International Journal for Novel Research in Economics , Finance and Management
www.ijnrefm.com
ISSN:3048-7722
Volume 4, Issue 4, July-Aug-2026, PP: 01-10

An Edge AI Framework for Real-Time Mobile Phone Detection and Identity Recognition

G.Chandram[1], Manasa Sai[2]

[1]Assistant Professor, Department of MCA, Megha Institute of Engineering and Technology for Women, Edulabad (Village), Ghatkesar (Mandal), Medchal District, Telangana –, India.

[2]MCA Student, Department of MCA, Megha Institute of Engineering and Technology for Women, Edulabad (Village), Ghatkesar (Mandal), Medchal District, Telangana –, India.

***Abstract* – The unauthorized use of mobile phones within restricted environments such as examination halls, government offices, hospitals, research laboratories, and other secure facilities presents serious concerns related to security, privacy, operational efficiency, and regulatory compliance. Conventional mobile phone detection techniques primarily rely on monitoring radio frequency or network signals, making them ineffective when devices operate in airplane mode, remain disconnected from communication networks, or employ signal suppression mechanisms. To overcome these limitations, this paper presents Phone Patrol, an intelligent image-based surveillance framework developed for detecting unauthorized mobile phone usage in mobile-free environments. The proposed system integrates You Only Look Once version 5 (YOLOv5) for real-time mobile phone detection and You Only Look Once version 9 (YOLOv9) for face recognition, enabling accurate identification of individuals violating mobile usage policies. The framework is deployed on a Raspberry Pi 5, providing a compact and cost-effective Internet of Things (IoT) solution capable of processing live video streams and automatically issuing SMS notifications through the Twilio API whenever a violation is detected. Experimental evaluation demonstrates excellent detection performance, with the YOLOv5 model achieving a mean Average Precision (mAP50) of 0.992, confirming its capability to accurately recognize mobile phone usage under real-time operating conditions. Although the current implementation is intended as a prototype and requires further optimization for large-scale deployment, the proposed framework demonstrates the effectiveness of combining deep learning, edge computing, and automated notification services to enforce mobile-free policies. The developed system provides a reliable, network-independent surveillance solution that strengthens security, improves policy compliance, and supports intelligent monitoring in restricted environments.**

Keywords - **Edge Artificial Intelligence (Edge AI), Mobile Phone Detection Identity Recognition, Real-Time Object Detection, Computer Vision, Deep Learning, Convolutional Neural Networks (CNNs)**

I. INTRODUCTION

The rapid advancement of smartphones and wireless communication technologies has transformed the way individuals communicate, access information, and perform daily activities. Mobile devices have become indispensable tools across educational institutions, healthcare facilities, government organizations, financial sectors, and industrial environments because of their ability to provide instant communication and internet connectivity. Despite these advantages, unrestricted mobile phone usage within designated no-mobile zones has introduced significant challenges related to security, privacy, operational efficiency, and policy enforcement. Sensitive environments such as examination halls, government offices, hospitals, military facilities, research laboratories, libraries, conference rooms, and places of worship require strict control over mobile phone usage to prevent unauthorized communication, information leakage, and unnecessary distractions. Consequently, ensuring compliance with mobile-free policies has become an important concern for organizations responsible for maintaining secure and distraction-free environments.

The presence of unauthorized mobile devices in restricted areas can create numerous operational and security risks. In educational institutions, mobile phones facilitate unfair examination practices through communication, internet access, and unauthorized information sharing. Within government organizations and corporate offices, smartphones may be used to capture confidential documents, record sensitive discussions, or transmit classified information outside secure premises. Similarly, hospitals and medical centers often restrict mobile phone usage because electromagnetic interference can affect the operation of critical medical equipment while unnecessary phone usage may disturb healthcare delivery. In industrial facilities and research laboratories, unrestricted mobile device usage may compromise confidential projects and introduce additional workplace safety concerns. These challenges emphasize the necessity of intelligent surveillance systems capable of automatically identifying mobile phone usage without requiring continuous human supervision.

Conventional mobile phone detection techniques primarily rely on radio frequency (RF) monitoring, electromagnetic signal analysis, and cellular network detection. These methods identify mobile devices by detecting communication signals transmitted between smartphones and nearby cellular towers or wireless communication infrastructure. Although signal-based approaches have demonstrated reasonable performance under normal operating conditions, they exhibit several practical limitations. Modern smartphones frequently operate in airplane mode, remain disconnected from cellular

networks, or utilize Wi-Fi-independent applications that significantly reduce detectable radio frequency emissions. Furthermore, signal shielding techniques, electromagnetic interference, and increasingly sophisticated communication technologies often reduce the effectiveness of traditional RF-based detection systems. Consequently, relying exclusively on network signals limits the ability to identify unauthorized mobile phones operating silently within restricted environments.

Recent advances in Artificial Intelligence (AI) and Deep Learning (DL) have introduced highly effective alternatives for intelligent visual surveillance. Computer vision techniques enable automatic interpretation of digital images and video streams, allowing systems to recognize objects, human activities, and environmental events without depending on communication signals. Deep learning models have demonstrated remarkable performance in applications such as object detection, facial recognition, autonomous surveillance, traffic monitoring, industrial inspection, healthcare imaging, and biometric authentication. Their capability to automatically learn complex visual features from large image datasets has significantly improved detection accuracy compared with traditional handcrafted feature extraction methods.

