2025 Volume: 01 Issue: 07 www.ijrpa.com ISSN 2456-9995 Review Article

International Journal Research Publication Analysis

Page: 01-20

STATE-OF-THE-ART TECHNIQUES USED FOR CRIME DETECTION IN INDOOR ENVIRONMENTS: A COMPREHENSIVE REVIEW

***Gloria Ebare. Amadi, Prof. B. B. Olofin, Dr. Okorie Kingsley .M**

Department of Computer Sciences organization, Enugu State University of Science and Technology, Enugu, Nigeria.

Article Received: 03 November 2025

Article Revised: 23 November 2025

Published on: 13 December 2025

***Corresponding Author: Gloria Ebare. Amadi**

Department of Computer Sciences organization, Enugu State University of Science and Technology, Enugu, Nigeria. DOI: https://doi-doi.org/101555/ijrpa.3815

ABSTRACT

The growing need for enhanced security and real-time environmental monitoring in a retail store has led to the development of intelligent surveillance systems. On that basis, this study presents a comprehensive literature review of state-of-the-art techniques used for crime detection in indoor environments. The work examines various intelligent methodologies and tools for surveillance systems on cloud-integrated Internet of Things (IoT) platforms. Each approach is evaluated based on its strengths and weaknesses concerning accuracy and real-time performance. The findings of the study revealed that Artificial Intelligence (AI)-driven models most especially Convolutional Neural Networks (CNNs) offer high accuracy and robust behaviour recognition but require significant computational resources. Then on the other hand, IoT-based systems provide a cost-effective and energy-efficient solution for basic surveillance and environmental sensing but often lack the intelligence to detect complex human behaviour. The study concludes that a hybrid surveillance framework integrating the strengths of both AI and IoT technologies can offer a good solution for intelligent indoor monitoring system in resource-constrained retail environments.

KEYWORDS: Environment Monitoring; Internet of Things (IoT); Artificial Intelligence; CNN; Crime Detection; Indoor Surveillance.

INTRODUCTION

Globally, shoplifting costs billions of dollars annually (Arroyo et al., 2015). Retail theft, sometimes referred to as shoplifting, is a crime that lowers a company's profitability. In order to evade discovery, it entails taking goods from a store without paying for them (Arroyo et al., 2015; Ansari and Singh, 2021). Among the many illegal actions that take place in businesses, shoplifting is one that gets a lot of attention. Additionally, it includes stealing goods from a business by concealing them in one's bag, pockets, or clothes and then leaving the store without paying for them (Muneer et al., 2023). For business owners as well as other organisations like law enforcement, the government, and the legal system, shoplifting is a serious problem.According to a research by the National Association for Shoplifting Prevention, one in eleven people steals something. Furthermore, according to reports, these burglars are only caught once out of every 48 theft cases (Arroyo et al., 2015).

Artificial intelligence-powered solutions for shoplifting detection have attracted a lot of attention due to the shortcomings of conventional security measures (Ansari and Singh, 2023; Kirichenko et al., 2022a). Despite being widely used, video surveillance systems produce enormous volumes of data that security guards are unable to process in real time, opening the door for automated detection systems (Arroyo et al., 2015; Muneer et al., 2023; Pazho et al., 2023).Real-time theft detection, staff warnings, and actionable insights, including pinpointing high-risk locations and peak shoplifting periods, might all be made possible by AI systems coupled with current security infrastructures (Kirichenko et al., 2022b).

Three major obstacles still stand in the way of vision-based shoplifting detection studies, despite advancements in AI-based computer vision. First, research on shoplifting detection is hampered by a scarcity of real-world datasets. The intricacies of real shoplifting episodes are not adequately captured by existing datasets (Ansari and Singh, 2022; Arroyo et al., 2015; Muneer et al., 2023), which frequently rely on data compiled from internet sources or staged situations using actors (Sultani et al., 2018).These datasets' application in real-world contexts is limited since they lack the contextual details unique to each retail venues. Second, progress is further hampered by the intricacy and paucity of tagged data. Because shoplifting incidents are uncommon, unexpected, and challenging to document in real-world settings, supervised learning techniques that depend on labelled data are less successful (Pazho et al., 2023). Third, when identifying shoplifting habits, privacy and bias problems are crucial considerations (Noghre et al., 2023; Hirschorn and Avidan, 2023). Because raw video footage may capture identifying consumer information, using it creates privacy issues. As a result, this research thoroughly examines the implementation of AI-driven intelligent interior surveillance systems. The contributions of the study are:

1. The study discusses relevant related works on indoor environment surveillance for the recognition of security threats considering human behavioural patterns using various intelligent techniques. This review highlights some of the strengths and weaknesses of the techniques applied by these authors
2. The work conceptually discussed some of the tools and techniques that can be used for the enhancement of automated indoor surveillance system such as CCTV and deep learning technologies.
3. Finally, key observations of the reviewed works are identified and discussed further in this work. These observations serve as a guide to highlight on the important factors that affect the technologies and how they can be improved on in future research works

LITERATURE REVIEW

With the rise in security concerns and retail theft, smart surveillance systems have become crucial for monitoring indoor environments, in preventing shoplifting. Traditional security methods like basic alarm systems are ineffective in detecting shoplifting in real time, this has led to integration of artificial intelligence (Sochima et al., 2025), deep learning (Kekong et al., 2019) and computer vision (Ebere et al., 2025) to enhance accuracy and responsiveness in detecting suspicious behaviour.

