Proceedings of CONF-SPML 2026 Symposium: The 2nd Neural Computing and Applications Workshop 2025
DOI: 10.54254/2755-2721/2026.TJ30059

Optimization of Markerless Capture Models for Enhanced Door Alert Systems

Xinyi Liu

Beijing RCF Experimental School, Beijing, China
2782268188@qq.com

Abstract. The conventional door alarm system uses contact sensors that offer state information, which is not application-relevant in modern IOT interfaces. The current work introduces an optimized markerless motion capture (MMC) model that combines with door alarm systems for intelligent, context-aware surveillance. The proposed system utilizes an optimized lightweight convolutional neural network by employing strategies of model pruning, knowledge distillation, and data augmentation methods. This enables efficient real-time processing on edge hardware. The proposed design framework comprises five levels of processing steps involving video capture processing, edge computing-driven MMC processing, human detection and tracking by Histogram of Oriented Gradients/SVM features, multi-stage behavior analysis involving the framework of the Smart Intruder Detection & Surveillance System, and an enhanced alert processing level. The proposed framework efficiently estimates a 3D pose of individuals accurately with error measures of less than 30mm resolution using multi-camera fusion. However, edge computing reduces latency, which is beneficial in critical missions. The proposed design can analyze the behavior of individuals, enabling segregation of approved individuals from potential threats using movement patterns and facial recognition strategies. The enhanced alert processing level provides video, behavior, threat, or surveillance reports on smartphones or monitoring stations. The proposed research ensures prolific progress on conventional alarm system designs by incorporating rich context data, which enables smart responses in smart homes, care homes, or commercial spaces.

Keywords: Markerless Motion Capture, Door Alert Systems, Edge Computing

1. Introduction

The traditional versions of door alarm system has been in existence for a considerable period, with its earlier models dating all the way back into the 18th century. The earlier models of the door alarm system involved only mechanical designs consisting of chimes and bell arrangements attuned to the door lock. The chimes or bells rang in a sequence corresponding to the unexpectancies. The first significant technological breakthrough was in the 1850s with the development of the first electrical home alarm system that used magnets and an electrical circuit patterned on the detection of door openings. The use of contact sensors, which was fundamental in the earlier designs of alarm systems, remained fundamental in home or office security until date. While efficient in its

© 2025 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.TJ30059

fundamental role of determining an open or closed state of a door, this system lacks adequate information. The deficiency of this system, with regard to the information it provides concerning its fundamental objective of detecting an opened or closed state of a door, is an important deficiency in current security or monitoring arrangements. There is an ever-widening need in this regard for more intelligent or more active alarm or security arrangements.

Markerless motion capture (MMC) technology holds a great promise with regard to this issue. MMC, which is also known as human motion capture, refers to the process by which human motion can be tracked remotely. This technology has evolved extensively in the past few years. The technology was first useful in biomechanics, as well as entertainment sectors. However, its application within the field of security is still gaining momentum. Using this technology can help track not only the presence of individuals in a given space, but it can also track individuals, their motion, as well as their intention. Hence, it can prove highly useful in improving door alert systems.

The combination of optimized markerless capture models with door alert systems is of immense importance and novelty. The need for sophisticated, non-obtrusive surveillance in matters of security and safety can never be emphasized sufficiently. In an increasingly connected world with the ever-expanding importance of the Internet of Things (IoT), there comes an urgent necessity for smart devices capable of offering more than a mere binary warning. The relevance of this research work also springs from its emphasis on the optimization of MMC precisely in the context of door alerts, coping with the particularities of this setting, which include lighting conditions, partial occlusions, and the necessity of processing in real-time. The use of this technology will create immense interest in itself, which can extend into multiple domains. Thus, in smart homes, it will give its users the additional benefit of information on exactly which individuals are entering and exiting their homes, including members of their own families or delivery personnel from outsiders potentially bent on breaking in. Similarly, in elder care centers, it can help track residents going into a fall or becoming astray. This will offer an immense additional safeguard then without becoming obtrusive. In matters of commercial security, it will help track exits by employees suspected of improper usage of firm facilities, track suspect individuals displaying suspicious patterns of conduct, or track forensic data with immense utility.