Among modern object detection algorithms, the You Only Look Once (YOLO) family has emerged as one of the most efficient real-time detection frameworks. Unlike conventional region-based detection methods that require multiple processing stages, YOLO performs object localization and classification simultaneously within a single neural network pass, enabling high-speed inference while maintaining excellent detection accuracy. YOLOv5 offers substantial improvements in computational efficiency, lightweight architecture, and detection performance, making it highly suitable for deployment on embedded devices with limited computational resources. Similarly, YOLOv9 introduces enhanced feature extraction capabilities that improve recognition accuracy for fine-grained visual tasks such as face identification. The combination of YOLOv5 for mobile phone detection and YOLOv9 for facial recognition provides an effective solution for simultaneously detecting unauthorized device usage and identifying the responsible individual.

The increasing popularity of edge computing has further accelerated the deployment of intelligent surveillance systems on compact embedded platforms. Devices such as the Raspberry Pi 5 provide sufficient computational capability to execute lightweight deep learning models while maintaining low power consumption, compact size, and affordable deployment costs. Edge-based processing eliminates the dependency on continuous cloud communication, reduces response latency, and improves privacy by performing inference locally. These characteristics make Raspberry Pi an attractive hardware platform for developing standalone surveillance systems capable of operating continuously within restricted environments.

In addition to accurate detection, effective enforcement of mobile-free policies requires rapid communication with security personnel or authorized administrators. Automated notification mechanisms significantly reduce response time by immediately informing responsible authorities whenever unauthorized mobile phone usage is detected. Cloud communication services such as the Twilio API enable real-time transmission of SMS alerts, allowing surveillance systems to automatically notify designated recipients together with relevant information regarding detected policy violations. Integrating intelligent detection with automated alert generation transforms passive monitoring systems into proactive security solutions capable of supporting immediate corrective action.

The proposed research introduces Phone Patrol, an intelligent surveillance framework specifically designed for enforcing mobile-free policies through image-based deep learning techniques. The framework integrates YOLOv5 for real-time mobile phone detection and YOLOv9 for facial recognition to accurately identify individuals using unauthorized mobile devices. The complete system is implemented on a Raspberry Pi 5, enabling continuous processing of live video streams acquired from a USB camera while maintaining a compact and cost-effective Internet of Things (IoT) architecture. Upon successful detection of both the mobile phone and the associated individual, the system automatically generates SMS notifications through the Twilio API, enabling rapid policy enforcement without requiring continuous manual observation.

The primary objective of this research is to develop a reliable, network-independent surveillance system capable of detecting unauthorized mobile phone usage regardless of the communication status of the device. By integrating advanced object detection, facial recognition, embedded edge computing, and automated notification services within a unified framework, the proposed solution enhances security, strengthens policy compliance, minimizes manual monitoring effort, and improves operational efficiency across various mobile-free environments. The experimental evaluation demonstrates that combining YOLOv5, YOLOv9, Raspberry Pi 5, and Twilio-based alert generation provides an effective and practical solution for intelligent surveillance in modern restricted-access facilities.

II. LITERATURE SURVEY

The rapid growth of smartphone technology has significantly increased the need for intelligent surveillance systems capable of enforcing mobile-free policies in secure environments. Unauthorized mobile phone usage in locations such as examination halls, hospitals, government

organizations, research laboratories, and industrial facilities poses serious risks related to confidentiality, security, productivity, and operational safety. Consequently, researchers have explored numerous detection techniques ranging from conventional signal-based approaches to advanced computer vision and deep learning methods. Recent developments in artificial intelligence have further accelerated the adoption of image-based surveillance systems that provide more reliable and flexible mobile phone detection independent of wireless communication signals.

Early research primarily focused on signal-based mobile phone detection. Ataro et al. proposed mobile phone detection circuits based on resistor-capacitor and reed-switch mechanisms capable of identifying nearby mobile devices through electromagnetic characteristics. Their approach demonstrated the feasibility of detecting mobile phones operating in silent mode and, under certain conditions, devices in flight mode. However, the detection range remained limited, and the proposed circuits were unable to distinguish user identity or provide intelligent monitoring capabilities required for large-scale restricted environments. These limitations highlighted the need for more sophisticated surveillance systems capable of identifying both the mobile device and the individual using it.

Another widely investigated approach involved monitoring electromagnetic signals emitted by active mobile phones. Mbaocha introduced an intelligent mobile phone detector that identified devices by analyzing communication signals generated during network activity. Although the proposed framework successfully detected actively transmitting phones, its performance depended heavily on the presence of detectable radio frequency emissions. Consequently, devices operating in airplane mode, switched off, or disconnected from cellular networks remained undetected. Such limitations significantly reduce the effectiveness of signal-based techniques in modern restricted environments where users intentionally disable communication services to avoid detection.