Afreen et al. (2023) developed a Smart Surveillance System for High-Security Areas (SS-HSA) classified into booster classifier, Naïve bayes classifier, KNN classifier, SVM classifier, random forest classifier, and decision trees to detect and analyze intrusion in an environment. The study analysed the algorithms and discovered that decision trees performed more than the other algorithms with an accuracy of 97% and was very optimal for SS-HSA therefore recommending it for future researches. However, this approach provided high accuracy, its reliance on decision trees limited adaptability to dynamic environments. To address this, Priyanka et al. (2023), adopted a Machine Learning (ML) technique incorporating Inceptionv3 and ResNet50 algorithms for theft detection. Inceptionv3 extracted features from the images at different scales and resolutions and allows network to detect patterns and structures in images that may not be visible with a single set of filters. ResNet a convolutional neural network was used to enhance the overall performance of a neural network. The algorithms were trained on data collected from real-time video from Closed-Circuit Television (CCTV) with a batch size of 32 along with 80 epochs resulting in 99.86% and 98.44% accuracy respectively. The results showed that accuracy with Inceptionv3 was better and can be used for theft detection in supermarkets, highlighted that a surveillance system with fewer human interventions will be a cost-effective product and showcased the effectiveness of deep learning in surveillance.

Despite the success of ML-based techniques, Kaviyaraj et al. (2020) introduced a python language classified into camera coding and main coding to detect robbery/theft motion. The images were captured in camera coding while in main coding, the user controlled the authenticated users and acted as an interface to the user. The system analysed videos images and the images with moving objects; once an object is detected, alerts were automatically sent to the user improving a surveillance system where the CCTV camera and target for small-scale user are developed. The method detected motion in a live stream environment and generated alerts by sending SMS to the user. The system was tested with image sequences and human activities but was limited to a low field. The study recommends that future researches should be done in high fields enhancing user interaction and sequence alignment in the training phase thereby reducing the rate of shoplifting. To overcome scalability issues, Munagekar (2018), adopted a canny edge detection algorithm categorized into grabbing frame, edge pixel processing and theft detection and alert to detect shoplifting in a diamond factory. The algorithm underwent various stages in order to detect intrusion which includes: smoothing, finding gradient, double thresholding and edge detection. The algorithm detected an intrusion and raised an alarm to the owner before sending them the theft images. It was used to preserve properties of an image for image processing and image edge calculation and criteria. The study discovered that the proposed algorithm detects theft swiftly and automatically and most importantly simple to use.

However, datasets and real-world testing limits surveillance accuracy, Reid et al. (2021), developed a dataset of signal attributes based on real-world shoplifting videos using different machine learning models like Support Vector Machine (SVM), K-Nearest Neighbour (KNN), Adaboost to predict shoplifting and the results showed that KNN performed better than the other algorithms with an accuracy of 90.22%. The study encountered certain challenges such as the dataset used to test the method was quite small, some of the videos were cropped and did not show when a customer walks in and out of a shop making it difficult to detect any of the social signal attributes. The study further recommends that future datasets should remained uncropped and more than one human labeller should be used to improve the integrity of the datasets in order to detect shoplifting in an environment. Sai et al., (2024), proposed an ESP32-CAM classified into a Passive Infrared (PIR) sensor and an Arduino Uno which detects someone entering and sends notification to the individual through GSM module. The Arduino, PIR sensor acquires data, a single chip microcomputer controls the GSM module while the ESP-CAM displays the information. The data were trained and the result was in form of SMS that was sent to a registered number. The registered mobile got notified when an unknown person enters the room. The study discovered that the proposed method was an effective and low-cost technique for indoor monitoring which can be integrated with colourful microcontrollers and microprocessors making it suitable for different operations and also a smart security surveillance system involving IoT to enhance security and improve the accuracy of security systems by analysing data and predicting potential security breaches before they occur thereby recommending it for future researches. Similarly, Shaikh et al, (2024) adopted a Connected Environmental Monitoring System classified into 4 layers (sensing, communication, cloud and user interface layer). The layers collected environmental data, fastens the data transmission, analyze data and relate with the system to visualize the collected data. The outcome from the study indicates that the system demonstrated the effectiveness in real time monitoring of an environment. It was further recommended that more researches should be done on indoor environment with additional sensors to improved insights. Sivakumar et al, (2022), adopted a Computer Vision-Based Roadside Occupation Surveillance System (CVROSS) to pre-processed collected data by analysing it through the use of Application Programming Interfaces (APIs). The CVROSS offer clients different types of data perception where aimage was used to introduce continuous traffic circumstances out and about inside the framework time which gave customers a view of the areas of interest. The study discussed the aim of CVROSS concerning reports about street utilization by different types of vehicles generated for street clients, operations organizations, and the public to better understand traffic conditions in the areas under observation during a given time period. Addressing the gap, Jimenez et al, (2022), analysed the accuracy of the different sensors commonly used on Smart Environment Monitoring (SEM) for living environments. The study classified the sensors based on temperature, relative humidity and air quality. The results from the analysis conducted on the sensors showed that the most commonly used sensor for measuring temperature was the DHT22 which uses a thermistor to measure the temperature and a capacitive sensor for relative humidity, also, the accuracy of the air quality was measured on Carbon Monoxide (CO), Carbon Dioxide (CO_2), Volatile Organic Compounds (VOC) and Ozone (O_3). The study overlooked the accuracy of the sensors thereby indulging more researches to be done on the sensors to enhance indoor environment monitoring. Aligning with Prasad et al. (2024) adopting Internet of Things (IOT) as an environmental monitoring system classified into sensor nodes, wireless communication to collects information from surrounding environmental condition and sends the data wirelessly using microprocessors and ESP826 Wi-fi module to the server. The study discovered that the system connects wirelessly to gateway via Bluetooth alongside different parameters, and was a success. The system was flexible, scalable and adapted for both indoor and outdoor use giving it an edge over others. The study recommends introducing several machine learning techniques that may improve the approach.

