

Article

Vision-Based Activity Recognition for Unobtrusive Monitoring of the Elderly in Care Settings

Rahmat Ullah ^{1,*} , Ikram Asghar ² , Saeed Akbar ³ , Gareth Evans ⁴ , Justus Vermaak ⁵, Abdulaziz Alblwi ⁶  and Amna Bamaqa ⁶ 

¹ School of Computer Science and Electronic Engineering, University of Essex, Colchester CO4 3SQ, UK

² Department of Computing & Games, Teesside University, Middlesbrough TS1 3BX, UK; i.asghar@tees.ac.uk

³ Department of Physical & Numerical Sciences, Qurtuba University of Science & Information Technology, Peshawar 25150, Pakistan; saedakbr@qurtuba.edu.pk

⁴ Centre of Excellence in Mobile and Emerging Technologies, University of South Wales, Pontypridd CF37 4BD, UK; gareth.evans@southwales.ac.uk

⁵ Skystm Limited, Barry CF62 5BE, UK; justus@henleycarecentre.com

⁶ Department of Computer Science, Applied College, Taibah University, P.O. Box 344, Madinah 41477, Saudi Arabia; ablwi@taibahu.edu.sa (A.A.); abamaqa@taibahu.edu.sa (A.B.)

* Correspondence: rahmat.ullah@essex.ac.uk

Abstract: As the global population ages, robust technological solutions are increasingly necessary to support and enhance elderly autonomy in-home or care settings. This paper presents a novel computer vision-based activity monitoring system that uses cameras and infrared sensors to detect and analyze daily activities of elderly individuals in care environments. The system integrates a frame differencing algorithm with adjustable sensitivity parameters and an anomaly detection model tailored to identify deviations from individual behavior patterns without relying on large volumes of labeled data. The system was validated through real-world deployments across multiple care home rooms, demonstrating significant improvements in emergency response times and ensuring resident privacy through anonymized frame differencing views. Upon detecting anomalies in daily routines, the system promptly alerts caregivers and family members, facilitating immediate intervention. The experimental results confirm the system's capability for unobtrusive, continuous monitoring, laying a strong foundation for scalable remote elderly care services and enhancing the safety and independence of vulnerable older individuals.

Keywords: activity recognition; action recognition; daily living activity recognition; elderly care; fall detection; remote monitoring



Academic Editor: Daniele Giansanti

Received: 26 March 2025

Revised: 30 April 2025

Accepted: 1 May 2025

Published: 4 May 2025

Citation: Ullah, R.; Asghar, I.; Akbar, S.; Evans, G.; Vermaak, J.; Alblwi, A.; Bamaqa, A. Vision-Based Activity Recognition for Unobtrusive Monitoring of the Elderly in Care Settings. *Technologies* **2025**, *13*, 184. <https://doi.org/10.3390/technologies13050184>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

According to recent statistics from the World Health Organization, the global population is aging rapidly. By the end of this decade, one in every six people worldwide will be over the age of 60. This demographic shift is anticipated to grow, with projections suggesting that there will be over 2 billion people aged 60 and above by the year 2050. Historically, trends of aging populations were predominantly observed in developed and high-income countries. However, current projections indicate that aging will also become a significant phenomenon in low and middle-income countries. By 2050, nearly 80% of the elderly population will reside in these regions [1]. It is crucial to investigate and implement care-based and health monitoring technologies in the living environments of the ageing population, such as care homes and independent homes, to ensure timely care is provided when needed [2].

The aging population presents a complex mix of challenges and opportunities. As life expectancy increases, individuals have more opportunities to engage in new learning experiences, embark on different career paths later in life, and contribute their knowledge and skills to society. However, the full participation of elderly individuals in these opportunities heavily depends on their health. Many older adults face a range of medical issues, including sensory impairments, mobility restrictions, and chronic conditions such as diabetes and dementia, which can significantly impact their quality of life [3].

Monitoring the daily activities of elderly individuals is crucial for assessing their health status. These activities, encompassing daily routines, physical movements, eating habits, sleeping patterns, and more, serve as vital indicators of their overall health [4]. By carefully observing and tracking these activities, caregivers, family members, and healthcare providers can detect any deviations from normal patterns, providing early warning signs of potential health issues. Moreover, monitoring physical activity can pinpoint changes in mobility and balance, potentially flagging early signs of conditions that could lead to falls or other serious health complications [5].

In today's rapidly advancing technological landscape, integrating artificial intelligence (AI) and computer vision into healthcare monitoring systems can significantly enhance the efficiency and effectiveness of care. These technologies offer the revolutionary potential to improve the management of healthcare facilities, particularly those catering to vulnerable groups such as the elderly or those with chronic health conditions [6]. Traditional monitoring methods often struggle to detect emergencies or significant health changes in real time, highlighting the need for more innovative solutions to these challenges. The COVID-19 pandemic, for instance, has exposed high mortality rates among older adults in care facilities, largely due to inadequate monitoring and care, underscoring the importance of robust remote monitoring systems that can ensure continuous care and timely interventions [7].

Research into wearable sensors for monitoring human activity provides a complex perspective, emphasizing the benefits and limitations of such technology, especially for the elderly. While wearable sensors can offer detailed insights into an individual's health and activity levels, they may not always be reliable for elderly users who might forget to wear them or find them uncomfortable. This can lead to significant gaps in data collection and reduce the effectiveness of monitoring based solely on these devices [8,9]. Furthermore, due to unique environmental challenges, traditional deep learning techniques are often inadequate for accurately identifying activities within specific indoor settings, such as care homes. These models rely on large volumes of labeled data to identify and classify activities indicative of health risks.

Indoor Human Activity Recognition requires models that are specifically tailored to the typical activities of the environment. Therefore, a one-size-fits-all model is not appropriate for specialized settings, as noted in a review by Bhola et al. [8]. Anomaly detection models are a great alternative to traditional machine learning techniques for human activity detection. These models can effectively operate with unlabeled data.

This paper proposes a vision-based monitoring system designed to enable effective remote monitoring in elderly care settings. The key contributions of our work are as follows:

- A novel computer vision system is developed that integrates cameras and infrared sensors to monitor daily activities of elderly individuals without requiring wearable devices.
- We implement lightweight anomaly detection algorithms capable of identifying deviations from normal behavior patterns without relying on large volumes of labeled data. This makes the system highly suitable for real-world care home environments where critical events are rare and collecting large volumes of labeled anomaly data is not impractical.

- The system is evaluated through deployments in real elderly care settings, demonstrating significant improvements in emergency response times and overall resident safety.
- The system maintains resident privacy by employing frame differencing techniques that anonymize video outputs while retaining critical movement information for accurate activity recognition.

In contrast to many deep learning-based approaches requiring substantial computational resources, extensive labeled datasets, and complex maintenance, the proposed lightweight system leverages simple frame differencing and anomaly detection techniques. This design ensures real-time performance, minimizes installation and operational costs, preserves resident privacy, and offers high scalability for elderly care environments. Our system demonstrates that in safety-critical applications like remote elderly monitoring, purposeful algorithmic simplicity can translate into greater practicality, reliability, and faster emergency responses than overly complex models.

2. Related Work

Recent computer vision and machine learning studies demonstrate significant advancements across various applications, from aerial surveillance to autonomous vehicle navigation and human-computer interaction. A comprehensive review of the current landscape in video-based human activity recognition systems, examining the essential components such as object segmentation, feature extraction, activity detection, and classification, is conducted in [10]. The survey spans three primary application areas: surveillance, entertainment, and healthcare. While noting the significant advancements in the field, the paper also identifies persistent challenges and technical hurdles that impede practical, real-world implementation. A comprehensive scoping and in-depth examination of audio and video-based human activity recognition systems within healthcare settings is presented in [11]. The review provides a critical evaluation of the performance, robustness, and scalability of these systems, as well as their privacy and security implications. Despite their potential, there are still challenges in implementing such systems into clinical practice due to privacy concerns and the complexity of real-world implementation.

