International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0082
e ISSN: 2584-2854
Volume: 04 Issue: 03
March 2026
Page No: 520 - 525

Real-Time Edge-Based Burglary Detection and Automated Alerting Using Deep Learning Framework

Mr. K. Srinivasan[1], Lochan Narayan B[2], Melvin Joel K[3], Sasikumar[4]
[1]Associate Professor – Electronics and Communication Engineering, Sri Sairam Engineering College, Chengalpattu, Tamil Nadu.
[2,3,4] UG Scholar – Electronics and Communication Engineering, Sri Sairam Engineering College, Chengalpattu, Tamil Nadu.
Emails: *srinivasan.ece@sairam.edu.in[1], secl23ec07@sairamtap.edu.in[2], sec22ec119@sairamtap.edu.in[3], sec22ec244@sairamtap.edu.in[4]*

Abstract
Burglary stands out as the major concern in terms of security, especially in India's major cities, where the National Crime Records Bureau reports over 100,000 burglaries. This paper proposes an edge-intelligent system that could potentially deter burglars using a real-time CNN-based deep learning model. The system is divided into two stages. Stage 1 uses a face recognition model to identify authorized individuals; hence, no action is taken if a recognized person is identified. Stage 2 is initiated; this stage detects burglary action using a burglary action detection model. This model is deployed on the edge device to prevent cloud security threats. Once a high confidence level is received indicating a burglary action, the system sends an SMS notification and a video feed using the Twilio REST API. Experimental results show the reliability of the system for real-time residential applications.
Keywords: *CNN, REST API, SMS, Edge-intelligent system.*

1. Introduction
The conventional methods of surveillance, such as observing the environment or using basic motion detection, are no longer sufficient in the modern world's fast-changing security environment, as these methods produce a number of false alarms while failing to read the behavior of the individuals involved. Modern-day security issues call for systems that not only detect the movement of individuals but also evaluate their behavior to differentiate betweennormal and abnormal behavior. This is attributed to the complex nature of crime, which calls for an instantaneous, real-time, and proactive reaction to the issues that may occur. This can be achieved by using artificial intelligence, specifically deep learning technologies such as YOLO (You Only Look Once), to develop intelligent security systems that can be considered the digital version of human security guards. YOLO is a state-of-the-art real-time object detection system for multiple objects in a frame of a video stream, known for its speed and accuracy. Its ability to process images extremely fast allows it to achieve extremely fast detection speed, which is an important aspect in the detection of intruders in real-time video streams. The concept of automated security surveillance systems, which is the least intrusive method of security, will inevitably use technologies that enable the machine to detect individuals or suspicious activities. The use of AI in the system eliminates the basic function of merely recording the activities in the environment; instead, it monitors the activities in real-time through the camera feed to detect intrusions or other activities, such as loitering in the environment and alert the security personnel or other notified individuals in real-time. This minimizes the human effort involved in constantly observing the activities in the environment while also eliminating false alarms resulting from the movements of pets or other insignificant activities in the environment. Furthermore, the system's interaction with the alarm system allows the individual to receive alerts with the corresponding video clips to help in the prompt

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0082
e ISSN: 2584-2854
Volume: 04 Issue: 03
March 2026
Page No: 520 - 525

decision-making process in cases of theft or vandalism.