While IOT-based environmental monitoring improved scalability and adaptability, real-time air pollution control remained a concern. Zulkifli et al. (2020), adopted Rapid Application Development (RAD) to control air pollution. The study conducted 3 experiments which is the connectivity test of sensors with Arduino Uno R3, network connection of GPRS GSM SIM900A with cloud storage and web-site design to generate PDF report from web site and integrate test of cloud storage between web sites and android applications. The study provided a real-time web-based cloud application which allow the DoE to monitor, update and display the air quality data on industrial site. The study further recommended that future research should be done to improve the quality of data transmitted and display on the web-based and android application in real time. Zafar et al. (2018), designed an android application using a Google App-Inventor Integrated Development Environment (IDE) and Java programming language which communicates with the microcontroller through Thing Speak cloud. The android application monitors the environment where the hardware was deployed using a smart phone to implement a home automation system so that the monitored values of temperature and humidity can be used to trigger actions and control the devices for heating or cooling via the mobile application. The study emphasis on the environmental monitoring system for real-time monitoring of temperature and humidity of surrounding environment but recommended that future research should be done on indoor environment monitoring. Addressing security concerns, Ifkat et al. (2023), introduced a Haar cascade algorithm to detect an image and matches the face with existing database to recognize if the face matches and sends alerts. The study demonstrated that face recognition-based human or criminal detection systems are safe. The study further employed raspberry Pi 3 to extract image from its surroundings to match the image from the stored database. It was discovered that the proposed system was used as a security surveillance system which was not only used for criminal detection but also for attendance monitoring, home security, business and car parks and its recognition can be increased using raspberry Pi Infra-red camera module.

For enhanced security in distress situations, Lohith et al. (2021), adopted a smart Closed-circuit Television (CCTV) for monitoring features, noise detection and face identification. The proposed system monitors, identify a person, detect noise, in and out detection and matches it with existing data in order to detect any fraudulent acts or intrusion. The system was used for technological advancement because of its high computing power with a low capacity, it further recommends a high-performance device, additional features like incident, fire and lethal weapon detection for a smart surveillance monitoring environment. This aligns with Abhijith et al. (2024), presented a study on IoT surveillance for real-time distress and fire detection. The proposed method detects both fire and human screams offering an innovative safety in an environment using microphone and sound processing algorithms. A Bluetooth module was adopted which enables interaction between users and robot to receive commands from a smartphone or computer thereby controlling the movements wirelessly. The study suggested that using machine learning and computer vision algorithms can differentiate between humans, vehicles, animals and objects and can also detect specific events like intrusion and loitering.

Shaikh et al. (2024), further contributed to environmental monitoring by designing and developing an environment using the NodeMCU ESP8266 and a variety of sensors including the MQ-135 for air quality, the DHT22 for temperature and humidity, the BMP280 for atmospheric pressure and altitude, and a sound sensor for detecting noise levels. The data from these sensors were transmitted to the ThinkSpeak cloud platform for visualizing in real time, analysed and stored for future use. The system served as a monitoring tool and also as a step to building a smarter and responsive environment that can manage environmental risks more effectively. Zahari and Zaaba, (2017), introduced an Intelligent Responsive Indoor System (IRiS) classified into recognition module, data training module and behaviour module to detect shoplifting intentions which had the capabilities to recognize the face of customers that enters a shop, upload the images of potential shoplifters, track the customers with more than one Closed circuit Television (CCTV) and generate an alarm whenever a shoplifter is identified using video cameras and computers. This method encountered several challenges ranging from high complexity of image processing, the accuracy issues and false alarm rates which are expected to improve from time to time. The study recommends improvement on

IRiS in future researches on integration of emotions, movement and gesture in order to refine the identification of potentials shoplifters. The summary of the literatures reviewed in this study are presented in Table 1.

Table 1: Summary of Literatures.

Author (Year)	Technique	Strengths	Weaknesses
Hamdy et al., (2022)	CNN (Deep Learning) for real-time shoplifting detection	High accuracy (99.86%) Real-time detection	Requires high computational power Needs large labeled datasets
Chandini et al., (2022)	IoT + AI + Cloud (ESP32-CAM, ThingSpeak)	Real-time notifications Simple integration with cloud platforms	Depends on stable internet Limited edge processing capabilities
Kumar et al., (2022)	Arduino Uno + ESP8266 + Sensors (IR, MQ135) + Blynk IoT	Energy-efficient and scalable Useful for environmental monitoring	Lacks visual surveillance Not effective for complex human behaviours
Rashid et al., (2021)	IoT-based real-time video surveillance (ESP32-CAM with app)	Lightweight and portable Live video streaming to mobile	Vulnerable to network disruptions Limited by onboard memory and processing
Deepthi et al., (2020)	AI-enhanced shoplifting detection system using CCTV	Integrates with existing infrastructure	Limited in detecting non-visual or deceptive shoplifting
Sharma et al., (2020)	CNN with Inceptionv3 and ResNet50 for suspicious activity detection	Suitable for real-time surveillance	Large model size High latency on low-end devices

Ali et al., (2019)	Gesture recognition + surveillance camera	Adds context to intent Useful for posture-based anomaly detection	Prone to false alarms in crowded environments

CONCEPT OF INTELLIGENT SURVEILLANCE SYSTEM

Both computer-based and tiny device-based methods are used in the intelligent video surveillance system to send out alerts when an anomalous scenario is captured on camera. The computer-based intelligent video surveillance system has been investigated and performs a variety of detections. Nevertheless, they are not the best for usage in real-world settings because to their high power consumption, installation and maintenance expenses, and potential for personal information leakage. Multi-camera object tracking is a prime example of a traditional computer-based intelligent video surveillance system. Digital Signal Processor (DSP)-based IP cameras input footage, which are then encoded using the Audio Video Coding Standard (AVS) and sent to the IP network using Real-Time Streaming Protocol (RTSP) (Kim et al., 2021).According to Zhang et al. (2012), the IP network employs GPU to reduce processing time in order to do distributed processing.