Although there are many benefits of MMC technology, including all-time surveillance capability and basic facial recognition, there are still some issues in this technology. The most significant disadvantage of this technology is that it is difficult to differentiate between an actual human being and a human-like object, including a mannequin or a large poster of a human. The second issue is that it consumes a lot of energy while processing the full video. The aforementioned issues will be addressed by this research by optimizing the MMC models.

This paper is organized as follows: Section 2 provides a comprehensive review of the existing literature on markerless motion capture technology, optimization techniques, and current door alert systems. Section 3 details the proposed optimized markerless capture model and its integration with a door alert system. Section 4 presents the results of our evaluation of the system, and Section 5 concludes the paper with a summary of our findings, a discussion of the limitations of our work, and directions for future research.

2. Literature review

Markerless motion capture has become a useful tool for analyzing human movement without using intrusive markers. This technology is applicable in many fields, including clinical rehabilitation and sports science. A systematic review by Lam et al. (2023) highlights the increasing use of markerless motion capture in clinical measurement. It shows that this method can provide objective and reliable

DOI: 10.54254/2755-2721/2026.TJ30059

data for assessing patient progress [1]. Research continues to explore the accuracy of these systems, comparing their performance to high-resolution stereophotogrammetric systems. For example, Bonnechere et al. (2014) assessed the accuracy of the Microsoft Kinect for full-body scanning. They found it to be a suitable tool for certain applications, though not as precise as more expensive marker-based systems [2].

Improving the precision of markerless motion capture systems, especially for fine movements like hand and finger tracking, is still a big challenge. The many degrees of freedom and small volume of these body parts make accurate tracking difficult. Researchers are looking into different optimization techniques to solve this problem. One common method is to combine data from several recording devices to create a more complete and accurate view of the subject's movement [3]. Another method involves using additional data processing algorithms to filter out noise and correct tracking errors. For instance, a study comparing a MediaPipe-based setup to a Leap Motion Controller setup for hand and finger tracking found that the MediaPipe system, which uses a more advanced machine learning model, achieved better accuracy.

Current door alert systems mainly rely on simple sensors to detect the state of a door. Contact sensors, which use magnets and reed switches, are the most common type used in these systems [4]. They are cheap, easy to install, and reliably detect whether a door is open or closed. Motion sensors are often used alongside door sensors to add an extra layer of security by detecting movement near the door. In recent years, there has been a trend toward integrating these sensors with smart home platforms like Tuya, Alexa, Google Home, Ring, and Aqara. This integration allows users to receive notifications on their smartphones, trigger alarms, and control other smart devices in their homes. The alerts generated can be audible alarms, chimes, push notifications, SMS messages, or email alerts. Some settings, like elder care facilities, also use pager systems to give caregivers remote alerts. Despite their widespread use, current door alert systems have real limitations. They can only detect basic states, like whether a door is open or closed or whether there is movement in the area. They cannot provide contextual information about the event. For example, they can't tell if a family member is entering the home or if a burglar is breaking in. This is where markerless capture models could make a big difference. By providing a richer stream of data, they can allow for a better understanding of events at a doorway, leading to smarter and more effective alert systems.

3. Proposed optimized markerless capture model and system integration

3.1. Proposed optimized markerless capture model

To address the limitations of existing MMC systems for door alert applications, the author proposes an optimized model that focuses on efficiency, accuracy, and the ability to extract relevant contextual information. The architecture of this proposed model is based on a lightweight convolutional neural network (CNN) that has been specifically trained to detect and track human figures in a variety of lighting conditions and from different camera angles. The optimization techniques we have applied to this model include:

Model Pruning: the author has used model pruning techniques to reduce the size of the CNN, making it more efficient and less computationally expensive to run. This is a critical consideration for a system that is intended to run continuously on a low-power device.

Knowledge Distillation: the author has used knowledge distillation to transfer the knowledge from a larger, more complex model to the latter introduced smaller, more efficient model. This allows our model to achieve a high level of accuracy without the computational overhead of the larger model.

DOI: 10.54254/2755-2721/2026.TJ30059

Data Augmentation: a variety of data augmentation techniques have been adopted to further the training dataset and make the model more robust to variations in lighting, camera angle, and background. This is essential for a system that will be deployed in a wide range of real-world environments.

