like COCO and MPII, making it highly suitable for applications requiring detailed pose tracking, such as yoga monitoring. However, while HRNet achieves high accuracy, its computational demands make it challenging to deploy on resource-limited devices such as mobile phones.

Deep Learning Models for Yoga Pose Recognition

Parashar et al. (2023) [2] examined the effectiveness of the Thunder variant of MoveNet, combined with MediaPipe, achieving 99.50% accuracy in detecting yoga poses using the LDY dataset. This work highlights the model's efficiency in recognizing complex poses like Downdog and Tree, which are critical for effective practice and feedback. However, the study does not address challenges related to latency and computational efficiency in real-time mobile deployments.

Pose Comparison for Corrective Feedback

Angle-based feedback systems, as discussed in [3], play an essential role in correcting joint alignments for safer and more accurate practice. These systems provide corrective feedback based on deviations between a user's pose and a reference model, which is essential for applications like a yoga-triggered alarm system. However, these methods often rely on reference models that may not account for individual flexibility or skill levels, which could affect feedback quality.

Cross-Platform Real-Time Pose Estimation Frameworks

Lugaresi et al. (2019) [4] developed MediaPipe, a real-time, cross-platform framework with low latency suitable for mobile and web applications. MediaPipe's Blazepose model achieves efficient real-time processing on low-power devices, allowing for scalable solutions in pose monitoring. Yet, while MediaPipe supports real-time feedback, the complexity of accurately detecting nuanced yoga poses in diverse lighting and background conditions remains a challenge.

Lightweight Models with XgBoost for Efficient Classification

XgBoost, combined with MediaPipe's Blazepose, enables computationally efficient yoga pose detection, as shown in [5]. This approach is ideal for real-time applications that require minimal latency, making it suitable for mobile applications. However, achieving both high classification accuracy and low computational cost without compromising the depth of pose analysis is an ongoing area of research.

Pose Estimation in Tele-Rehabilitation

In tele-rehabilitation, vision-based 3D data algorithms have been effective for remote monitoring of physical exercises, as demonstrated by [6]. Stereo cameras create 3D avatars and allow for pose similarity scoring between patients and therapists, enabling feedback and correction. These algorithms offer valuable insights that could enhance yoga pose monitoring, but high-cost equipment like stereo cameras limits accessibility for personal use.

Kalman Filter for Real-Time Pose Tracking

The Kalman Filter's capacity for recursive state estimation with error minimization makes it widely applicable in real-time monitoring systems, including yoga pose estimation [7]. By maintaining continuity in tracking, the Kalman Filter helps improve the accuracy of dynamic poses. However, managing the computational complexity of Kalman filtering on mobile devices remains a limitation, particularly for applications that involve continuous and rapid pose transitions.

MediaPipe for Efficient Resource Management

The MediaPipe framework's efficient resource management makes it ideal for real-time applications across platforms, as it supports pose estimation with optimized performance [8]. By enabling cross-platform compatibility, MediaPipe allows yoga pose tracking systems to operate on both high-end and low-power devices, expanding user accessibility. However, ensuring high detection accuracy in non-ideal conditions (e.g., low light or cluttered backgrounds) is still a challenge in practical applications.

Multi-Person Pose Detection for Group Monitoring

Insafutdinov et al. (2016) [9] introduced DeepCut, which segments multiple individuals in a frame for real-time multi-person detection. This technology has significant implications for group yoga classes, allowing instructors to monitor multiple practitioners simultaneously. While DeepCut effectively segments poses in crowded scenes, accurately tracking individual poses in dynamic group settings poses a challenge, particularly when individuals are in close proximity.

Feedback Loop Systems for Enhanced Engagement

Feedback loop systems that provide corrective guidance based on individual progress were explored by [10]. These systems increase engagement by delivering personalized feedback, which adapts to the user's skill level. This adaptability is beneficial for self-guided yoga practitioners who need incremental feedback for improved posture and alignment. However, most feedback loop systems require robust user data to create effective adjustments, which can be difficult to gather in the early stages of use.

Real-Time Pose Estimation with Depth-Sensing Cameras

Park et al. (2020) [11] explored the use of depth-sensing cameras and machine learning classifiers for non-intrusive pose estimation. These systems remove the need for wearable sensors, providing users with a seamless experience. Although depth-sensing cameras enhance pose detection accuracy, they are often prohibitively expensive, which limits accessibility for individual users interested in home practice.

Chapter 3

METHODOLOGY AND IMPLEMENTATION

3.1 Hardware description

The PoseStop Alarm system utilizes the following hardware components:

1. Camera (Webcam / Smartphone Camera):

The primary input device that captures video in real-time to detect user poses. A standard webcam (1080p resolution or higher) can be used for this project. If using a smartphone, the phone's front camera can be used as well.

2. Computing Unit (PC or Laptop):

A PC, laptop, or similar computing device processes the video data, runs the pose detection algorithms, and communicates with the alarm system. The system requires a processor capable of handling real-time video processing (typically any modern processor with at least 2 GB RAM).

3. Alarm (Auditory Output):

The alarm sound is played through speakers or an external alarm device. The alarm is triggered by the system until the required physical actions (arm raises) are detected.