The fall detection system based on body contours is the second system. This system uses real-time video input from cameras to enable computers to identify falls and transmit the data to a hospital server via its network for storage and monitoring. Aspect ratios and tilt angle properties of a human body contour are used by the fall detection system to determine the fall with less processing, and the Gaussian mixture model (GMM) in the input images is used to recognise objects (Lin et al., 2013). Each camera's input footage is sent to the central server via the computer-based intelligent video surveillance system.Intruder, fire, loitering, and fall detection are just a few of the detections that are carried out in tandem by the central server. The intruder detection system using Raspberry Pi and Arduino is a prime example of a traditional intelligent surveillance system built on tiny devices. The system uses MOG2 to identify movements in input images and uses the sizes of the images to identify human beings. The system uses Fisherface-based face recognition to identify intruders and Haar-like characteristics to identify faces. A relevant user receives an email notice when an intruder is detected, and they may watch the video remotely via a web interface (Tsourma and Dasygenis, 2016).

CCTVTechnology and Crime Prevention

Increased formal monitoring levels inside a target region are a component of the Situational Crime Prevention (SCP) technique known as CCTV (Cornish and Clarke, 2003; Welsh and Farrington, 2009a). By decreasing the quantity of criminal possibilities and raising the perceived risk of offending through physical environment alteration, SCP aims to reduce crime (Clarke, 1995). The rational choice viewpoint, which views crime as "purposive behaviour designed to meet the offender's commonplace needs," is the primary foundation for situational prevention of crime (Clarke, 1997). According to the rational choice approach, criminals take into account a number of "choice structuring properties," such as the possible benefits and inherent dangers associated with committing a certain crime.CCTV's main goal is thought to be to set off a perceptual mechanism that influences an offender's decision to structure their property in a way that discourages them from committing crimes (Ratcliffe, 2006).

Figure 1: CCTV footage of Supermarket. (Source: Grok)

According to study findings published in the literature, the main expected advantage of CCTV is the reduction of crime. Most assessments examine the effectiveness of CCTV by comparing the crime rate before and after camera installation. CCTV can prevent crimes in different ways, even though this study goal appears to emphasise deterrent effects (Piza et al., 2014a) (Welsh and Farrington, 2009b). Researchers have shown that possible mechanisms of CCTV-generated crime reduction include enhanced citizen awareness, publicity, greater natural monitoring, and higher offender apprehension (Gill and Spriggs, 2005).Additionally, CCTV can help police after crimes are committed, particularly by offering visual evidence for use in criminal investigations (Ashby, 2017), enhancing personnel's response to emergencies (Ratcliffe, 2006), and obtaining early guilty pleas from criminals (Owen et al., 2006). We must also recognise that CCTV may lead to more crimes being reported because it can identify crimes that would not have been reported to the police otherwise (Winge and Knutsson, 2003). Additionally, CCTV may increase public vulnerability by giving people a false sense of security, which leads them to become less vigilant or cease taking precautions in public places (Armitage et al., 1999).

In a more recent analysis of seven randomised and natural CCTV tests, Alexandrie (2017) found that although parking lots and suburban subway stations showed no change in crime, public streets and urban subway stations saw decreases of between 24% and 28%. Welsh and Farrington's (2009a) results were somewhat different from Alexandrie's (2017) findings. This discrepancy may be explained by smaller impact sizes linked to quasi-experiments, different research sites (i.e., nations), and variable integration with police procedures as contextual factors. Alexandrie's (2017) claim that integration with police activities may decide the impacts of CCTV is supported by recent study findings (La-Vigne et al., 2011; Piza et al., 2015; Piza et al., 2014b). However, Alexandrie (2017) only cited a tiny number of research, which only makes up a small percentage of the body of information regarding CCTV.

Deep Learning for Crime Detection in Indoor Environment

The technology is intended to offer a comprehensive and efficient method of identifying crimes in crowd monitoring and real-time surveillance. It promotes community engagement and improves public safety by utilising deep learning algorithms. This is a summary of the intelligent system's main elements and characteristics (Gunjal et al., 2024).

Real-Time Surveillance Component

CCTV equipment will be positioned thoughtfully across the infrastructure, especially at transport hubs, shopping centres, and key infrastructure zones. The real-time observation component relies on deep learning algorithms to determine the location and behaviours of individuals. This makes it possible for the system to recognise people who are displaying suspicious behaviour, brandishing hazardous weapons, or acting aggressively. When the framework notices what may be a criminal movement, such as someone taking up a weapon, it instantly delivers a warning. To inspire a response, this notice is sent to key law enforcement agencies or security professionals. The framework incorporates protections such as the privacy of those who are not involved in unlawful activity. It complies with stringent protective regulations to guarantee appropriate use.

Crowd Monitoring Component

Users are urged to submit video footage of busy places and public events to improve the system's performance. The information is safely monitored and its legitimacy is confirmed. Sophisticated video analysis methods powered by deep learning models are used to analyse and assess user-generated video footage. This includes identifying events, detecting anomalies, and detecting objects. By actively participating in crowd monitoring, users assist ensure public safety. Their assistance is crucial to increasing the scope of the monitoring and raising the situational awareness of the community. When the crowd monitoring system finds potentially illegal activities in the submitted footage, it sends out alerts.