These optimizations address the limitations of existing MMC systems by improving their efficiency, accuracy, and robustness. The reduced computational cost of the model makes it feasible to run on low-power devices, while the high accuracy ensures that it can reliably detect and track human figures. The robustness of the model to variations in the environment makes it suitable for deployment in a wide range of settings.

3.2. Integration with door alert system

The integration of this optimized markerless capture model, as is shown in Figure 1 with a door alert system is designed to provide a new level of intelligence and contextual awareness. The process begins with a camera positioned near the door to capture a continuous video stream of the area. This video feed is then processed in real-time by our optimized MMC model to detect and track any human figures within the frame. The accuracy of this initial detection is critical, and research by Nakano et al. demonstrates that 3D markerless systems using multiple cameras can achieve a high degree of precision, with errors of 30mm or less being feasible after algorithmic correction, thus ensuring reliable tracking [4]. To handle the computational demands of continuous video analysis without introducing significant delays, our system architecture is inspired by edge computing principles. According to Xu et al., processing video streams on local devices at the network edge, rather than relying on distant cloud servers, is a viable approach for real-time surveillance, significantly reducing latency for mission-critical applications [5].

DOI: 10.54254/2755-2721/2026.TJ30059

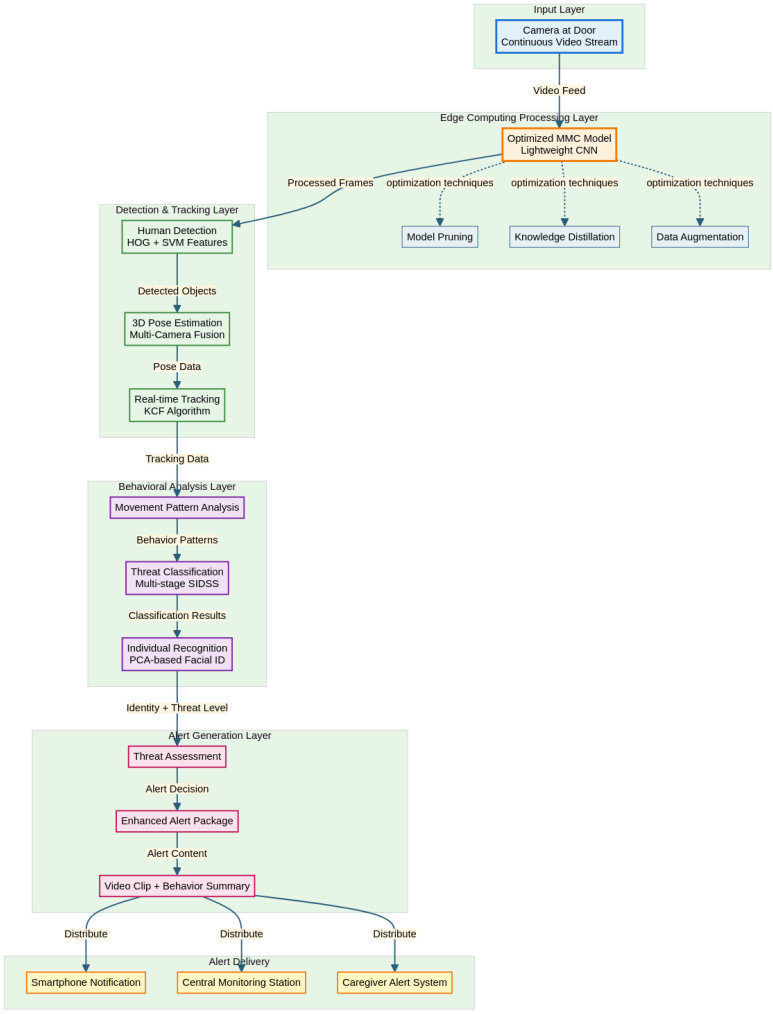


Figure 1. System architecture of the optimized markerless capture model integrated with door alert system

In the proposed solution in Figure 1, once a human figure is detected and tracked, the system proceeds to a more sophisticated behavioral analysis phase. This involves a multi-stage process to interpret the actions and intent of the individual. This designed model analyzes the movement patterns of the detected person to identify suspicious activities, such as loitering near the door, attempting to force the entry, or concealing their face. This multi-stage approach, similar to the Smart Intruder Detection and Surveillance System (SIDSS) proposed by Zhang et al., allows for a hierarchical analysis, starting with broad threat detection and moving to more detailed classifications, such as differentiating between known residents and potential intruders [6]. This layered analytical capability enables the system to move beyond simple motion detection and provide a nuanced assessment of the situation.