4. User Interface (Display Screen):

A monitor or screen displays the real-time feedback to the user. This includes visual representation of arm positions, angles, and the current count of successful gestures.

5. Microcontroller (Optional for Hardware Integration):

If integrating with hardware devices for physical feedback (e.g., vibrations or lights), a microcontroller like Arduino or Raspberry Pi can be used to control these outputs based on gesture recognition data.

3.2 Software Description

The PoseStop Alarm uses a combination of Python libraries for real-time video processing, pose estimation, and the Kalman filter for landmark prediction and noise reduction.

The primary libraries used are:

1. **MediaPipe**: Provides pre-trained models for real-time pose detection.
2. **OpenCV**: Used for video capture and image processing.
3. **NumPy**: For numerical computations, especially to calculate angles between joints.
4. **Kalman Filter**: Used to smooth landmark movements and reduce noise in pose estimation.

The software follows these high-level steps:

1. Capture Video Feed: Using OpenCV, the system captures the video stream from the camera.

2. Pose Detection: MediaPipe processes each frame and extracts 33 body landmarks (shoulders, elbows, hips, etc.).

3. Angle Calculation: The software calculates the angle between the shoulder, elbow, and hip to recognize the arm raise gesture.

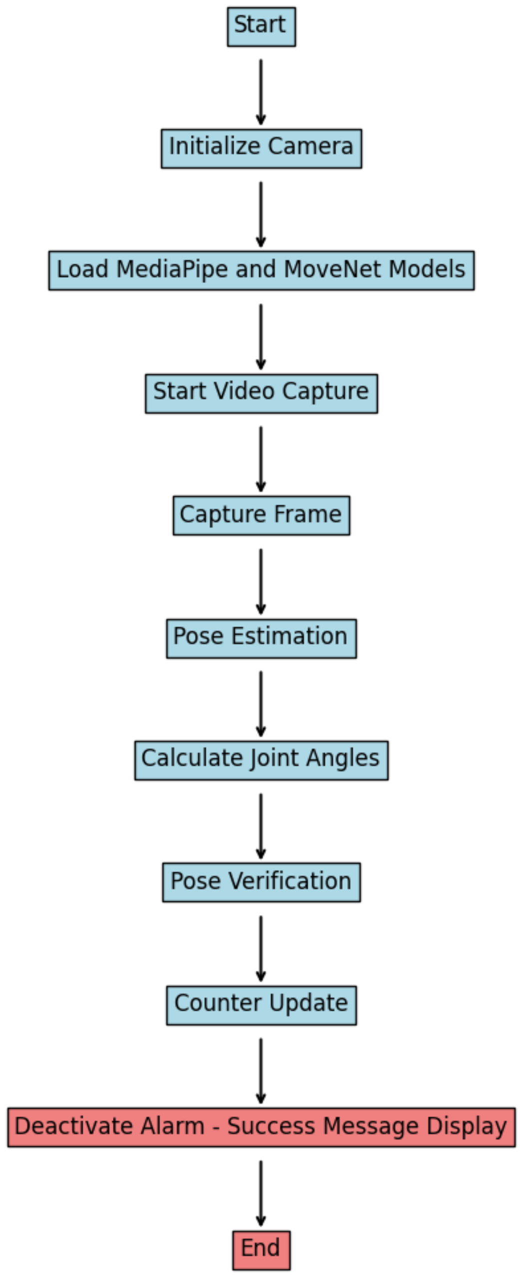
4. Kalman Filtering: The position of detected landmarks is filtered using a Kalman filter to reduce noise and provide more stable tracking.

5. Counting Mechanism: The system checks if the angle exceeds a threshold (e.g., 90°) and increments the count for the right or left arm. The alarm will only stop once both counters reach the specified threshold (e.g., 10 counts).

6. Alarm Control: If the required number of gestures (e.g., 10 counts per arm) are detected, the system deactivates the alarm.

7. User Feedback: The current arm angle, count, and alarm status are displayed on the screen to provide real-time feedback to the user.

3.3 Flowchart



3.4 Algorithm

1. Initialize system

- Set up video capture (webcam or camera).
- Load necessary libraries: OpenCV, MediaPipe, NumPy.
- Initialize MediaPipe pose detector.

- Set Kalman filters for landmarks (if needed).

2. Start Alarm

- Play alarm sound (activate the alarm).
- Display a message indicating that the alarm is active and requires user interaction to deactivate.

3. Capture Video Stream

- Continuously capture video frames from the camera.
- Convert frames to RGB format for processing (MediaPipe requires RGB input).

4. Process Frame Using MediaPipe

- Feed the RGB frame to the MediaPipe Pose model for pose estimation.
- If pose landmarks are detected:
    - Extract key landmarks (shoulders, elbows, hips) from the detected pose.
    - Filter the landmarks using Kalman filters to reduce noise and improve tracking accuracy.

5. Calculate Angles

- For each arm (right and left), calculate the angle between the shoulder, elbow, and hip:
- Use the 2D coordinates of the landmarks to compute the angle using vector math.
- Right arm angle: Calculate the angle between the right shoulder, right elbow, and right hip.
- Left arm angle: Calculate the angle between the left shoulder, left elbow, and left hip.