Deep Learning Techniques for Indoor Surveillance

This section presents the various deep learning-based algorithms that can be used for the surveillance of an indoor environment such as Convolutional Neural Network (CNN) and You Only Look Once (YOLO).

CNN

One type of deep neural network that is most frequently used in deep learning is the Convolutional Neural Network (CNN/ConvNet), which is seen in Figure 2. CNN does not involve matrix multiplications, which is what we typically think of when we think of a neural network. It employs a unique method known as convolution. Convolution, as used in mathematics, is a mathematical operation on two functions that yields a third function that describes how one function's form is altered by the other. In contrast to a greyscale image, which contains just one plane, an RGB image is only a matrix of pixel values with three planes. The Convolved Feature's spatial size is decreased by the Pooling layer.By lowering the dimensions, this will lower the amount of processing power needed to process the data. Average pooling and maximum pooling are the two forms of pooling.

YOLO

YOLO may be a powerful deep learning system used to locate protests in real time from images and videos. The input image (aimage or video frame of a crime scene) is divided into a grid of squares by YOLOv3. A number of bounding boxes are predicted by each grid cell around items it finds. To accommodate diverse item dimensions, these pre-defined boxes, also known as anchor boxescome in a range of sizes. Every bounding box is given a confidence score by YOLOv3. This score shows how accurate the anticipated size and placement are, as well as how convinced the model is that the box contains an object.

Additionally, the model forecasts the likelihood of every item type inside the bounding box. It can distinguish between a person, a weapon, or another pertinent thing in the scene thanks to this.Although YOLO employs the straightforward CNN architecture depicted in Figure 3 for object identification and image classification, it lacks its own architecture.

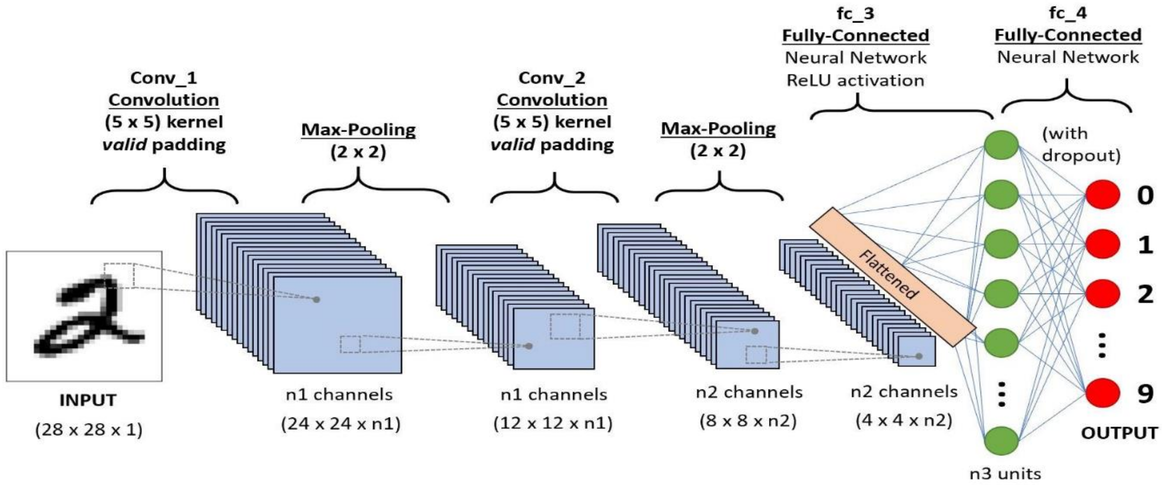

Figure 2: Architecture of CNN Algorithm. (Ratan, 2024)

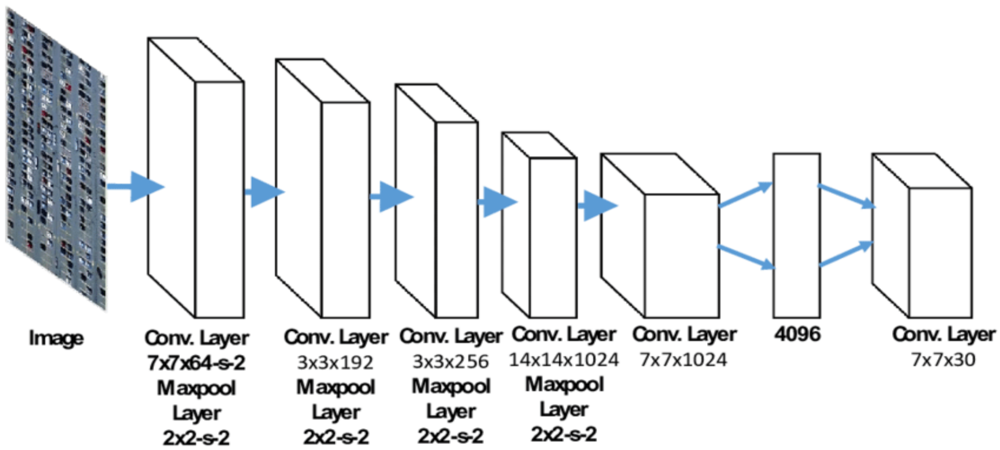

Figure 3: Architecture of YOLO Algorithm. (Malaainine et al., 2021)

RESEARCH DISCUSSION

The study intensively reviewed the technology of indoor automated surveillance using various AI-driven techniques. Here are some observations derived from the review on indoor environmental monitoring:

1. The review revealed that deep learning models like CNNs and transfer learning algorithms like Inceptionv3 and ResNet50, have proven highly effective in shoplifting

detection. These architectures are capable of extracting deep and complex features from surveillance footage, allowing them to achieve remarkably high accuracies.