2.1. Vision-Based Human Activity Recognition Systems

A Diminutive Multi-Dimensional Locality Coding Convolutional Neural Network (DMLC-CNN) is proposed for human activity recognition from UAV-captured videos [12]. The model introduces a novel feature extraction and encoding process suitable for complex and dynamic backgrounds in aerial surveillance, resulting in superior accuracy compared to traditional methods. Similarly, deep learning techniques recognize in-bed movements from video streams in a clinical setting in [13]. The paper proposed a 3D MoCap combined with skeleton-based action recognition as a promising technique for improving the accuracy of clinical in-bed monitoring systems. However, the effectiveness of the proposed methods heavily relies on the availability of high-quality, labeled datasets for training and validation, which are scarce in clinical contexts.

A system was designed to provide real-time feedback to stroke patients performing daily activities necessary for independent living. It used overhead cameras and histogram-based recognition methods to monitor the position and movement of the patient's hands and the objects they manipulate [14]. The key events are recognized and interpreted in the context of a model of the coffee-making task. A computer vision-based human activity recognition system for assisted living systems was proposed in [15]. The proposed system addresses some of the limitations of traditional sensor-based systems by using a numerical dataset derived from key-joint angles, distances, and slopes. This approach, using data from 20,000 images to detect five human postures, sitting, standing, walking, lying, and falling

enabled efficient training and high-accuracy detection. A video sensor-based HAR system for elderly care was proposed in [16]. The proposed system used depth imaging sensors to generate human skeletons and joint information for precise activity recognition. This approach outperforms traditional RGB sensor-based systems by providing richer data necessary for detailed analysis of elderly activities within smart environments. An AI-driven privacy-preserving monitoring system for elderly care was proposed in [17]. Their system employs the YOLOv8 model to detect individuals in real-time and replaces them with anonymized 2D avatars, ensuring privacy protection. However, the proposed system's performance can be less effective under varying environmental conditions like lighting and obstacles, which can hinder sensor accuracy and activity recognition. Moreover, it requires high computational resources for complex algorithms, raising concerns about its feasibility in real-time applications.

2.2. Wearable Sensor-Based Human Activity Recognition Systems

Wearable technology is attracting considerable interest for its ability to track and improve physical activity in older adults [18]. An extensive review of the advancements in smart home technologies specifically tailored for elderly healthcare is presented in [19]. Their work underscores the integration of e-health and assisted living technologies within the Internet of Things (IoT) framework. The authors discuss critical components such as sensors, actuators, and communication networks that are instrumental in monitoring and assisting the elderly. A study was conducted by [20], which showed that wearable devices, along with mobile-based intermittent coaching, can significantly boost physical activity and health outcomes, especially in pre-frail older adults. Similarly, Smirnova et al. found that accelerometers, commonly used in wearable technology, provide reliable measurements of physical activity and are better predictors of five-year mortality in older adults than traditional methods [21]. A similar study identified wearable devices as crucial facilitators of physical activity, especially for older adults with cardiovascular diseases or associated risks [22].

A comprehensive review is conducted on the use of wearable sensors in human activity recognition (HAR), mainly through deep learning approaches [23]. The paper highlights the wide range of applications for wearable technologies, from fitness and lifestyle enhancement to critical roles in healthcare and rehabilitation. An ensemble measurement-based deep learning model is proposed in [24], which utilizes four Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM)-based models for smartphone sensor-based Human HAR problems. The model was tested on three publicly available datasets and demonstrated commendable performance in activity prediction with reasonable accuracy. However, the lack of advanced processing techniques for time-series sensor data potentially impacts its performance effectiveness.

2.3. Fall Detection Systems and Low-Light Activity Recognition

Recent advancements in human activity monitoring have also emphasized the significance of fall detection, particularly for the elderly, using different machine and deep learning models. For instance, a modified NASNet model incorporating Local Binary Patterns (LBP) features distinguishes itself by achieving an impressive accuracy, precision, recall, and F1 score of 99% [25]. A similar study developed a vision-based fall detection system capable of operating under low-light conditions. Their method involves enhancing images and tracking objects using a dual illumination estimation algorithm and then utilizing YOLOv7 + Deep SORT/w DUAL for fall detection [26]. A framework for Human Activity Recognition in low-light conditions using a dual stream network that combines Convolutional Neural Networks and Transformers is proposed in [27]. This cloud-assisted

IoT framework enhances video frame quality at the edge and utilizes an optimized network architecture for efficient and accurate activity recognition in low-lighting scenarios.

However, despite these positive findings, research remains inconclusive regarding the absolute validity and reliability of these devices for evaluating physical activity and health-related outcomes in older adults, highlighting the need for further research and adaptation to suit this age group better [28]. In conclusion, substantial evidence supports the effectiveness of wearable devices in monitoring and enhancing physical activity among older adults. There is a simultaneous recognition of the necessity for continued research, particularly in terms of device adaptability and reliability in this age group.

3. Materials and Methods

This study used a co-design approach. The research process started with multiple requirements workshops between the research team, care industry experts, and end users to understand the care industry's needs. The research aimed to develop a system that provides a way to demonstrate the possibilities of applying computer vision techniques within the care industry. Based on the nature of this study, the prototyping approach was best suited for the system development. The data to test the system's performance was collected through a pilot study, which records a person's movements in a predefined scenario. The data were captured with the help of cameras and infrared sensors, which monitor an individual's movements. The pilot study data marked a first step in determining the suitability and efficiency of the developed system.

3.1. Proposed Method

A series of computer vision algorithms and data handling methods are used. These algorithms are pivotal for the real-time monitoring and activity analysis of elderly residents in care environments. By using these techniques, the system provides detailed insights into daily activities and quickly identifies any deviations that may indicate potential health or safety concerns. Using raw image data requires extensive data cleaning and pre-processing techniques. This process required lots of experimentation to ensure optimal results while providing the user with enough flexibility and control. An illustrative diagram of this process can be seen in Figure 1, which outlines the sequence from frame capture to motion detection.

Frame differencing was used to detect the movement in data obtained using camera and infrared sensors. In this method, the captured video frames are converted into grayscale. The absolute difference between consecutive frames is calculated to highlight areas where motion has occurred. This difference image undergoes a binary thresholding process, simplifying the image to binary values representing motion and no-motion areas. This step is crucial for isolating relevant motion data from background noise.

Once motion is detected through frame differencing, the system uses a moving average filter to assess the continuity and significance of detected movements. This helps distinguish between incidental movements, such as a curtain fluttering, and significant activities like walking or falling. Logging an activity as 'true motion' is made based on the consistency and duration of detected movement across several frames. The Algorithm 1 gives pseudocode for the Frame Differencing Algorithm. The algorithm processes each frame to identify significant movements by comparing differences, applying a binary threshold, and analyzing contours for robust motion detection.