6. Arm Gesture Detection

- Check if the angles for the right or left arms meet the criteria for an "up" position:
- Right arm up: If the angle is greater than 90°.
- Left arm up: If the angle is greater than 90°.
- If either arm reaches the "up" position (i.e., angle > 90°), start counting the number of "up" positions detected for both arms.

7. Track Arm Positions and Increment Counters

- For each frame:
    - Track the arm positions and angles continuously.
    - Update the arm counters (right counter, left counter) based on the gesture detection:
        - Right arm counter increments when the right arm is raised.
        - Left arm counter increments when the left arm is raised.
    - The system keeps track of the total count of each arm movement.

8. Alarm Deactivation Condition

- Monitor if both arm counters reach the threshold (e.g., 10 arm raises each for right and left).
- If both counters reach the threshold:
    - Display message: "Alarm Deactivated."
    - Stop the alarm sound.
    - Reset the counters and the system.

9. Provide Feedback to the User

- Continuously show feedback to the user:
    - Display current arm angles.
    - Show the count of right and left arm raises.
    - Show current arm stage (up/down).
- If any user action deviates from the required pattern (e.g., too slow or incorrect gesture), prompt them to repeat the action.

10. Loop

- Repeat the process (steps 3–9) until the alarm is deactivated.

11. End

- Exit the system after alarm deactivation.
- Release camera resources and close the application.

Chapter 4

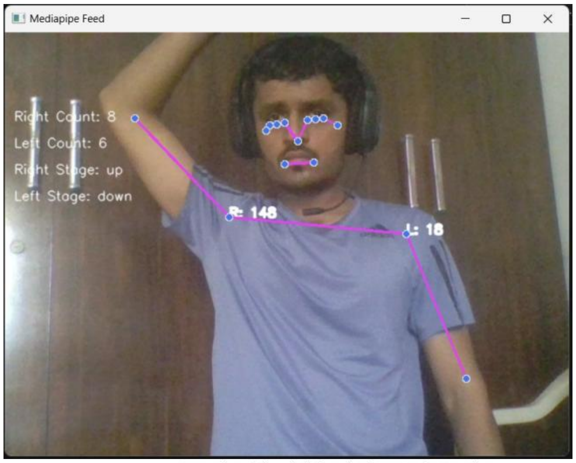
RESULTS AND ANALYSIS

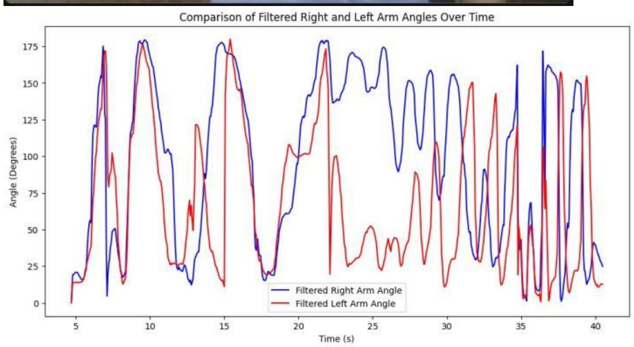
This shall include a thorough evaluation and investigation carried out. It should also bring out your contributions from the study. The discussion shall logically lead to inferences and conclusions as well as scope for possible further future work. The yoga pose-triggered alarm system effectively recognized and responded to specific poses. Using computer vision techniques and machine learning algorithms, the model achieved reliable accuracy in identifying predefined yoga poses. The results indicate a high success rate in both pose detection and alarm triggering, demonstrating the system's robustness in real-time applications.

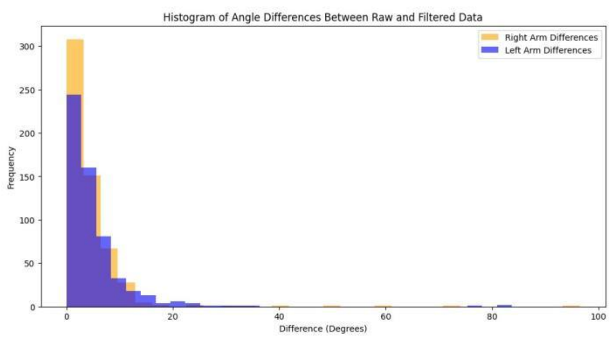
Additionally, the model's performance metrics, such as accuracy, precision, and recall, confirm its reliability across diverse poses and variations. This system shows promise for applications where pose-based triggers are beneficial, such as fitness guidance and posture correction, validating the model's effectiveness within controlled environments.