2. Despite the dominance of deep learning, traditional machine learning algorithms continue to play a significant role in surveillance applications. ML classifiers have been employed in various studies and have demonstrated competitive performance. These models are often favoured for their simplicity, interpretability and lower computational requirements, making them ideal for systems with limited processing power.
3. Another major trend is the integration of Internet of Things (IoT) technologies, which enhances the responsiveness and flexibility of surveillance systems. Devices like the ESP32-CAM and Arduino Uno are frequently used to create low-cost, real-time surveillance solutions. These systems can transmit images or alert notifications instantly, often over Wi-Fi or GSM networks, and are especially useful in resource-constrained environments. Their ability to combine video monitoring with sensor data (e.g., motion, gas, or smoke detection) makes them versatile tools for both security and environmental monitoring.
4. Cloud-based platforms such as ThingSpeak and Blynk have also been widely adopted to support remote access and real-time alerts which enable centralized monitoring and control from mobile devices or computers, making surveillance systems more scalable and user-friendly.
5. A notable limitation observed across many studies is the use of non-diverse or limited datasets where several models were trained and tested on cropped, synthetic, or small-scale datasets, which may not reflect real-life complexities.
6. CCTV and AI-based systems have shown high effectiveness in detecting visible crimes like shoplifting, robbery or assault. However, they are less effective against crimes like fraud or vandalism without supplementary intelligence which means that while vision-based systems excel at behaviour recognition, they may require integration with other data sources for broader crime detection capabilities.

Finally, this review has resolved that several challenges remain unresolved such as privacy concerns related to constant monitoring, the high costs associated with traditional CCTV infrastructure and the tendency of AI models to produce false alarms. Additionally, ensuring system scalability and adaptability across different environmental conditions continues to be a critical issue in the deployment of intelligent surveillance systems.

CONCLUSION

The review explored a range of intelligent surveillance systems combining emerging technologies such as Artificial Intelligence (AI), and Internet of Things (IoT).The review highlights that AI-powered video surveillancecurrently offers the most effective means for detecting shoplifting due to its ability to learn and recognize complex patterns of human behaviour. Several research works were reviewed comprehensively where each applied different techniques to achieve either shoplifting detection, indoor environmental monitoring or both.The integration of cloud platforms provided real-time monitoring and alert capabilities, making these systems more user-friendly and suitable for remote access. However, their performance is often limited by internet dependency and data latency.Despite their differences, all the reviewed systems contribute valuable insights into the development of smart surveillance technologies tailored for retail safety and environmental awareness.

Meanwhile according to the study, IoT-based systems provide a cost-effective and scalable alternative for environmental monitoring and basic motion detection, but they fall short in accurately identifying theft or human behaviour anomalies without AI support.An ideal intelligent surveillance solution for shoplifting detection and indoor monitoring should therefore combine the strengths of both AI and IoT. Future research should focus on hybrid models that are both computationally efficient and behaviourally intelligent, enabling real-time, autonomous detection and response systems suited for resource-constrained environments like small retail stores.

REFERENCES

1. Abhijith, R., Bhujarkar, O. N., Ms. Spoorthy, M. U., Karan, S., & Sudha, P. N. (2024). Literature survey on IoT surveillance for real-time distress & fire detection. International Advanced Research Journal in Science, Engineering and Technology, 11(12). https://doi.org/10.17148/IARJSET.2024.111246
2. Afreen, H., Kashif, M., Shaheen, Q., Alfaifi, Y. H., & Ayaz, M. (2023). IoT-based smart surveillance system for high-security areas. Applied Sciences, 13(15), 8936. https://doi.org/10.3390/app13158936
3. Alexandrie, G. (2017). Surveillance cameras and crime: A review of randomized and natural experiments. Journal of Scandinavian Studies in Criminology and Crime Prevention, 18(2), 210–222.

4. Ansari, M. A., & Singh, D. K. (2021). An expert eye for identifying shoplifters in mega stores. In Proceedings of the International Conference on Innovative Computing and Communications (ICICC 2021) (Vol. 3, pp. 107–115). Springer.
5. Ansari, M. A., & Singh, D. K. (2022). An expert video surveillance system to identify and mitigate shoplifting in megastores. Multimedia Tools and Applications, 81(16), 22497–22525.
6. Ansari, M. A., & Singh, D. K. (2023). Optimized parameter tuning in a recurrent learning process for shoplifting activity classification. Cybernetics and Information Technologies, 23(1), 141–160.
7. Armitage, R., Smyth, G., & Pease, K. (1999). Burnley CCTV evaluation. In K. A. Painter & N. Tilley (Eds.), Surveillance of public space: CCTV, street lighting and crime prevention (Crime prevention studies Vol. 10, pp. 129–150). Monsey, NY: Criminal Justice Press.
8. Arroyo, R., Yebes, J. J., Bergasa, L. M., Daza, I. G., & Almazán, J. (2015). Expert video-surveillance system for real-time detection of suspicious behaviors in shopping malls. Expert Systems with Applications, 42(21), 7991–8005. https://doi.org/10.1016/j.eswa.2015.06.007
9. Ashby, M. P. (2017). The value of CCTV surveillance cameras as an investigative tool: An empirical analysis. European Journal on Criminal Policy and Research, 23(3), 441–459. https://doi.org/10.1007/s10610-017-9341-6
10. Clarke, R. V. (1995). Situational crime prevention: Its theoretical basis and practical scope. In M. Tonry & N. Morris (Eds.), Crime and justice: A review of research (Vol. 19, pp. 91–150). Chicago, IL: University of Chicago Press.
11. Clarke, R. V. (1997). Introduction. In R. V. Clarke (Ed.), Situational crime prevention: Successful case studies (2nd ed., pp. 1–43). New York: Harrow and Heston.
12. Cornish, D. B., & Clarke, R. V. (2003). Opportunities, precipitators and criminal decisions: A reply to Wortley's critique of situational crime prevention. In M. Smith & D. B. Cornish (Eds.), Theory for situational crime prevention (Crime prevention studies Vol. 16, pp. 41–96). Monsey, NY: Criminal Justice Press.
13. Ebere Uzoka Chidi, E Anoliefo, C Udanor, AT Chijindu, LO Nwobodo (2025)” A Blind navigation guide model for obstacle avoidance using distance vision estimation based YOLO-V8n; Journal of the Nigerian Society of Physical Sciences, 2292-229; https://doi.org/10.46481/jnsps.2025.2292