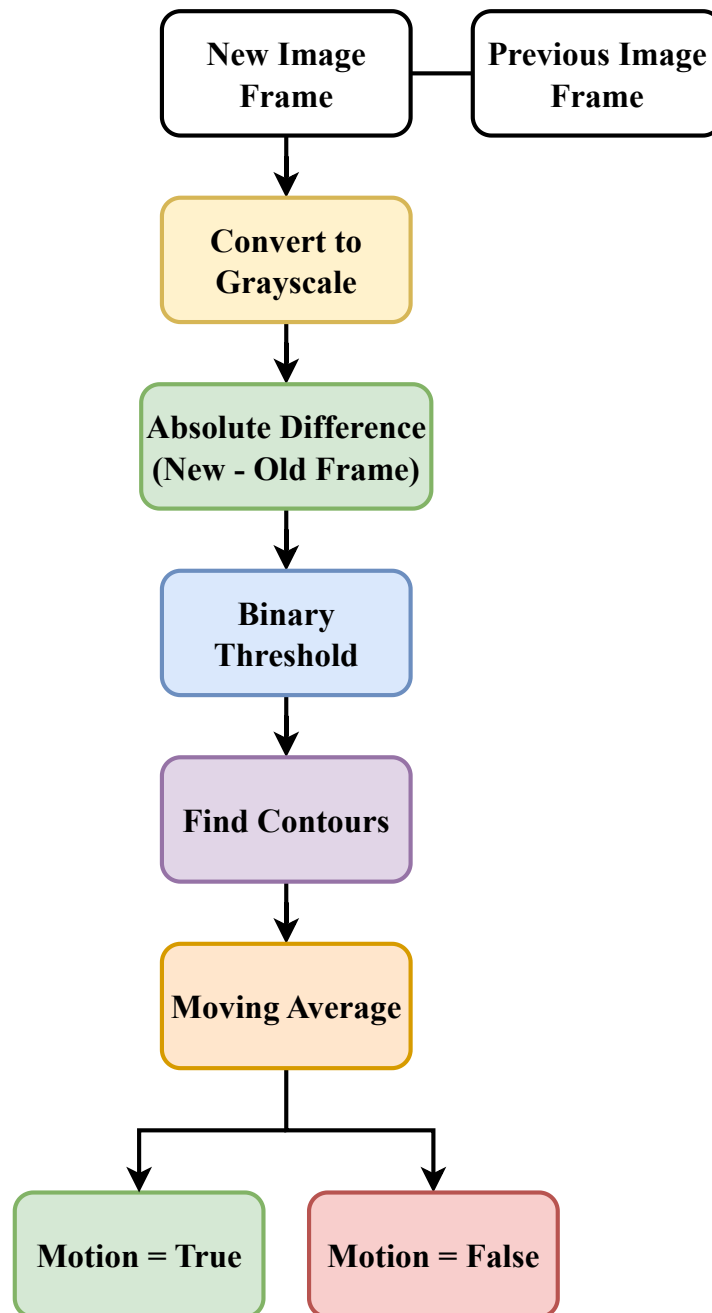


Figure 1. Workflow of the frame differencing algorithm. The process starts with converting video frames to grayscale and then computing the absolute difference between consecutive frames. The resulting difference image is then held in a binary format, isolating significant movement areas.

Several key parameters and techniques are utilized to optimize the detection and analysis of motion. These parameters include the detection threshold, binary threshold, and window length, each contributing uniquely to the system's efficacy.

Thresholding in computer vision is used to segment images, effectively separating the foreground from the background [29]. The detection threshold is crucial for determining the existence of motion based on the number of contours identified in processed images. This parameter was adjusted from 1 to 125, where 1 indicates a higher sensitivity to motion, capturing finer movements. Conversely, a higher threshold value makes the system less sensitive and useful in environments where incidental movements are common. This flexibility

allows the system to be finely tuned in real time to suit different operational conditions, ensuring that motion detection is accurate and relevant to the specific care setting.

Algorithm 1: Frame Differencing Algorithm used in motion detection

Result: Detect motion in video frames and mark frames containing significant motion

```

1 Initialize: previous_frame  $\leftarrow$  NULL;
2 foreach frame in video stream do
3   Convert frame to grayscale;
4   if previous_frame is not NULL then
5     absolute_difference  $\leftarrow$  ComputeAbsoluteDifference(frame, previous_frame);
6     binary_image  $\leftarrow$  ApplyBinaryThreshold(absolute_difference);
7     contours  $\leftarrow$  FindContours(binary_image);
8     if contours exist then
9       motion_average  $\leftarrow$  CalculateMovingAverage(contours);
10      if motion_average indicates significant motion then
11        Mark frame as containing motion;
12        Store or process frame as required;
13      end
14    end
15  end
16  previous_frame  $\leftarrow$  frame;
17 end

```

Binary thresholding is applied following frame differencing to distinguish between static and dynamic elements within the frame. The binary image results in pixels set to '1' (white) where differences are detected and '0' (black) elsewhere. Adjusting the binary threshold impacts the sensitivity of motion detection; for example, a lower threshold captures slight differences between frames, enhancing sensitivity, while a higher threshold requires more significant differences for detection, reducing sensitivity. This parameter is pivotal for adjusting the system's response to varying activity levels and environmental conditions.

The window length parameter addresses the temporal aspect of motion detection. By considering multiple frames at once, the system can more reliably determine whether the detected motion is consistent and significant rather than transient or background noise. The algorithm maintains an array of motion values where each new frame's motion result (either '0' or '1') is added, and the oldest value is removed. If the sum of this array exceeds a predefined limit (e.g., five in a ten-frame window), the system confirms the presence of sustained motion. This approach minimizes false positives and enhances the reliability of the system in dynamically changing environments. This pseudocode in Algorithm 2 represents the moving average algorithm used for motion detection.

The proposed algorithm adopts a sliding window approach to aggregate and analyze motion values from consecutive video frames. Denote v_t as the binary motion value at time t , where v_t is 0 (no motion) or 1 (motion detected). These values are stored in an array A , with a predefined window length W , typically set to $W = 10$. Each new frame updates A by appending v_t and removing the oldest value, effectively shifting the array by one position.

Algorithm 2: Moving Average Algorithm for Motion Detection

Data: motionWindow, windowLength, motionData
Result: averageMotion indicating the motion status

```

1 for each new frame do
2   motionWindow[currentIndex] ← Cast motionData to integer;
3   if currentIndex ≥ windowLength then
4     | Reset currentIndex to 0;
5     | Reset motionSum to 0;
6   end
7   for i ← 0 to windowLength-1 do
8     | motionSum ← motionSum + motionWindow[i];
9   end
10  if motionSum ≥ (windowLength/2) then
11    | averageMotion ← 1;
12  else
13    | averageMotion ← 0;
14  end
15 end

```

The mathematical representation of the update mechanism for array A at time t is expressed as:

$$A_t = v_{t-W+1}, v_{t-W+2}, \dots, v_t \quad (1)$$

The decision-making process involves calculating the sum S_t of the elements within A_t . If $S_t > 5$, the algorithm concludes that sustained motion is present within the window, indicating a reliable detection of actual motion:

$$S_t = \sum_{i=t-W+1}^t v_i \quad (2)$$

For example, after processing 10 frames, if the majority of frames within A indicate motion (e.g., $[1, 1, 1, 1, 1, 0, 0, 0, 0]$), then $S_t = 6$ and the algorithm signals motion detection by setting “averageMotion” to 1, suggesting true motion. Conversely, if fewer than half the frames show motion, such as in $[1, 0, 0, 0, 1, 1, 0, 0, 0, 0]$ with $S_t = 3$, “averageMotion” is set to 0, indicating no significant motion is detected. This distinction is crucial for minimizing false positives from transient or incidental movements.

The algorithm’s effectiveness and accuracy are further refined through empirical testing of various binary threshold settings to identify the optimal configuration that accurately reflects true motion in specific environments. This tailored optimization is vital for adapting the system to meet the unique requirements and safety protocols necessary for monitoring scenarios, such as in care homes for the elderly, ensuring both security and comfort.

3.2. Anomaly Detection

Anomaly detection is a technique used to identify deviations from expected patterns, similar to outlier detection in machine learning. This method finds extensive applications in various sectors, such as intrusion detection in network systems, fraud detection in financial sectors, identification of health anomalies in medicine, and more [30]. Even though it is widely used, its integration into smart home technologies is still in its nascent stage. Anomaly detection addresses the challenges associated with the need for labeled data, as it effectively processes unbalanced or singular-class datasets. In this work, the anomaly de-

tection model is designed to identify deviations from typical behavior patterns considered normal for each resident. This model is simpler than deep neural networks, making it easier to train and requiring less data. The main concept behind this model is to establish a normal activity profile for each resident based on their routine daily activities. This profile acts as a baseline for identifying any anomalous behavior.

To implement the anomaly detection, daily activity data were collected for each resident to develop a comprehensive profile of their normal behavior. This involved monitoring various activities throughout the day and capturing patterns of movement and interaction within the care environment. Once this profile is established, the anomaly detection model continuously monitors the incoming data against the established norms. When the system detects activities that deviate significantly from the normal profile, it flags these as anomalies. This real-time detection allows for immediate responses, where the system can generate alerts for caregivers or family members, prompting timely intervention.