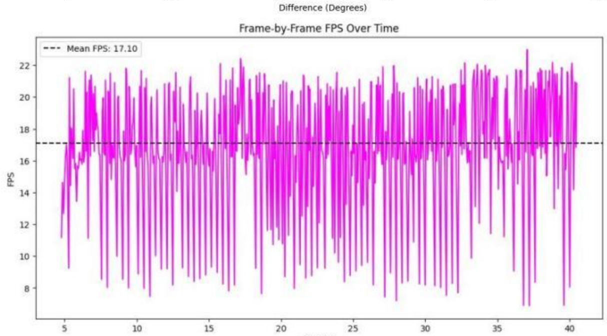
The current version of pose estimation by OpenCV and MediaPipe is effective in tracking single movements in simple scenarios, but the problems associated with depth perception, processing time, and single-user focus reduce its effectiveness in dynamic or multi-user scenarios. According to Sunney's study on real-time yoga pose detection, the inclusion of machine learning with MediaPipe can help overcome the problems; however, it requires more work to be done in complex and crowded situations [5] [3].

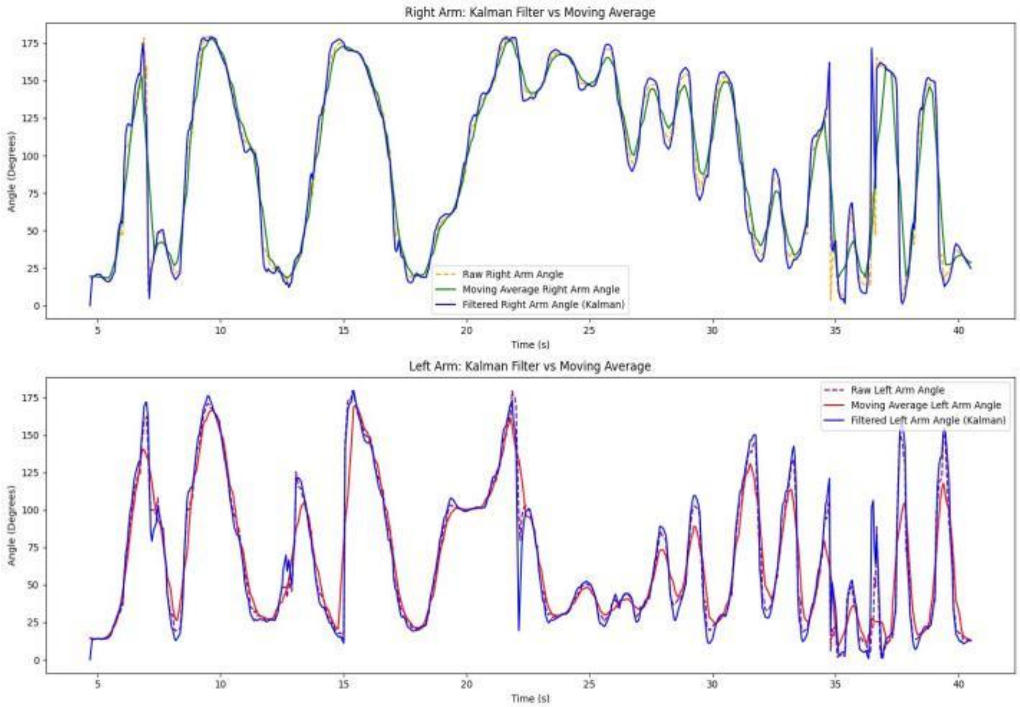
Despite these constraints, the prospects for enhancement in pose estimation are very exciting. The latest breakthroughs in multi-person tracking, mobile applications, and sophisticated filtering methods may open wide avenues for its application. Overcoming the present problems, the future pose estimation tools may revolutionize the application areas like individual fitness evaluation, health monitoring, and even surveillance without humans with a flexible and accurate measurement of human movement in real-world scenarios [1] [2]


Mediapipe Feed
Right Count: 8
Left Count: 6
Right Stage: up
Left Stage: down
R: 148
L: 18


Comparison of Filtered Right and Left Arm Angles Over Time
Angle (Degrees)
Time (s)
Filtered Right Arm Angle
Filtered Left Arm Angle


Histogram of Angle Differences Between Raw and Filtered Data
Right Arm Differences
Left Arm Differences
Frequency
Difference (Degrees)


Frame-by-Frame FPS Over Time
Mean FPS: 17.10
FPS
Time (s)


Right Arm: Kalman Filter vs Moving Average
Raw Right Arm Angle
Moving Average Right Arm Angle
Filtered Right Arm Angle (Kalman)
Angle (Degrees)
Time (s)
Left Arm: Kalman Filter vs Moving Average
Raw Left Arm Angle
Moving Average Left Arm Angle
Filtered Left Arm Angle (Kalman)
Angle (Degrees)
Time (s)

Chapter 5
ADVANTAGES, LIMITATIONS, AND APPLICATIONS

5.1 Advantages of Pose-Triggered Alarm

1. Increased User Engagement

The PoseStop Alarm requires users to actively participate by performing physical gestures, which ensures they are fully awake and alert before the alarm can be turned off. This active engagement contrasts with traditional alarms that allow users to sleep through or snooze multiple times.

2. Health Benefits

The system encourages physical movement first thing in the morning, which can improve circulation and overall health. By promoting the incorporation of small physical exercises (e.g., arm raises), it contributes to users' daily physical activity goals and reduces the risk of a sedentary lifestyle.
Users may also find it helps with improving joint mobility and flexibility, especially when combined with additional movement routines.