14. Gill, C., Weisburd, D., Telep, C. W., Vitter, Z., & Bennett, T. (2014). Community-oriented policing to reduce crime, disorder and fear and increase satisfaction and legitimacy among citizens: A systematic review. Journal of Experimental Criminology, 10(4), 399–428. https://doi.org/10.1007/s11292-014-9210-y
15. GunjalB. L., Lende, A., Wandre, G., Thombal, A., & Rahane, C. (2024). CCTV-based crime detection using deep learning. International Journal of Innovative Research in Management, Pharmacy and Sciences (IJIRMPS), 12(2), March–April 2024.
16. Hirschorn, O., & Avidan, S. (2023). Normalizing flows for human pose anomaly detection. In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 13545–13554).
17. Ifkat, S., Mandal, A. K., Mandal, R. K., Abubeker, N., & Kumar, R. (2023). IoT-based smart surveillance system. International Journal of Research Publication and Reviews, 4(4), 4530–4543. https://doi.org/10.55248/gengpi.4.423.38048
18. Jimenez, J., Carni, D. L., Trojman, L., Prócel, L., & Lamonaca, F. (2022). Low-cost smart living environment monitoring systems from a metrological point of view. IEEE.
19. Kaviyaraj, R., Ramya, S., &Yogapriya, D. (2020). Smart surveillance with artificial intelligence: Robbery/theft motion detection. International Journal of Creative Research Thoughts (IJCRT), 8(7), 1321–1328.
20. Kekong P.E, Ajah I.A., Ebere U.C. (2019). Real-time drowsy driver monitoring and detection system using deep learning based behavioural approach. International Journal of Computer Sciences and Engineering 9 (1), 11-21
21. Kim, J. S., Kim, M. G., & Pan, S. B. (2021). A study on implementation of real-time intelligent video surveillance system based on embedded module. EURASIP Journal on Image and Video Processing, 2021, Article 76. https://doi.org/10.1186/s13640-021-00576-0
22. Kirichenko, L., Radivilova, T., Sydorenko, B., & Yakovlev, S. (2022a). Detection of shoplifting on video using a hybrid network. Computation, 10(11), 199.
23. Kirichenko, L., Sydorenko, B., Radivilova, T., & Zinchenko, P. (2022b). Video surveillance shoplifting recognition based on a hybrid neural network. In 2022 IEEE 17th International Conference on Computer Sciences and Information Technologies (CSIT) (pp. 44–47). IEEE.
24. La Vigne, N. G., Lowry, S. S., Markman, J. A., & Dwyer, A. M. (2011). Evaluating the use of public surveillance cameras for crime control and prevention. Washington, DC: URBAN Institute. Retrieved from

https://www.urban.org/research/publication/evaluating-use-public-surveillance-cameras-crime-control-and-prevention

25. Lin, B. S., Su, J. S., Chen, H., & Jan, C. Y. (2013). A fall detection system based on human body silhouette. In IEEE International Conference on Intelligent Information Hiding and Multimedia Signal Processing (pp. 49–50).
26. Lohith, P. S., Shivasai, G., Goud, L. V., Lowkya, C. H., & Khan, S. I. (2021). Smart surveillance using computer vision. International Journal of Innovative Science and Research Technology, 6(6).
27. Malaainine, M.,Lechgar, H.,&Rhinane H., (2021). YOLOv2 Deep Learning Model and GIS Based Algorithms for Vehicle Tracking. Journal of Geographic Information System. 13. 395-409. 10.4236/jgis.2021.134022.
28. Munagekar, M. S. (2018). Smart surveillance system for theft detection using image processing. International Research Journal of Engineering and Technology (IRJET), 5(8), 1198–1202.
29. Muneer, I., Saddique, M., Habib, Z., & Mohamed, H. G. (2023). Shoplifting detection using hybrid neural network CNN-BiLSMT and development of benchmark dataset. Applied Sciences, 13(14), 8341. https://doi.org/10.3390/app13148341
30. NoghreG. A., Pazho, A., Katariya, V., &Tabkhi, H. (2023). Understanding the challenges and opportunities of pose-based anomaly detection. In Proceedings of the 8th International Workshop on Sensor-Based Activity Recognition and Artificial Intelligence (pp. 1–9).
31. Owen, K., Keats, G., & Gill, M. (2006). A short evaluation of the (economic) benefits of the Milton Keynes CCTV system in managing police resources. Tunbridge Wells, England: Perpetuity Research & Consultancy International (PRCI) Ltd.
32. Pazho, A., AlinezhadNoghre, G., Katariya, V., &Tabkhi, H. (2023). Understanding the challenges and opportunities of pose-based anomaly detection. In Proceedings of the 8th International Workshop on Sensor-Based Activity Recognition and Artificial Intelligence (pp. 1–9).
33. Pazho, A., AlinezhadNoghre, G., Purkayastha, A. A., Vempati, J., Martin, O., &Tabkhi, H. (2023). A survey of graph-based deep learning for anomaly detection in distributed systems. IEEE Transactions on Knowledge and Data Engineering, 36(1), 1–20.
34. Pazho, A., Neff, C., AlinezhadNoghre, G., Ardabili, B. R., Yao, S., Baharani, M., &Tabkhi, H. (2023). Ancilia: Scalable intelligent video surveillance for the artificial intelligence of things. IEEE Internet of Things Journal, 10(17), 14940–14951.