2. Existing System

The existing system uses an old combination of traditional cameras, simple motion sensors, and infrared detectors. However, these systems are not sufficient for detecting thefts accurately or for stopping burglaries in real-time. Normal CCTV systems just record footage but do not give a proper interpretation of what is actually being shown on their screens. Hence, people have to watch their screens in real-time or go through their recordings when something suspicious happens. Many IP cameras claim to be "smart" as they can detect motion or recognize people. However, that is where their intelligence ends. They can inform you that someone is present or that motion is detected somewhere but they cannot understand the context, behavior, or intent behind that motion. Hence, we can imagine how many false alarms you can receive from pets, shadows, insects, rain on the camera lens, vehicles passing in front of the camera, or even sudden changes in brightness from headlights or sun glare. Even with existing cameras that have object detection built into them, it's still hard to detect behavioral patterns. It doesn't detect someone messing with a lock, pushing against doors, prying with instruments, or attempting to break into a house with metal rods or heavy equipment. Moreover, it doesn't know if it's the homeowner, a family member, a friendly neighbor or if it's actually a burglar. As a result, homeowners receive notifications for someone they know or for a neighbor walking by, which leads to ignoring alerts. Environmental factors like rain, fog, dust, insects, and temperature changes may interfere with motion and IR sensor readings. Motion sensors are useful, but cannot determine if detected motion indicates a red flag or simply indicates normal activity. Therefore, motion sensors have difficulties in differentiating between normal household activity and activity that may indicate suspicious activity. Similarly, IR sensors that detect temperature changes are equally ineffective in confirming whether a person is present or not in a given space due to interference with other warm objects or drafts, leading to false positives. Moreover, conventional security systems don't monitor a series of incidents or events over a period of time. This means they won't notice a person entering the door repeatedly, lingering around with suspicious intentions, or trying to cover up the camera. This is because these systems are mostly isolated and don't integrate well into automation systems or intelligent communication channels. If something out of the ordinary is detected, they will mostly respond by sending an alarm or recording a video on the device, on the cloud, or on a local drive without providing any kind of real-time response or advice. In fact, with a cloud-based model, there can be additional delays because of the transmission and processing of a video, which can be critical at a particular moment. Most conventional systems don't have an automatic response feature either. There is no live video feed to the homeowner, no direct contact with law enforcement, and no activation of local audible alarms. Consequently, the existing system provides a false sense of security. It is unable to identify actions associated with a particular owner, identify genuine patterns of intrusion, or differentiate between normal activities and suspicious activities. This has led to a flood of false alarms, among other issues. More so, significant events that occur outside the pre-defined detection areas may go undetected. As such, it is not reliable, especially without smart alerts, making it impossible for it to be used for the prevention of theft or unauthorized access within smart homes.

3. Proposed System

This system describes an intelligent system for surveillance that employs a deep learning framework for the identification of burglary-related activities around the home entry points. Unlike other systems that utilize motion sensors for the purpose of surveillance, this system employs a fine-tuned deep learning algorithm for the identification of specific burglary-related activities. These include the use of a crowbar for prying the door open, the use of heavy objects for hitting the door, the use of tools for picking the lock, as well as the use of force for manipulating the entry points. The system employs a Raspberry Pi, an edge computing device which is connected to the camera, allowing it to constantly monitor the surroundings of the door. A

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0082
e ISSN: 2584-2854
Volume: 04 Issue: 03
March 2026
Page No: 520 - 525

key feature of this system is that it employs a two-stage workflow for the reduction of false alarms. The first stage begins with face recognition. When the system identifies the person as the homeowner or any of the residents, it remains silent and does not begin processing the possibility of a burglary. However, if the person is unknown or not registered in the system, the process proceeds to the next stage. In the second stage, the burglary action detector using a deep learning framework is activated. This technology monitors the video sequence of frames for any action that may resemble an attempted burglary by the identified person. This action includes the use of too much force on the door frame by the burglar.

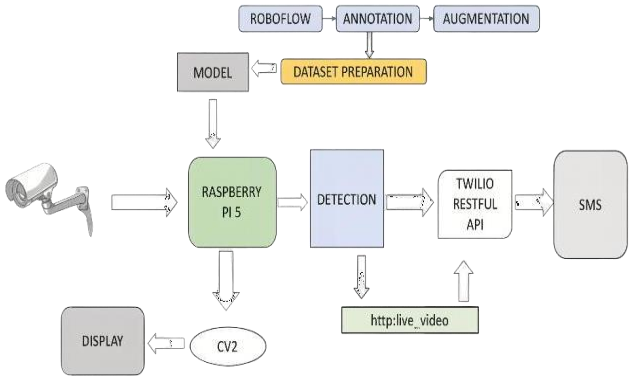

Figure 1 Architecture of the Proposed System

A person using metal tools to pick the lock, someone acting strangely near the lock mechanism and holding a lock pick tool, another person attempting to pry the hinges of the door away from the frame, and someone attempting to force the door open using a hammer.