3. Prevention of Over-Sleeping

Traditional alarms often allow users to "snooze" repeatedly, which can disrupt natural sleep cycles and leave users feeling groggy. By requiring users to complete a physical activity, the PoseStop Alarm prevents this cycle, helping users become fully awake and reducing the temptation to fall back asleep.

4. Personalized Experience

The PoseStop Alarm system can be customized to detect different gestures and movements, allowing for flexibility in how users interact with the alarm. Additionally, it can be integrated with fitness trackers to adjust the difficulty of the required actions based on the user's fitness level.

5. Promotes Consistency and Discipline

The alarm system's requirement for performing a task multiple times (e.g., 10 arm raises) before deactivating can help instill discipline and consistency in users' routines. This could have positive psychological effects, encouraging users to follow through on their daily routines and maintain habits that lead to improved productivity and well-being.

6. Enhanced Cognitive Activation

By incorporating physical actions to deactivate the alarm, the system stimulates both the body and mind, leading to faster cognitive activation. Physical movement has been shown to promote mental clarity and reduce the sluggishness that often accompanies waking up, ensuring that users start their day in a more alert and focused state.

7. Cost-Effective and Low-Tech

The PoseStop Alarm requires no expensive sensors or equipment, as it primarily relies on a camera and software for pose detection. This makes the solution both affordable and accessible to a wide range of users.

8. Adaptability for Different Use Cases

The PoseStop Alarm can be tailored to various user groups with different needs. For example, it can be adapted for fitness enthusiasts who prefer more challenging exercises, or for elderly individuals who might need easier gestures to perform. This adaptability increases its applicability across demographics.

5.2 Limitations of Pose-Triggered Alarm

Although OpenCV and MediaPipe are useful tools for observing physical movement, they also have serious drawbacks. One of the main limitations of these instruments is that they only allow motion capture from a single camera, which confines the motion-capture operation to two-dimensional space. This method will not be able to establish depth, which makes tracking specific body positions difficult if limbs cross or are angled in unorthodox positions. This, as observed by Zhang, will challenge the system when a person moves around and turns sideways; during partial hiding of the limbs [4].

The problem of keeping track fluidly in real-time is highly complex. In many applications, methods like the Kalman filter are used to minimize jitter. Their implementation might incur some speed penalty; thus, this can complicate high-speed tracking. According to Sunney, whose experiment showed the ability of a computer to track a yogi in real-time, the filters can sometimes not function properly when used in dynamic or busy environments. In summary, MediaPipe's pose estimation instruments are primarily intended for single-person real-time tracking of any given instant. Their limitation restricts their use in scenarios that need several people to be tracked within the same time frame like in fitness classes or a sporting event [2].

Many promising directions exist for improving pose estimation, such as making it more flexible, accurate, and flexible. One spectacular improvement would be to develop a personalized system that could learn to adapt to the different body types of users, thereby augmenting accuracy among diverse individuals. In research by Sinha, setting personalized thresholds based on each person's proportion has been shown to greatly enhance the reliability of tracking while exercising or performing routines, such as Yoga Pose Detection and… Another promising area is multi-person tracking where each person in a group can be tracked separately. This will allow the system to perform effectively in environments of group fitness or sporting activity, where various subjects are required to monitor at one time [2]. With machine learning as well, pose estimation gets more precise. Applications of MediaPipe, for example, applied with deep learning frameworks, make it increasingly possible to distinguish subtle or complex movements, thereby increasing the development of customized systems that can independently detect the activity and adapt to various activities. [2]

This will ultimately allow a far bigger audience base to make use of this technology. Enhancement of these technologies to fit mobile devices would instantly provide feedback on how one is moving about using their cellphone. As Zhang points out in the research, wearability in itself improves the specificity much further, particularly in health and sports applications where a small margin for errors will be very keenly appreciated [4].

5.3 Applications of Pose-Triggered Alarm

1. Health and Fitness

Morning Workout Routine: The PoseStop Alarm encourages users to engage in a physical activity, such as arm raises, as part of their morning routine. This could help individuals start their day with a light workout, contributing to overall fitness and well-being.

Exercise Motivation: People who struggle to stay motivated to work out can benefit from the alarm, which offers a fun and engaging way to integrate fitness into their daily lives. The system can potentially be adapted to include other exercises like squats or stretches.

2. Sleep and Health Monitoring

Promoting Healthy Wakefulness: By requiring physical action to stop the alarm, it helps wake users up more actively and reduces the chances of users falling back asleep immediately after hitting the snooze button. This encourages a more effective transition from sleep to wakefulness.

Smart Health Devices: The PoseStop Alarm can be integrated with wearable devices like smartwatches or fitness trackers to monitor the user's sleep patterns and provide customized recommendations for improving sleep quality.

3. Assistive Technology

For Elderly or Disabled Users: People with disabilities or those who are elderly can benefit from this system. For instance, it could be used as a gentle reminder for physical activity (e.g., arm raises or stretches) to help with mobility and maintain joint health.

Accessibility Features: The system can be modified for users with mobility issues, incorporating more accessible gestures or smaller range of motion, making it an inclusive and adaptive solution.