This method is particularly effective in environments where individuals have consistent daily routines, making it easier to spot deviations that could indicate potential health issues or emergencies. The implementation of this anomaly detection model not only enhances the safety and well-being of the elderly residents by providing immediate alerts but also contributes to a more personalized approach to care as the system adjusts and learns from each individual's behavioral patterns.

4. Results and Analysis

4.1. Data Used in Experiments

To validate the proposed system, we conducted a pilot study in a simulated elderly care environment using volunteer participants. A total of four volunteers were recruited to perform daily activities and occasional anomaly events across four separate monitored rooms. Each room was equipped with cameras and infrared sensors integrated with our vision-based monitoring system. The volunteers were asked to perform a set of typical activities of daily living (ADL), including sitting, standing, walking, lying down, and moving between rooms, as well as to simulate abnormal scenarios such as prolonged inactivity, sudden falls, and absence from expected zones.

Across the sessions, over 150 h of continuous monitoring footage were recorded and analyzed. Data were collected over multiple sessions to ensure variability in lighting, room conditions, and participant behavior. No personal identifiable information was recorded, as privacy protection was a fundamental design aspect of our system. The monitoring relied solely on anonymized frame differencing outputs, where individuals were represented as motion contours rather than identifiable images.

This dataset allowed us to evaluate the system's ability to: (1) accurately detect motion and non-motion events; (2) identify deviations from normal activity patterns using anomaly detection algorithms; and (3) measure system robustness under different environmental conditions, such as lighting changes or background movements.

Although no formal numerical metrics were computed at this stage, qualitative validation was performed by manually cross-verifying system-generated activity logs against video footage, confirming that detected motion events and room transitions matched expected real-world behaviors.

4.2. System Workflow and Monitoring Process

The basic functionality of the proposed system system is depicted in the flow diagram as shown in Figure 2. The diagram shows a step-by-step process for monitoring, collecting data, and updating activity logs. As shown in the flow diagram, when monitoring begins, a new frame is received and processed for each room (Room 1 to Room 4). For each room,

the system checks the frame for motion (frame differencing) and determines whether there is motion (True or False). This information is added to a rolling window if motion is detected, which likely represents an ongoing log or buffer. Subsequently, the activity log is updated, and the user interface (UI) is updated to reflect the new data.

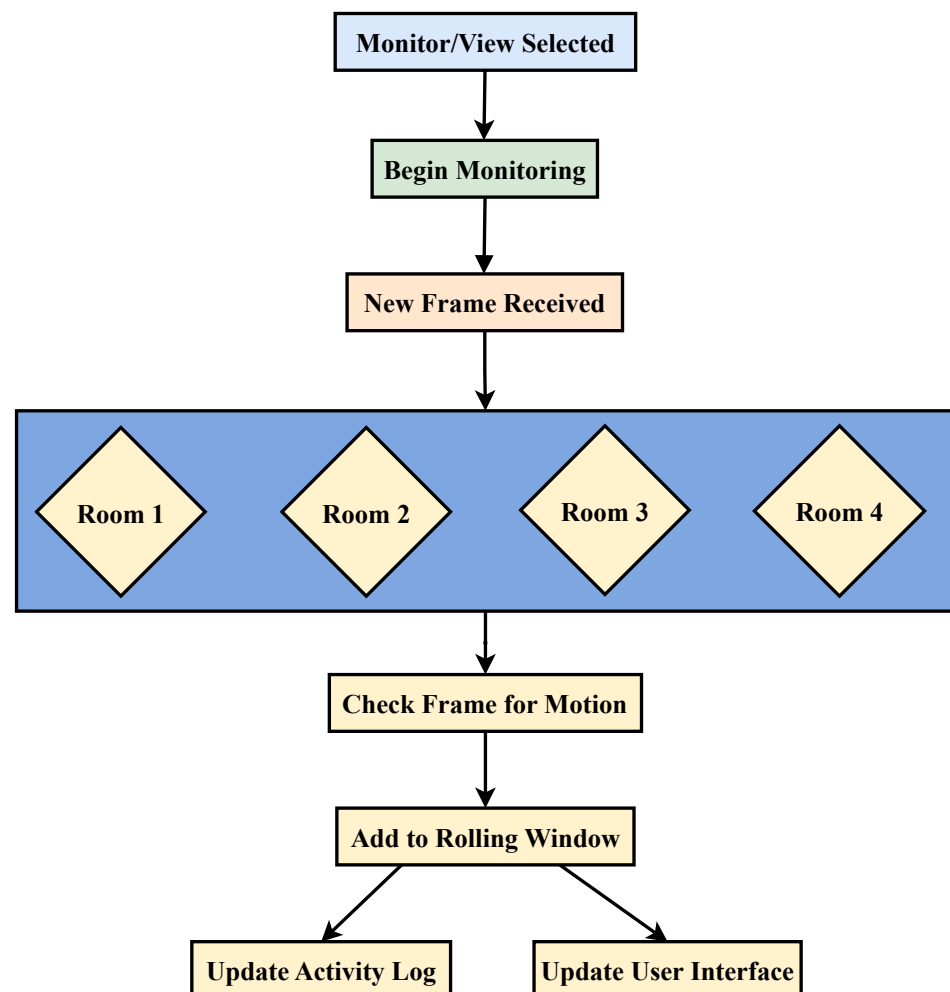
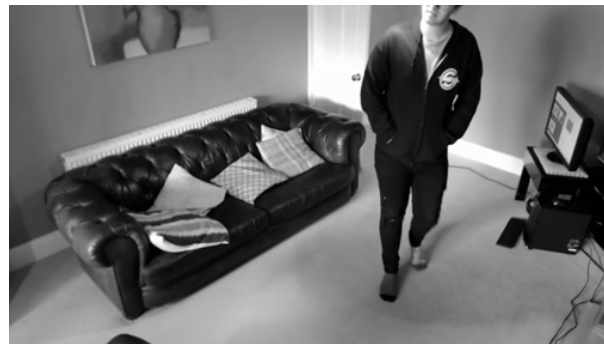


Figure 2. Schematic representation of the system's operational flow, illustrating the process from user-initiated monitoring to motion detection and data logging across multiple rooms.

In a care setting, it is crucial to have monitoring systems that can adapt to different environments and meet the specific needs of the care providers. To enhance its adaptability to various care settings, the system incorporates several adjustable parameters that allow fine-tuning of its operations. One such parameter is the activity update rate, which controls the temporal resolution of activity logging. A lower value means that activity is logged more frequently, capturing more detail about short-duration movements. A higher value would reduce the system's responsiveness but could be beneficial in reducing data storage requirements or minimizing the noise from insignificant movements.

Another critical adjustable feature is the detection threshold. The detection threshold is a crucial parameter for tuning the system's motion detection algorithm. A lower threshold makes the system highly responsive to even slight movements, which could be essential in scenarios where the safety of individuals is a priority. In contrast, a higher threshold reduces sensitivity, which might be suitable in environments where minor movements are not critical to detect or where there is a need to filter out false positives caused by incidental movements.

The detection threshold is used to facilitate frame differencing, where the image is divided into binary classes: pixels with unchanged frame values are set to 0 (black), indicating no motion, and pixels with different frame values are set to 1 (white), indicating motion. This binary segmentation is crucial as it distinctly highlights areas of movement. This parameter can also be used to tweak the motion detection sensitivity (1 being sensitive, 125 being less sensitive). This can be tuned in real-time while the application is running. The image shown in Figure 3 demonstrates frame differencing using a binary threshold value of 3. A lower threshold value is far more sensitive to detecting a difference between the frames, making overall motion detection more sensitive.



(a)



(b)

Figure 3. Illustration of the effect of detection threshold settings on motion detection outputs. (a) The original room environment. (b) Enhanced focus on relevant movements.

For our experiment, a binary threshold value of 16 was used. This value was determined to provide optimal performance for reliable motion detection in the demonstrated environment. It is important to note that this threshold value is not universally applicable; it will need to be adjusted and fine-tuned based on the specific requirements and conditions of each scenario to ensure the system functions effectively.