Table 1 Class Distribution of the Burglary Detection Model

Activity ID	Burglary Action Description
A1	Door prying using crowbar or metal tools
A2	Lock manipulation / lock picking with tools
A3	Striking door using heavy objects (hammer/rod)
A4	Forced door entry using excessive physical force

When any of these burglarious behaviors are detected, the system will activate the alarm and send an instant alert to the home owner using the API in the form of SMS, along with the video feed using via an API in the form of a link, in order for the homeowner to be made aware of the situation in real time and assess the degree of threat to their home. This system, therefore, aims to provide the homeowner with awareness of the situation in order for the homeowner to take the necessary actions in order to further improve the security of their home. As the proposed system is focused on burglarious behaviors rather than general motion, this reduces the rate of false alarms since non-threatening behaviors are filtered out, thereby keeping the homeowner informed of the situation in their home when there is an actual threat. This system, which is driven by the Raspberry Pi, therefore remains in an affordable league and is highly scalable, hence easy to deploy in homes without the need for expensive computing power. This solution presents a novel, instant, and context-based solution to the issue of unauthorized entry into the home and the enhancement of home security using deep learning.

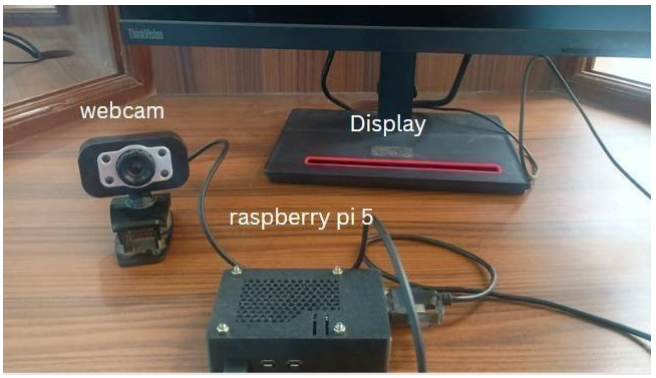

Figure 2 Hardware Setup

4. Methodology

The proposed burglar detection system proposes a two-stage intelligent surveillance system for the detection of burglars by reducing false alarms in the residential area. In the first stage of the system, a face recognition model for the identification of authorized persons is used. A face recognition model for the identification of authorized persons is trained with a dataset of images of authorized persons. During the

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0082
e ISSN: 2584-2854
Volume: 04 Issue: 03
March 2026
Page No: 520 - 525

operation of the system, the video stream from the camera at the entrance is continuously captured by the Raspberry Pi. Then the video frames are fed into the face recognition module to initiate the face recognition model for the identification of the persons at the entrance. The face recognition model recognizes the face in the captured frames, and the face recognition model recognizes by matching the face with the faces of the authorized persons stored in the system. If the face matches the face of the authorized persons with a high level of confidence, the system will be in the Idle Monitoring Phase. If the detected face is not recognized in the database of identities, or if an unidentifiable face is detected, it automatically goes into the second stage, where the burglar detection module is activated. In this module, video images are passed through a pre-trained YOLOv8n object detection model, where it tries to detect signs of burglary, such as suspicious interactions with doors, manipulation of locks, or forced use of objects near doors. The transition from one stage to another is done through a conditional pipeline where the output of the face recognition system acts as a trigger to determine when to use the output of the burglary detection system. Using a two-stage system makes it more efficient because it only uses a deep learning model to detect burglars when an unknown face is detected. In this system, during training, both systems are trained independently. For face recognition, a set of images containing faces of authorized people is collected and preprocessed to detect and extract faces, and a model is trained to learn a discriminative embedding of each face.
These are stored as reference vectors.