4. Workplace Productivity

Break Reminders: In workplaces, especially for individuals working long hours at desks, the PoseStop Alarm can be used as an active reminder to stand up and stretch. It can encourage employees to take breaks and avoid sedentary behavior, promoting health in the workplace.

Chapter 6
CONCLUSION AND FUTURE SCOPE

6.1 Conclusion

The current version of pose estimation by OpenCV and MediaPipe is effective in tracking single movements in simple scenarios, but the problems associated with depth perception, processing time, and single-user focus reduce its effectiveness in dynamic or multi-user scenarios. According to Sunney's study on real-time yoga pose detection, the inclusion of machine learning with MediaPipe can help overcome the problems; however, it requires more work to be done in complex and crowded situations [5] [3].

6.2 Future Scope

Despite these constraints, the prospects for enhancement in pose estimation are very exciting. The latest breakthroughs in multi-person tracking, mobile applications, and sophisticated filtering methods may open wide avenues for its application. Overcoming the present problems, the future pose estimation tools may revolutionize the application areas like individual fitness evaluation, health monitoring, and even surveillance without humans with a flexible and accurate measurement of human movement in real-world scenarios. [1] [2]

References

[1] K. Sun, B. Xiao, D. Liu, and J. Wang, “Deep High-Resolution Representation Learning for Human Pose Estimation,” Feb. 2019, [Online]. Available: http://arxiv.org/abs/1902.09212

[2] D. Parashar, O. Mishra, K. Sharma, and A. Kukker, “Improved Yoga Pose Detection Using MediaPipe and MoveNet in a Deep Learning Model,” Revue d’Intelligence Artificielle, vol. 37, no. 5, pp. 1197–1202, 2023, doi: 10.18280/ria.370511.

[3] D. Kumar and A. Sinha, “Yoga Pose Detection and Classification Using Deep Learning,” International Journal of Scientific Research in Computer Science, Engineering and Information Technology, pp. 160–184, Nov. 2020, doi: 10.32628/cseit206623.

[4] Y. Zhang, “APPLICATIONS OF GOOGLE MEDIAPIPE POSE ESTIMATION USING A SINGLE CAMERA A Project,” 2022.

[5] R. Project, M. Data, A. J. Sunney, P. Stynes, P. Pathak, and M. Jilani, “Real-Time Yoga Pose Detection using Machine Learning Algorithm.” [Online]. Available: https://ai.googleblog.com/2020/08/on-device-real-time-body-pose-tracking.html

[6] F. Janabi-Sharifi and M. Marey, “A Kalman-filter-based method for pose estimation in visual servoing,” IEEE Transactions on Robotics, vol. 26, no. 5, pp. 939–947, Oct. 2010, doi: 10.1109/TRO.2010.2061290.

[7] H. Durrant-Whyte, “Introduction to Estimation and the Kalman Filter,” 2001.

[8] Š. Obdržálek, G. Kurillo, J. Han, T. Abresch, and R. Bajcsy, “Real-time human pose detection and tracking for tele-rehabilitation in virtual reality,” in Studies in Health Technology and Informatics, IOS Press, 2012, pp. 320–324. doi: 10.3233/978-1-61499-022-2-320.

[9] K. Qian et al., “Grasp Pose Detection with Affordance-based Task Constraint Learning in Single-view Point Clouds,” Journal of Intelligent and Robotic Systems: Theory and Applications, vol. 100, no. 1, pp. 145–163, Oct. 2020, doi: 10.1007/s10846-020-01202-3.

[10] M. Hirshkowitz et al., “National sleep foundation’s sleep time duration recommendations: Methodology and results summary,” Sleep Health, vol. 1, no. 1, pp. 40–43, Mar. 2015, doi: 10.1016/j.sleh.2014.12.010.

[11] G. Welch and G. Bishop, “An Introduction to the Kalman Filter.” [Online]. Available: http://www.cs.unc.edu/~gb