Lastly, the window length parameter adjusts the breadth of frame consideration. Increasing the window length parameter enhances the algorithm's robustness against false positives by considering a broader set of frames, thus allowing for a more reliable motion assessment. However, this parameter introduces a trade-off between the system's sensitivity and immunity to sporadic non-relevant movements. The array-based implementation facilitates a dynamic motion analysis across consecutive frames, providing a buffer that smooths out anomalies and ensures that detected motion is not an isolated incident but a consistent activity. These parameters collectively ensure that the system can be customized effectively, making it a versatile and reliable tool for enhancing the safety and monitoring capabilities within a smart care home.

4.3. Data Analysis

This subsection describes the data analysis modules of the system. The system provides a MonitorView module that is the heart of the application, encompassing advanced computer vision and motion detection capabilities. Users must first select a residence from the Residence page with a double-click action to access it. Upon selection, the user is directed to the MonitorView interface. Before initiating the monitoring process, the user must configure the system by visiting the settings page to select the appropriate video files intended for the motion detection test. Once the files are chosen, the user can start the motion detection scenario by clicking the green play button within the MonitorView, enabling the system to begin analyzing and detecting motion within the video feed.

In the proposed system, the data collection and analysis capabilities are demonstrated through the ‘Motion Active’ room display within the ‘Overview’ tab, as depicted in Figure 4. This interface provides an immediate and thorough understanding of an occupant’s activities. It employs a donut chart for a lucid representation of time spent in various rooms, offering a quick visual interpretation of the occupant’s spatial movements. The activity log, an integral part of this tab, catalogs detailed room activities alongside precise timestamps. This log is essential for granular analysis in the AnalysisView, where data stored in JSON format can be meticulously examined.

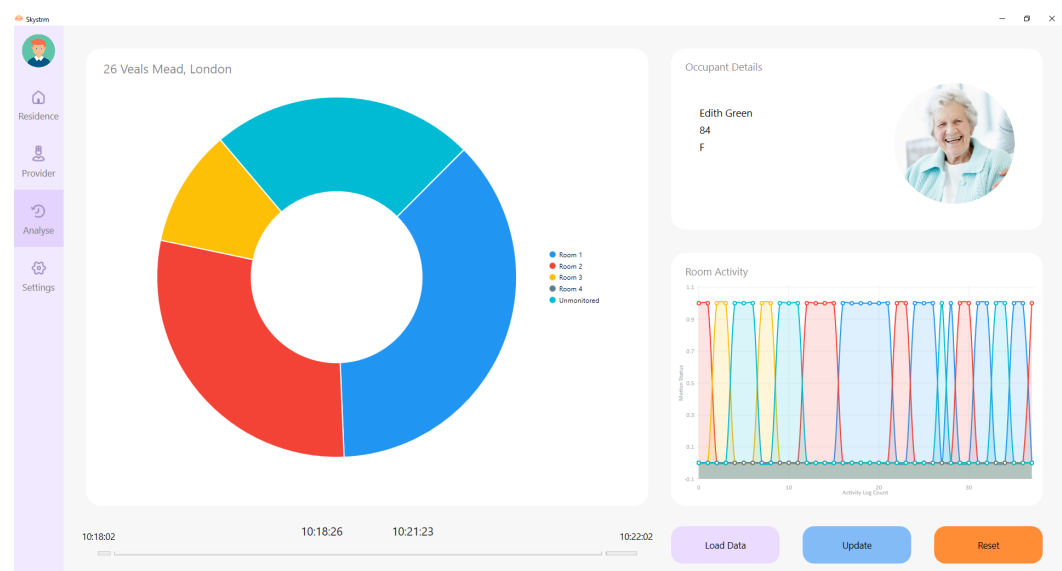


Figure 4. Data analysis of stored JSON file. The diagram shows the activities of one occupant/patient in different rooms.

Additionally, the system can monitor multiple rooms simultaneously via the ‘Rooms’ tab, as shown in Figure 5. Each active room is marked with a distinct green circle, enabling users to quickly identify areas of current activity. The real-time update functionality of these indicators is paramount for ensuring timely and accurate monitoring. This multi-room surveillance feature enhances the system’s utility by providing a comprehensive view of the entire monitored space, which is crucial for complete situational awareness and prompt response to any irregular activities or emergencies.

In deploying and evaluating the proposed system, considerable attention was given to the privacy and anonymity of the real-time data collected from cameras within care homes. The system’s user interface is designed with a selector switch that enables the toggling between the raw camera feed and the frame differencing video output, as shown in Figure 6.

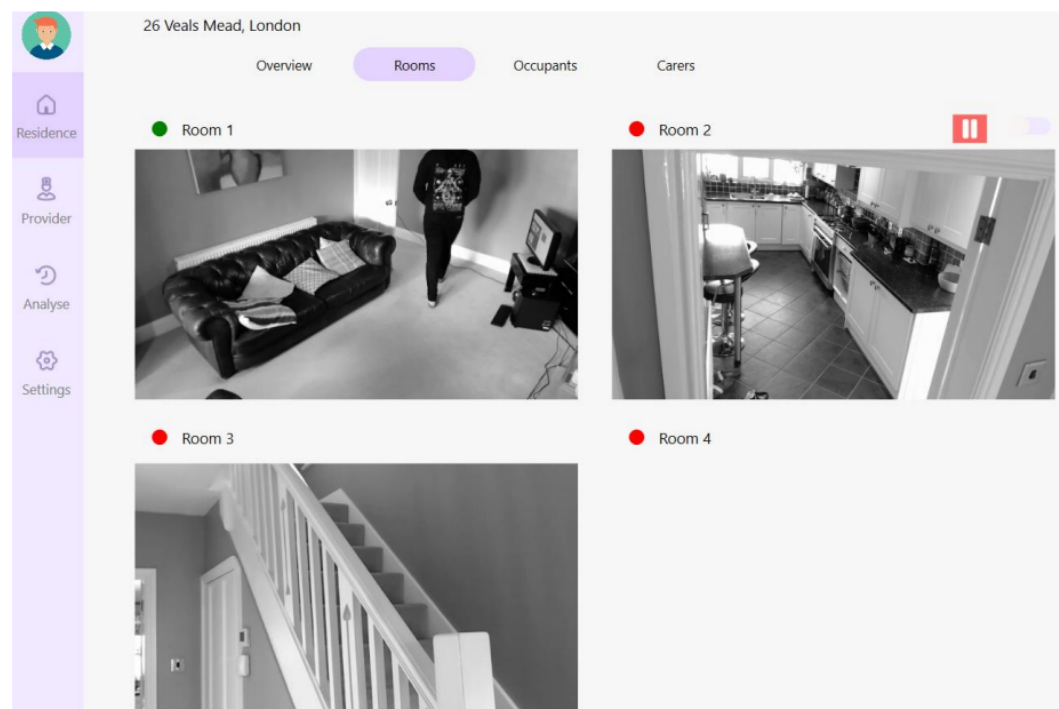


Figure 5. Simultaneous activity monitoring in differing rooms.

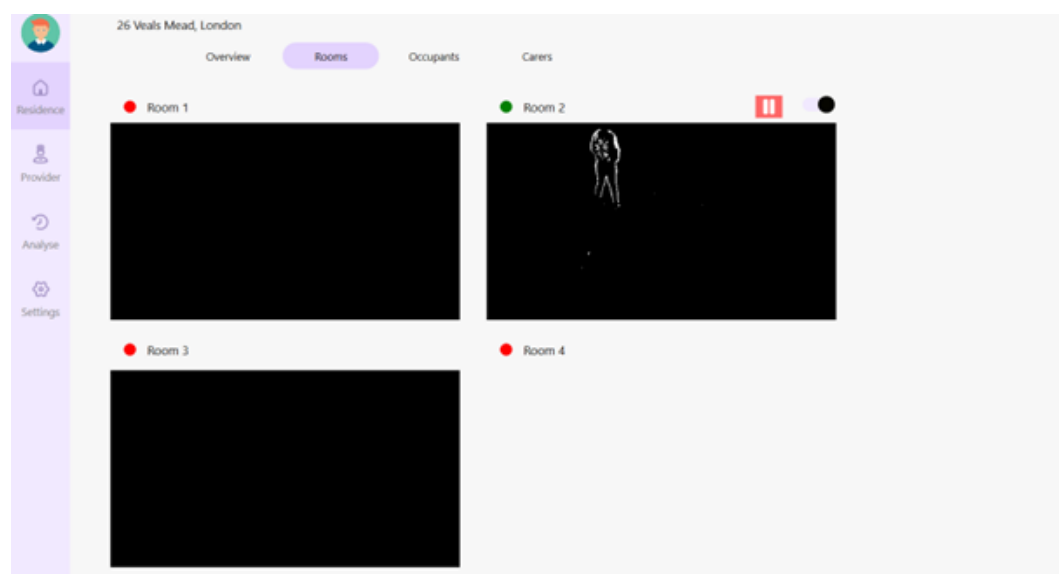


Figure 6. Toggling between the raw camera feed and the frame differencing video.