Table 2 Dataset Distribution for Burglary Detection Model

Dataset Split	Number of Images	Percentage
Training Set	15,000	75%
Testing Set	2,000	10%
Validation Set	3,000	15%
Total	20,000	100%

Once the inference is complete, the system compares the face vector to determine whether the subject is authorized to enter. To detect burglary cases, the module uses a dataset of images that contain annotations of simulated burglary cases. The module uses a transfer learning approach with the pre-trained

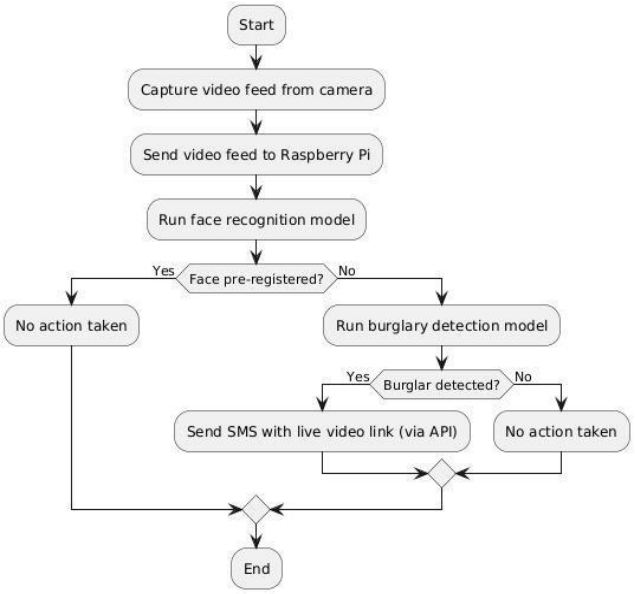

Figure 3 Flow Diagram

YOLOv8n model to train on the dataset for 100 epochs with images of size 640x640 pixels to learn the spatial features related to burglary cases. These features include door handling, tool usage, and attempts to force the door open. The AdamW optimizer is used with an initial learning rate of 0.001 to set the starting step size for the model's weights during the backpropagation algorithm. The choice of the learning rate is important since a high learning rate may cause the model to fail to converge or not converge at all. Adam optimizer also helps the model to generalize by using weight regularization to control the uncontrolled growth of the weights to avoid overfitting. A cosine learning rate schedule is used to maintain the training stability by starting with a high learning rate and reducing it later on. In addition to that, early stopping is used with a patience value of 20 epochs. This means that if the validation metrics do not improve over a period of 20 consecutive epochs, the training process is stopped to prevent overfitting. In order to make the model more robust to real-time surveillance scenarios, several data augmentation techniques are used. These include horizontal flipping with a probability of 0.5,

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0082
e ISSN: 2584-2854
Volume: 04 Issue: 03
March 2026
Page No: 520 - 525

mosaic augmentation with a probability of 0.5, and mixup augmentation with a probability of 0.1. In addition to that, several perturbations are used, including rotation up to 5 degrees, translation up to 5%, and scaling up to 30% helps the model generalize to variations in camera perspective, lighting conditions, and object positions. When the burglary detection model detects suspicious activity during real-time surveillance, the alert is sent to the notification module. It is responsible for sending alerts. Apart from that, the Raspberry Pi can instantly trigger a buzzer connected to the GPIO pins to loudly blast out an alarm sound to wake up people in the surrounding area. The live video stream continues to be hosted on the local streaming server. An alert message is sent to the user through SMS using Twilio. The Pi constructs the alert message and logs in to a mail server through SMTP to send the formatted alert message. The Pi acts as the sender. Twilio's cloud services receive the request, process the request through the mobile network infrastructure, and send the SMS to the registered mobile number. It is a multi-step real-time system to optimize the use of resources for embedded systems like the Raspberry Pi 5.