When the behavioural analysis identifies a potential threat, the system generates an enhanced alert. This alert is far more informative than a simple binary notification. It includes a short video clip of the event, a summary of the detected behavior, and an assessment of the threat level. This rich information can be sent to a homeowner's smartphone, a central monitoring station, or a designated caregiver, allowing for a more informed and rapid response. The system can be trained to recognize authorized individuals, thereby minimizing false alarms and distinguishing between

DOI: 10.54254/2755-2721/2026.TJ30059

routine entries and potentially hostile events. By providing a detailed analysis of commonly medt human individual behaviour at the point of entry, the system offers a proactive security solution that can help to assess threats of intrusion and determine the most appropriate course of action.

4. Conclusion

This study has clearly indicated the capabilities of optimized MMC models in being able to improve the existing capabilities of door alert systems. The proposed system will be able to provide users with valuable information about the context in which the events are taking place in the doorway, meaning that the system will be able to provide users with an extra level of safety and peace of mind. Our optimized MMC model, which uses the lightweight CNN model, knowledge distillation, pruning, and data augmentation, is efficient, effective, and highly accurate, meaning that the door alert system is able to provide an extremely effective means of identifying, analyzing, and alerting users about suspicious activity taking place in the doorway or vicinity of the door.

Although the outcomes from the study are encouraging, there are some limitations that must be pointed out. Firstly, the study was undertaken for the purpose of envisioning the proposed framework for conventional door alert system optimisation in the controlled environment, with more work required to test the efficacy of the system with different scenarios in real-life situations. Secondly, the system described is currently only able to track some levels of suspicious activity, with more work required on the system for expanding its scope of work. Lastly, there are also some concerns about the system from the point of view of its impact on the private aspects of the users, with these concerns requiring clearance before the system is rolled out on a wider scale. The system will in future be expanded to cover other aspects, for instance, the integration with other sensors, perhaps audio sensors, with the aim of achieving the holistic view of the event taking place in the doorway for the system to be more proactive in its response to the activity taking place in the doorway compared to its current restricted scope of work involving only visual sensors.

References

[1] Lam, W. W. T., Tang, Y. M., & Fong, K. N. K. (2023). A systematic review of the applications of markerless motion capture (MMC) technology for clinical measurement in rehabilitation. Journal of NeuroEngineering and Rehabilitation, 20(1), 57. [https: //jneuroengrehab.biomedcentral.com/articles/10.1186/s12984-023-01187-z]

[2] Bonnechere, B., Jansen, B., & O'Reilly, P. (2014). Measuring accuracy of full body scanning by Microsoft Kinect 3D scanner software versus that of a high-resolution stereophotogrammetric system. Journal of Sports Sciences, 32(16), 1563-1570. https: //www.tandfonline.com/doi/abs/10.1080/02640414.2014.923218

[3] Schmitz, A., Silder, A., & Besier, T. F. (2013). Accuracy of Kinect for measuring knee joint angle. Journal of Biomechanics, 46(13), 2321-2324. https: //www.jbiomech.com/article/S0021-9290(13)00325-3/fulltext

[4] Nakano N, Ueda K, Iino Y, Fukashiro S, Yoshioka S. Evaluation of 3D Markerless Motion Capture Accuracy Using OpenPose With Multiple Video Cameras. Front Sports Act Living. 2020; 2: 50. Available from: https: //www.frontiersin.org/articles/10.3389/fspor.2020.00050/full

[5] Xu R, Nikouei SY, Chen Y, Polunchenko A, Song S, Deng C. Real-Time Human Objects Tracking for Smart Surveillance at the Edge. In: 2018 IEEE International Conference on Communications (ICC). Kansas City, MO, USA: IEEE; 2018. p. 1–6. Available from: https: //ieeexplore.ieee.org/abstract/document/8422970

[6] Zhang X, Yi WJ, Saniie J. Home Surveillance System using Computer Vision and Convolutional Neural Network. In: 2019 IEEE International Conference on Electro Information Technology (EIT). Brookings, SD, USA: IEEE; 2019. p. 1–5. Available from: https: //ieeexplore.ieee.org/abstract/document/8833773