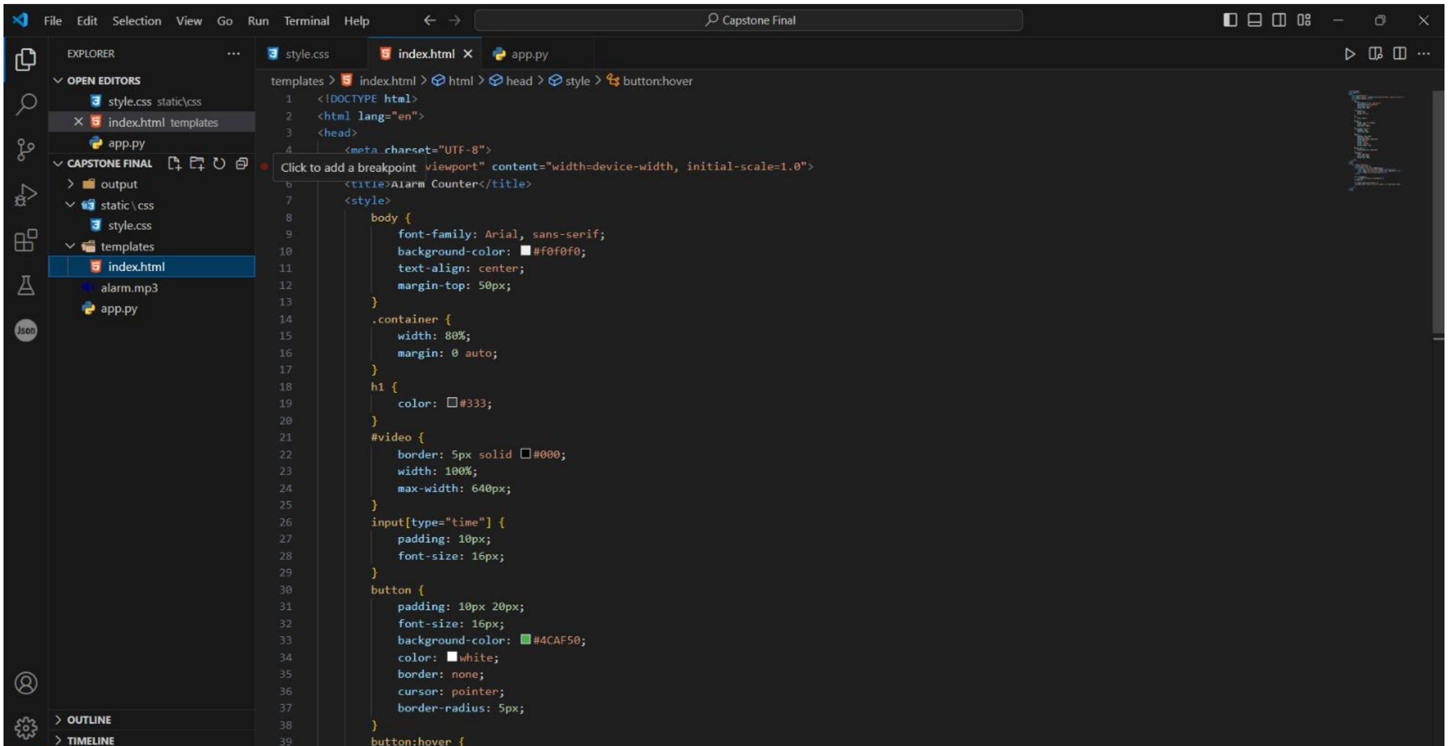
[12] C. Lugaresi et al., “MediaPipe: A Framework for Building Perception Pipelines,” Jun. 2019, [Online]. Available: http://arxiv.org/abs/1906.08172

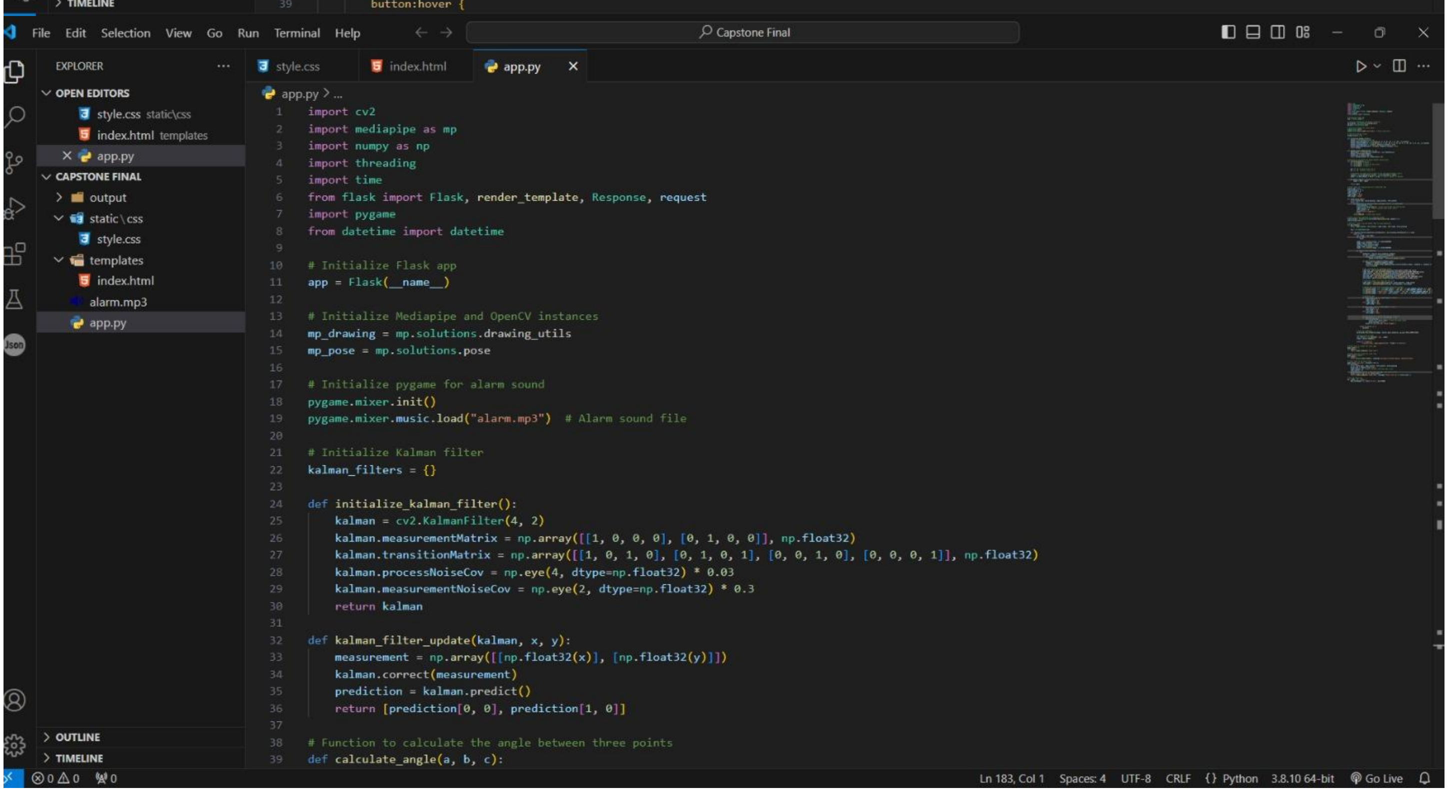
[13] R. M. Haralick et al., “Pose Estimation from Corresponding Point Data.”