5. Result and Discussion

Different residential entry scenarios were considered to evaluate the performance of the proposed system in real-time. The training dataset consisted of approximately 20,000 annotated images. The burglary detection model based on the YOLOv8n architecture demonstrated good generalization capability under different lighting conditions and camera viewing angles. During real-time testing, the system achieved an average detection performance ranging between 81% and 88% across all classes. The model obtained an average precision of approximately 83–85%, recall between 82–84%, and an F1-score in the range of 83–85%. These results indicate that the proposed system can effectively identify burglary-related activities while maintaining a relatively low rate of false detections in real-world conditions. The optimal confidence threshold for detection was experimentally determined to be approximately 0.53, which provided a balanced trade-off between precision and recall during inference. This threshold enabled stable detection performance while reducing unnecessary alerts during continuous monitoring. The lightweight architecture of YOLOv8n enabled efficient deployment on the Raspberry Pi 5 platform. The detection model is converted to NCNN for edge devices, and the system achieved an average inference speed of approximately 18 frames per second, allowing continuous real-time monitoring without noticeable delay.

Table 4 Performance Evaluation Metrics

Metric	Obtained Value
Detection confidence	81-88
Precision	83-85
Recall	82-84
F1-Score	87.3
Average Inference Speed	18 FPS

Conclusion

The proposed Real-Time Edge-Based Burglary Detection and Automated Alerting Using Deep Learning Framework effectively detects authorized individuals and intruders by incorporating a YOLOv8n-based action detection model and a face recognition model into a two-stage framework. This article focuses on context awareness by recognizing particular burglary actions, such as forced entry and lock manipulation. The intelligent surveillance system has been evaluated by considering various experiments, and it has been observed that it can detect intruders accurately and efficiently in real-time with minimal latency and a reduced false positive rate. Moreover, it can function efficiently on embedded systems. The integration of real-time SMS notifications and live video streaming has made this intelligent surveillance system highly efficient and effective. This intelligent surveillance system is a highly feasible solution for modern residential security systems since it is highly efficient and cost- effective.

References

[1]. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real- Time Object Detection,” Proceedings of the IEEE Conference on Computer Vision and

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0082
e ISSN: 2584-2854
Volume: 04 Issue: 03
March 2026
Page No: 520 - 525

Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 779–788.

[2]. A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection, arXiv preprint arXiv:2004.10934, 2020.

[3]. G. Jocher, A. Chaurasia, and J. Qiu, "YOLOv8: Ultralytics Real-Time Object Detection," Ultralytics, 2023. [Online].

[4]. R. Girshick, "Fast R-CNN," Proceedings of the IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 2015, pp. 1440– 1448.

[5]. S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real- Time Object Detection with Region Proposal Networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137–1149, Jun. 2017.

[6]. P. Viola and M. Jones, "Rapid Object Detection Using a Boosted Cascade of Simple Features," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Kauai, HI, USA, 2001, pp. 511– 518.

[7]. D. Singh, S. Panthri, and P. Venkateshwari, "Human Body Parts Measurement Using Human Pose Estimation," Proceedings of the 9th International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 2022, pp. 288–292, doi: 10.23919/ INDIACom 54597.2022.9763292.

[8]. Y. Yamaura, Y. Tsuboshita, and T. Onishi, "Head Pose Estimation for an Omnidirectional Camera Using a Convolutional Neural Network," Proceedings of the IEEE 13th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP), Aristi Village, Greece, 2018, pp. 1–5, doi: 10.1109/IVMSPW.2018.8448756.

[9]. C. Shang, C.-Y. Chang, and B. Dande, "Irregularity Detection of Daily Behaviour Patterns Based on Unsupervised Learning," Proceedings of the IEEE International Conference on Consumer Electronics – Taiwan (ICCE- Taiwan), Taoyuan, Taiwan, 2020, pp. 1–2, doi: 10.1109/ICCE-Taiwan49838.2020.9258319.

[10]. S. T. Gnanasekar, S. Yanushkevich, N. J. Van den Hoogen, and T. Trang, "Rodent Tracking and Abnormal Behaviour Classification in Live Video Using Deep Neural Networks," Proceedings of the IEEE Symposium Series on Computational Intelligence (SSCI), Singapore, 2022, doi: 10.1109/ SSCI51031. 2022.10022203.