35. Piza, E. L., Caplan, J. M., & Kennedy, L. W. (2014a). Is the punishment more certain? An analysis of CCTV detections and enforcement. Justice Quarterly, 31(6), 1015–1043.
36. Piza, E. L., Caplan, J. M., & Kennedy, L. W. (2014b). Analyzing the influence of micro-level factors on CCTV camera effect. Journal of Quantitative Criminology, 30(2), 237–264.
37. Piza, E. L., Caplan, J. M., Kennedy, L. W., & Gilchrist, A. M. (2015). The effects of merging proactive CCTV monitoring with directed police patrol: A randomized control trial. Journal of Experimental Criminology, 11(1), 43–69.
38. Prasad, S., Sharma, K., Karwasra, A., Raj, D., & Gupta, S. (2024). A review: IoT-based environmental monitoring system. International Journal of Creative Research Thoughts (IJCRT), 12(5).
39. Priyanka, N., Deeksha, S., Muskaan, F., Neha, S., & Preethi, K. (2023). Vision-based surveillance system for theft detection in supermarkets. International Research Journal of Modernization in Engineering Technology and Science, 5(5).
40. Ratan P., (2024) What is the Convolutional Neural Network Architecture?https://www.analyticsvidhya.com/blog/2020/10/what-is-the-convolutional-neural-network-architecture/ Accessed April 2025
41. Ratcliffe, J. (2006). Video surveillance of public places (Problem-oriented guides for police. Response guide series. Guide No. 4). Washington, DC: U.S. Department of Justice, Office of Community Oriented Policing Services. Center for Problem-Oriented Policing.
42. Reid, S., Coleman, S., Vance, P., Kerr, D., & O'Neill, S. (2021). Using social signals to predict shoplifting: A transparent approach to a sensitive activity analysis problem. Sensors, 21(20), 6812. https://doi.org/10.3390/s21206812
43. Sai, P., Ibrahim, M., & Siddiq, S. (2024). Smart security surveillance system using IoT. Vignan's Foundation for Science, Technology and Research.
44. Shaikh, A., Gurav, S., Jagdale, A., Kamble, P., & Chavan, S. (2024). IoT-based connected environmental monitoring system using NodeMCU. International Journal for Multidisciplinary Research (IJFMR), E-ISSN: 2582-2160.
45. Sivakumar, G. S., Sunitha, P., Sailaja, V., Srinivas, A., & Dash, S. (2022). Smart surveillance system using computer vision methods for smart cities with IoT-enabled networks. NeuroQuantology, 20(10), 10031–10048. https://doi.org/10.14704/nq.2022.20.10.NQ55976

46. Sochima V.E. Asogwa T.C., Lois O.N. Onuigbo C.M., Frank E.O., Ozor G.O., Ebere U.C. (2025)"; Comparing multi-control algorithms for complex nonlinear system: An embedded programmable logic control applications; DOI: http://doi.org/10.11591/ijpeds.v16.i1.pp212-224
47. Sultani, W., Chen, C., & Shah, M. (2018). Real-world anomaly detection in surveillance videos. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 6479–6488).
48. Tsourma, M., &Dasygenis, M. (2016). Development of a hybrid defensive embedded system with face recognition. In DCABES (pp. 154–157).
49. Welsh, B. C., & Farrington, D. P. (2009a). Public area CCTV and crime prevention: An updated systematic review and meta-analysis. Justice Quarterly, 26, 716–745.
50. Welsh, B. C., & Farrington, D. P. (2009b). Making public place safer: Surveillance and crime prevention. New York: Oxford University Press.
51. Winge, S., & Knutsson, J. (2003). An evaluation of the CCTV scheme at Oslo central railway station. Crime Prevention and Community Safety: An International Journal, 5(3), 49–59.
52. Zafar, S., Miraj, G., Baloch, R., Murtaza, D., & Arshad, K. (2018). An IoT-based real-time environmental monitoring system using Arduino and cloud service. Engineering, Technology & Applied Science Research, 8(4), 3238–3242.
53. Zahari, M. K. M., &Zaaba, Z. F. (2017). Intelligent responsive indoor system (IRIS): A potential shoplifter security alert system. Journal of ICT, 16(2), 262–282.
54. Zhang, S., et al. (2012). On the design and implementation of a high definition multi-view intelligent video surveillance system. In IEEE International Conference on Signal Processing, Communication and Computing (pp. 353–357).
55. Zulkifli, N. S. A., Satrial, M. R., Osman, M. Z., Ismail, N. S. N., & Razif, M. R. M. (2020). IoT-based smart environment monitoring system for air pollutant detection in Kuantan, Pahang, Malaysia. In The 6th International Conference on Software Engineering & Computer Systems. IOP Conference Series: Materials Science and Engineering, 769(1), 012014. https://doi.org/10.1088/1757-899X/769/1/012014