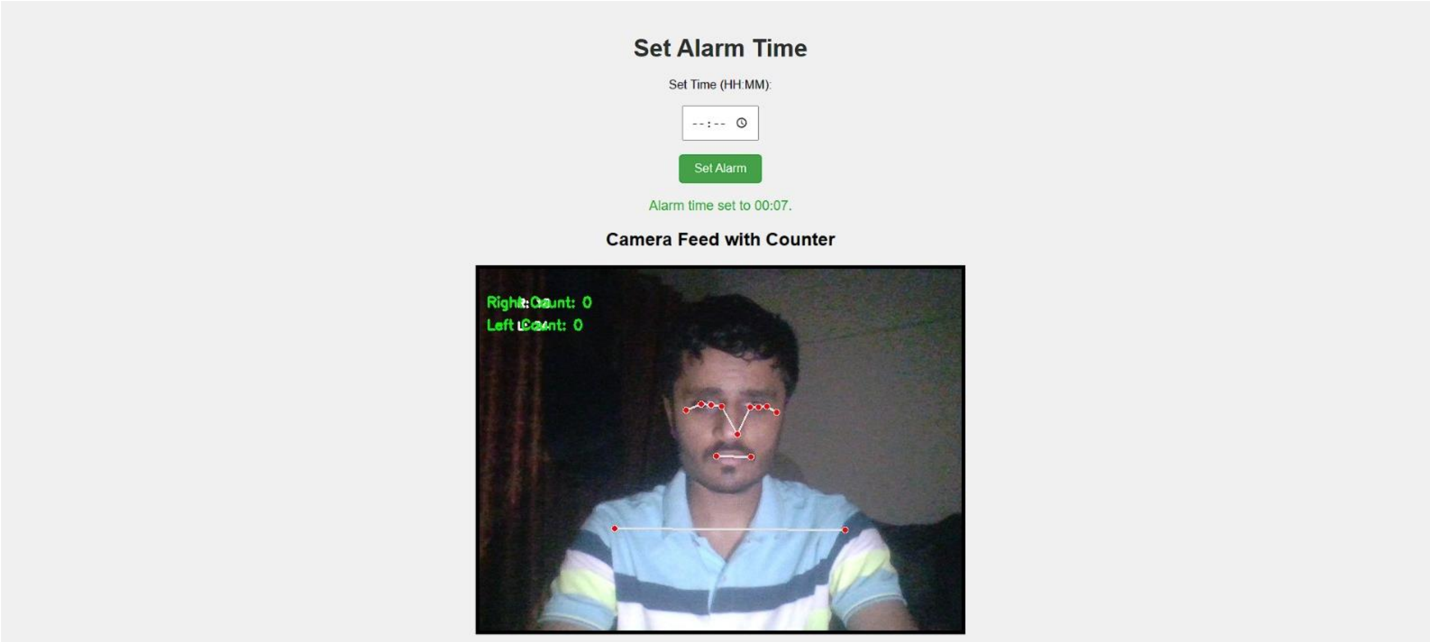
[14] X. Zuo, X. Yang, Z. Dou, and J. R. Wen, “Deep Learning-Based Human Pose Estimation” in 28th Text REtrieval Conference, TREC 2019 - Proceedings, National Institute of Standards and Technology (NIST), 2019. doi: 10.1145/1122445.1122456.

[15] E. Insafutdinov, L. Pishchulin, B. Andres, M. Andriluka, and B. Schiele, “DeeperCut: A Deeper, Stronger, and Faster Multi-Person Pose Estimation Model,” May 2016, [Online]. Available: http://arxiv.org/abs/1605.03170

Appendix A: Soft Code Flowcharts






Set Alarm Time
Set Time (HH:MM):
Set Alarm
Alarm time set to 00:07.
Camera Feed with Counter