This feature played a crucial role during the testing phase, allowing us to validate the system's performance without compromising resident privacy. More importantly, this functionality lays the groundwork for future operational use, where any demonstration or review of the recorded footage can be conducted in a manner that maintains the anonymity of individuals. The frame differencing view abstracts the individuals into non-identifiable moving shapes, ensuring that the system can be used in real-world settings without infringing on personal privacy. This approach to data handling not only meets ethical standards but also aligns with legal privacy requirements, underscoring the system's applicability in sensitive environments where the dignity and rights of individuals are paramount.

A vital aspect of the developed system's utility is its sophisticated data visualization capabilities, which are essential in presenting complex data streams in an accessible manner. The system employs a sophisticated user interface for data analysis and visualization,

designed to handle, display, and analyze the data collected during monitoring. Users can load data into the system, which then updates the UI to reflect the information in an interactive format. Features like adjustable sliders allow users to navigate through timestamps, facilitating detailed examination of specific events or periods. The AnalysisView system can reset and clear the UI, providing a fresh start for new data analysis without residual data from previous sessions. The Workflow of the Analysis View is depicted in Figure 7.

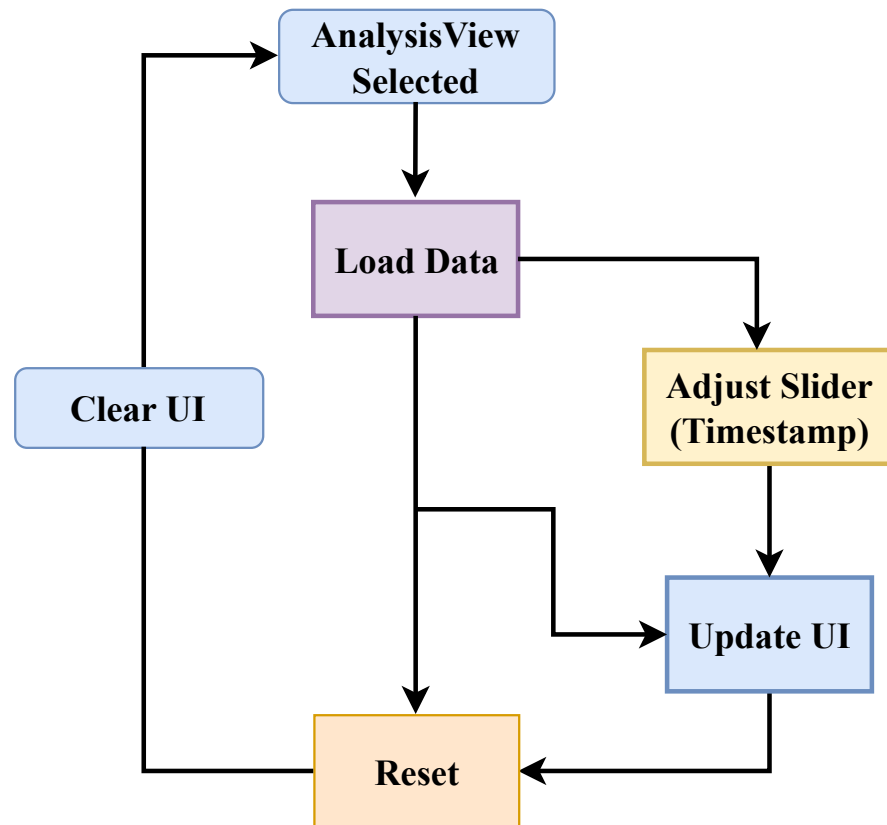


Figure 7. Data Handling and Visualization Workflow. The process begins with loading data into the system, which triggers updates to the user interface to reflect new information.

AnalysisView also supports extensive data handling capabilities. Data collected during monitoring sessions is stored in JSON format, containing detailed logs of activities along with associated metadata like timestamps and sensor readings. This format supports efficient data manipulation and retrieval, which is essential for analyzing long-term trends and immediate concerns in elderly care.

The combination of these algorithms and the AnalysisView interface equips caregivers and health professionals with the tools needed to monitor the well-being of elderly residents effectively, ensuring timely intervention and improved care outcomes. The detailed data flow through the AnalysisView interface and the precise steps involved in motion detection through frame differencing is crucial for understanding how the proposed system enhances elderly care through technology.

One of the primary visualization tools employed is the donut chart, which has proven highly effective for this application. This graphical representation offers an immediate overview of the distribution of an occupant's activities, identifying the rooms frequented over a user-defined period as shown in Figure 8a. Such a visual summary is invaluable for building activity profiles for occupants, particularly those living alone. By observing the proportion of time spent in various areas, caregivers can quickly discern normal

patterns and, crucially, pinpoint any anomalies that may suggest changes in behavior or potential emergencies.

Additionally, the system uses line graphs to present motion data over time, offering a dynamic and temporal view of resident activities. Each entry in the graph signifies a motion event, categorized as either ‘true’ (1) or ‘false’ (0), allowing for a clear depiction of activity levels and movements throughout the care home. A particularly insightful feature of the line graph is its ability to trace the transitions between different rooms, adding a layer of context to the activity data as shown in Figure 8b. This not only aids in the immediate understanding of daily routines but also serves as a reliable record against which any deviations can be measured.

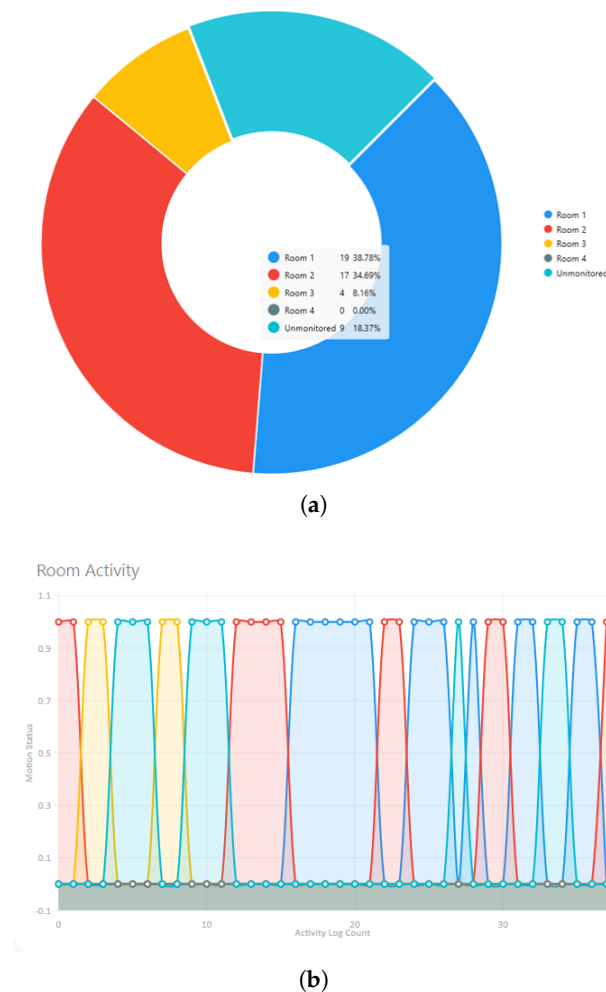


Figure 8. Data visualizations. (a) A donut chart representing the most frequently visited rooms within the last n period, where n is a user-defined parameter. (b) A line graph displaying a distinct record of the occupant's activity for the specified period.

During the testing phase, using pre-recorded scenarios, the line graphs were instrumental in validating the system's logs against expected outcomes. The exact match between the logged data and the footage's manual observations confirmed the motion detection's precision. The congruence between the visualized data and the actual events confirms the accuracy of the system's motion detection capabilities and its potential as a foundation for further development.

These visualization techniques simplify the interpretation of complex data for caregivers and underpin the system's potential for advanced applications, such as predictive analytics and deeper behavior analysis. By employing such intuitive data visualizations,

the system ensures that critical information is conveyed effectively, supporting timely decision-making and enhancing the overall care experience within elderly care homes.

4.4. Comparative Analysis with Existing Methodologies

In the context of existing methodologies for elderly monitoring, many systems utilize complex machine learning models or deep neural networks that often require substantial computational resources and large volumes of labeled training data [11,13,17]. By contrast, the proposed system employs a lightweight frame differencing-based approach combined with anomaly detection, providing reliable motion detection without the need for extensive data preprocessing or model training.

While formal quantitative comparisons based on detection rates or false positive rates were not computed at this stage, internal qualitative validation demonstrated the system's capability to detect motion and room transitions accurately. Compared to vision-based systems such as depth-sensor approaches [16] and privacy-preserving avatar systems [17], our solution achieves a balance between real-time performance, user privacy, and operational simplicity, making it particularly suitable for deployment in resource-constrained elderly care environments. Future work will involve conducting formal benchmarking studies against other state-of-the-art solutions using standardized datasets and metrics.

5. Discussion

Camera-based Human Activity Recognition uses video or images captured by cameras to identify and categorize human activities. Unlike wearable or ambient sensors, which often lack detailed information, camera sensors can provide comprehensive data [31]. They can determine a person's location, environment, facial expressions, poses, gestures, and even mood. The implementation of the vision-based system in elderly care settings represents a critical evolution in monitoring technologies. It blends computer vision with anomaly detection to enhance residents' safety and independence [32]. These systems underscore the potential of integrating sophisticated monitoring solutions into traditional care practices to improve the monitoring of elderly activities.

Preliminary findings from the system's deployment show that it can identify daily activities and detect deviations from established behavioral patterns. These features are essential for anticipating potential emergencies and enhancing the overall management of elderly care. Importantly, the use of simple anomaly detection models instead of complex deep learning-based approaches is a deliberate design choice. In elderly care environments, lightweight algorithms provide critical advantages: they reduce computational demands, minimize system maintenance complexity, improve interpretability for caregivers, and allow real-time operation without sacrificing accuracy. Using simple anomaly detection models in place of more complex machine learning or neural network-based algorithms ensures that the system remains efficient and user-friendly, reducing computational requirements while maintaining a high level of accuracy.

This purposeful simplicity also supports privacy preservation, as it eliminates the need to store or process large volumes of sensitive video data, relying instead on anonymized motion-based outputs. Furthermore, by avoiding deep learning models that require extensive labeled datasets and continuous retraining, the proposed system achieves greater robustness and scalability for practical deployment across multiple care settings.

A critical aspect of the system's functionality is the adjustability of parameters such as the activity update rate, detection threshold, and window length. These parameters are essential as they allow the system to be finely tuned to suit the specific environmental conditions and care needs. For example, adjusting the activity update rate can help manage the volume of data logged, enhancing the focus on significant movements rather than

minor or irrelevant ones. Similarly, the detection threshold is vital for tailoring the system's sensitivity, enabling it to detect subtle movements critical in high-risk situations or to avoid false alarms in less dynamic environments.

Furthermore, the binary threshold and window length adjustments play a significant role in activity recognition accuracy. They ensure that only meaningful activities are recorded and analyzed, reducing the likelihood of false positives and enhancing the system's reliability. Such adaptability improves the system's performance and applicability across various care settings, demonstrating a versatile approach to elderly care. Despite these advancements, the proposed system may face challenges in implementing new technologies in sensitive environments. Variations in individual behaviors and the dynamic nature of care settings can sometimes affect the accuracy of the readings, leading to potential false positives or missed detections. These issues highlight the need for ongoing refinements and algorithm adjustments based on continuous feedback and accumulated data.

Limitations

While the proposed system demonstrates practical applicability and strong performance in simulated care environments, several limitations exist. First, the current validation is based on qualitative analysis using pre-recorded scenarios, and no formal numerical evaluation (e.g., detection accuracy, false positive rate) has yet been conducted using annotated ground truth datasets. Second, the system has not been directly compared against other existing motion detection or fall detection frameworks due to the prototype nature of the work and limited access to standardized datasets. Third, while the algorithm is lightweight and well-suited for privacy-conscious environments, more complex behaviors may not be accurately captured without the integration of advanced learning techniques. Future work will focus on real-world deployment, quantitative benchmarking, and evaluation against state-of-the-art systems.

6. Conclusions and Future Work

This study presents an unobtrusive monitoring system for elderly people in care settings. The system uses advanced computer vision techniques and anomaly detection algorithms to monitor elderly people in care settings. The system tracks daily activities and intelligently analyzes patterns to identify deviations from normal routines. The system logs these events, which can be visualized later for further analysis. The real-time notifications allow immediate intervention when potential health risks are detected, significantly reducing response times and potentially preventing emergencies. The system's effectiveness in a real-world setting has set the stage for further advancements in the field. Future work will focus on refining the accuracy of the algorithms to decrease the likelihood of false positives and increase the system's reliability. Integration with additional sensor data, such as wearable devices or environmental sensors, could provide valuable context for activity recognition and anomaly detection. Furthermore, exploring advanced machine learning techniques, including deep learning and reinforcement learning, may unlock new insights and predictive capabilities within the system.

Author Contributions: Conceptualization, R.U., I.A. and J.V.; methodology, G.E.; software, G.E.; validation, R.U., I.A., J.V. and G.E.; formal analysis, J.V.; investigation R.U., I.A., S.A., A.A. and A.B.; resources, J.V.; data curation, G.E. and J.V.; writing—original draft preparation, R.U., I.A. and S.A.; writing—review and editing, R.U., I.A., S.A., A.A. and A.B.; visualization, R.U., I.A. and S.A.; supervision, J.V.; project administration, J.V. All authors have read and agreed to the published version of the manuscript.

Funding: The European Regional Development Fund (ERDF) and the Welsh Government funded this project. The expert knowledge sharing by CEMET and Skystrm colleagues is also acknowledged.

Institutional Review Board Statement: Ethical review and approval were waived for this study due to the fully anonymized nature of the dataset and the absence of sensitive personal information. All participants provided informed consent for their anonymized data to be used in research, in accordance with our organization's data privacy policies.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author due to privacy concerns.

Conflicts of Interest: Author Justus Vermaak was employed by the company Skystrm Limited. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Tan, M.P. Healthcare for older people in lower and middle income countries. *Age Ageing* **2022**, *51*, afac016. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Saied, I.; Alzaabi, A.; Arslan, T. Unobtrusive Sensors for Synchronous Monitoring of Different Breathing Parameters in Care Environments. *Sensors* **2024**, *24*, 2233. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Maresova, P.; Krejcar, O.; Maskuriy, R.; Bakar, N.A.A.; Selamat, A.; Truhlarova, Z.; Horak, J.; Joukl, M.; Vítkova, L. Challenges and opportunity in mobility among older adults—key determinant identification. *BMC Geriatr.* **2023**, *23*, 447. [\[CrossRef\]](#)
4. Awotunde, J.B.; Ayoade, O.B.; Ajamu, G.J.; AbdulRaheem, M.; Oladipo, I.D. Internet of things and cloud activity monitoring systems for elderly healthcare. In *Internet of Things for Human-Centered Design: Application to Elderly Healthcare*; Springer: Singapore, 2022; pp. 181–207.
5. Hussain, A.; Khan, S.U.; Rida, I.; Khan, N.; Baik, S.W. Human centric attention with deep multiscale feature fusion framework for activity recognition in Internet of Medical Things. *Inf. Fusion* **2024**, *106*, 102211. [\[CrossRef\]](#)
6. Eyiokur, F.I.; Kantarci, A.; Erakın, M.E.; Damer, N.; Ofli, F.; Imran, M.; Križaj, J.; Salah, A.A.; Waibel, A.; Štruc, V.; et al. A survey on computer vision based human analysis in the COVID-19 era. *Image Vis. Comput.* **2023**, *130*, 104610. [\[CrossRef\]](#)
7. Asghar, I.; Ullah, R.; Griffiths, M.G.; Evans, G.; Vermaak, J. Skystrm: An Activity Monitoring System to Support Elderly Independence Through Smart Care Homes. In Proceedings of the 2023 6th Conference on Cloud and Internet of Things (CIoT), Lisbon, Portugal, 20–22 March 2023; pp. 119–126.
8. Bhola, G.; Vishwakarma, D.K. A review of vision-based indoor HAR: State-of-the-art, challenges, and future prospects. *Multimed. Tools Appl.* **2024**, *83*, 1965–2005. [\[CrossRef\]](#)
9. Ullah, R.; Arslan, T. PySpark-based optimization of microwave image reconstruction algorithm for head imaging big data on high-performance computing and Google cloud platform. *Appl. Sci.* **2020**, *10*, 3382. [\[CrossRef\]](#)
10. Ke, S.R.; Thuc, H.L.U.; Lee, Y.J.; Hwang, J.N.; Yoo, J.H.; Choi, K.H. A review on video-based human activity recognition. *Computers* **2013**, *2*, 88–131. [\[CrossRef\]](#)
11. Cristina, S.; Despotovic, V.; Pérez-Rodríguez, R.; Aleksic, S. Audio-and Video-based Human Activity Recognition Systems in Healthcare. *IEEE Access* **2024**, *12*, 8230–8245. [\[CrossRef\]](#)
12. Sinha, K.P.; Kumar, P. Human activity recognition from UAV videos using a novel DMLC-CNN model. *Image Vis. Comput.* **2023**, *134*, 104674. [\[CrossRef\]](#)
13. Karácsony, T.; Jeni, L.A.; De la Torre, F.; Cunha, J.P.S. Deep learning methods for single camera based clinical in-bed movement action recognition. *Image Vis. Comput.* **2024**, *143*, 104928. [\[CrossRef\]](#)
14. Ghali, A.; Cunningham, A.S.; Pridmore, T.P. Object and event recognition for stroke rehabilitation. In Proceedings of the Visual Communications and Image Processing 2003, Lugano, Switzerland, 8–11 July 2003; Volume 5150, pp. 980–989.
15. Gaikwad, S.; Bhatlawande, S.; Shilaskar, S.; Solanke, A. A computer vision-approach for activity recognition and residential monitoring of elderly people. *Med. Nov. Technol. Devices* **2023**, *20*, 100272. [\[CrossRef\]](#)
16. Jalal, A.; Kamal, S.; Kim, D. A depth video sensor-based life-logging human activity recognition system for elderly care in smart indoor environments. *Sensors* **2014**, *14*, 11735–11759. [\[CrossRef\]](#)
17. Wang, C.Y.; Lin, F.S. AI-driven privacy in elderly care: Developing a comprehensive solution for camera-based monitoring of older adults. *Appl. Sci.* **2024**, *14*, 4150. [\[CrossRef\]](#)
18. Gillani, N.; Arslan, T. Intelligent sensing technologies for the diagnosis, monitoring and therapy of alzheimer's disease: A systematic review. *Sensors* **2021**, *21*, 4249. [\[CrossRef\]](#)

19. Majumder, S.; Aghayi, E.; Noferesti, M.; Memarzadeh-Tehran, H.; Mondal, T.; Pang, Z.; Deen, M.J. Smart homes for elderly healthcare—Recent advances and research challenges. *Sensors* **2017**, *17*, 2496. [\[CrossRef\]](#)
20. Jang, I.Y.; Kim, H.R.; Lee, E.; Jung, H.W.; Park, H.; Cheon, S.H.; Lee, Y.S.; Park, Y.R. Impact of a wearable device-based walking programs in rural older adults on physical activity and health outcomes: Cohort study. *JMIR MHealth UHealth* **2018**, *6*, e11335. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Smirnova, E.; Leroux, A.; Cao, Q.; Tabacu, L.; Zipunnikov, V.; Crainiceanu, C.; Urbanek, J.K. The predictive performance of objective measures of physical activity derived from accelerometry data for 5-year all-cause mortality in older adults: National Health and Nutritional Examination Survey 2003–2006. *J. Gerontol. Ser. A* **2020**, *75*, 1779–1785. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Jiang, Y.; Zeng, K.; Yang, R. Wearable device use in older adults associated with physical activity guideline recommendations: Empirical research quantitative. *J. Clin. Nurs.* **2023**, *32*, 6374–6383. [\[CrossRef\]](#)
23. Zhang, S.; Li, Y.; Zhang, S.; Shahabi, F.; Xia, S.; Deng, Y.; Alshurafa, N. Deep learning in human activity recognition with wearable sensors: A review on advances. *Sensors* **2022**, *22*, 1476. [\[CrossRef\]](#)
24. Bhattacharya, D.; Sharma, D.; Kim, W.; Ijaz, M.F.; Singh, P.K. Ensem-HAR: An ensemble deep learning model for smartphone sensor-based human activity recognition for measurement of elderly health monitoring. *Biosensors* **2022**, *12*, 393. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Umer, M.; Alarfaj, A.A.; Alabdulqader, E.A.; Alsubai, S.; Cascone, L.; Narducci, F. Enhancing fall prediction in the elderly people using LBP features and transfer learning model. *Image Vis. Comput.* **2024**, *145*, 104992. [\[CrossRef\]](#)
26. Zi, X.; Chaturvedi, K.; Braytee, A.; Li, J.; Prasad, M. Detecting human falls in poor lighting: Object detection and tracking approach for indoor safety. *Electronics* **2023**, *12*, 1259. [\[CrossRef\]](#)
27. Hussain, A.; Khan, S.U.; Khan, N.; Rida, I.; Alharbi, M.; Baik, S.W. Low-light aware framework for human activity recognition via optimized dual stream parallel network. *Alex. Eng. J.* **2023**, *74*, 569–583. [\[CrossRef\]](#)
28. Teixeira, E.; Fonseca, H.; Diniz-Sousa, F.; Veras, L.; Boppre, G.; Oliveira, J.; Pinto, D.; Alves, A.J.; Barbosa, A.; Mendes, R.; et al. Wearable devices for physical activity and healthcare monitoring in elderly people: A critical review. *Geriatrics* **2021**, *6*, 38. [\[CrossRef\]](#)
29. Yu, Y.; Wang, J.; Li, N. Foreground and background separated image style transfer with a single text condition. *Image Vis. Comput.* **2024**, *143*, 104956. [\[CrossRef\]](#)
30. Galvão, Y.M.; Castro, L.; Ferreira, J.; Neto, F.B.d.L.; Fagundes, R.A.d.A.; Fernandes, B.J. Anomaly Detection in Smart Houses for Healthcare: Recent Advances, and Future Perspectives. *SN Comput. Sci.* **2024**, *5*, 136. [\[CrossRef\]](#)
31. Kolkar, R.; Geetha, V. Issues and challenges in various sensor-based modalities in human activity recognition system. In *Applications of Advanced Computing in Systems: Proceedings of International Conference on Advances in Systems, Control and Computing*; Springer: Singapore, 2021; pp. 171–179.
32. Zhang, Y.; Liang, W.; Yuan, X.; Zhang, S.; Yang, G.; Zeng, Z. Deep Learning Based Abnormal Behavior Detection for Elderly Healthcare Using Consumer Network Cameras. *IEEE Trans. Consum. Electron.* **2023**, *70*, 2414–2422. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.