The emergence of computer vision has provided a promising alternative to traditional signal-dependent detection methods. Abdulhamid et al. demonstrated the practical deployment of computer vision algorithms on Raspberry Pi platforms for intelligent image analysis applications. Their work confirmed that embedded edge devices provide sufficient computational capability to execute lightweight image processing algorithms while maintaining low cost and energy consumption. These findings encouraged the development of compact surveillance systems capable of performing real-time visual monitoring directly on edge devices without continuous dependence on cloud infrastructure. The successful integration of computer vision with Raspberry Pi technology established an important foundation for intelligent embedded surveillance applications.

Researchers have also extensively investigated face recognition for security and access control applications. Owayjan et al. developed a face recognition-based security system capable of identifying authorized and unauthorized individuals entering protected areas. Their framework demonstrated that automated facial recognition significantly enhances surveillance efficiency by reducing manual identity verification while improving security enforcement. Similarly, Ayyappan et al. proposed a face recognition system for identifying criminals and missing children through the integration of facial recognition algorithms and web-based information retrieval. Although their methodology achieved promising identification accuracy, its dependence on web-scraped datasets introduced challenges related to image quality, dataset inconsistency, and computational complexity. These studies collectively demonstrate the effectiveness of facial recognition technology for intelligent security monitoring while revealing opportunities for more efficient real-time implementations.

Recent advances in deep learning have significantly improved real-time object detection capabilities. The You Only Look Once (YOLO) family of object detection algorithms has become one of the most widely adopted frameworks because of its ability to perform object localization and classification simultaneously within a single neural network inference. YOLO-based models provide an effective balance between computational efficiency and detection accuracy, making them highly suitable for surveillance applications requiring continuous real-time processing. Lightweight variants such as YOLOv5 have further improved inference speed and memory efficiency, enabling deployment on resource-constrained embedded devices while maintaining high object detection performance. These characteristics have made YOLO one of the preferred object detection frameworks for intelligent surveillance systems.

Continuous improvements in object detection architecture have led to the development of YOLOv9, which provides enhanced feature extraction capabilities and improved recognition performance compared with earlier YOLO versions. Recent investigations have demonstrated that YOLOv9 performs particularly well for fine-grained recognition tasks such as face detection and facial identification under varying lighting conditions, viewpoints, and environmental complexities. The improved detection precision achieved by YOLOv9 makes it highly appropriate for surveillance applications requiring accurate identification of individuals involved in security policy violations.

The adoption of edge computing has further strengthened the deployment of intelligent surveillance systems. Embedded computing platforms such as the Raspberry Pi 5 provide an attractive combination of low power consumption, compact design, affordability, and sufficient computational capability for executing lightweight deep

International Journal for Novel Research in Economics , Finance and Management
www.ijnrefm.com
ISSN:3048-7722
Volume 4, Issue 4, July-Aug-2026, PP: 01-10

learning models locally. Local inference reduces communication latency, minimizes cloud dependency, improves data privacy, and enables continuous surveillance even under limited network connectivity. These advantages have encouraged researchers to deploy object detection and facial recognition models directly on embedded IoT devices for practical security monitoring applications.

Another important advancement involves the integration of cloud communication services for automated security response. Instead of relying solely on visual monitoring, recent surveillance systems incorporate notification mechanisms capable of immediately informing responsible authorities when abnormal events are detected. Communication platforms such as the Twilio API enable automatic generation of SMS alerts, providing rapid notification together with violation details. Automated alert generation significantly reduces response time while improving the effectiveness of intelligent surveillance systems operating in restricted environments.

Despite the considerable progress achieved by previous research, several important limitations remain. Most traditional detection methods depend on radio frequency or electromagnetic signal analysis, preventing reliable detection of devices operating in airplane mode or without active network connectivity. Computer vision-based systems often focus exclusively on object detection without simultaneously identifying the individual responsible for the violation. Similarly, many face recognition frameworks operate independently of object detection, limiting their applicability for comprehensive policy enforcement. Furthermore, relatively few studies have integrated real-time mobile phone detection, face recognition, embedded edge computing, and automated SMS notification within a unified surveillance framework suitable for pratical deployment in mobile-free environments.

Motivated by these research gaps, the present study proposes Phone Patrol, an integrated deep learning-based surveillance framework that combines YOLOv5 for mobile phone detection, YOLOv9 for face recognition, Raspberry Pi 5 for edge-based real-time processing, and the Twilio API for automated SMS alert generation. Unlike conventional signal-based detection approaches, the proposed framework performs entirely image-based mobile phone detection, enabling reliable identification of unauthorized mobile device usage regardless of network connectivity or airplane mode. By integrating object detection, facial recognition, embedded computing, and intelligent notification services into a single surveillance architecture, the proposed system provides an efficient, scalable, and practical solution for enforcing mobile-free policies in secure environments while improving security, policy compliance, and operational efficiency.

III. PROPOSED METHODOLOGY

The proposed Phone Patrol framework is designed to provide an intelligent and automated solution for detecting unauthorized mobile phone usage in mobile-free environments through advanced deep learning, computer vision, and Internet of Things (IoT) technologies. Unlike conventional mobile phone detection systems that rely on monitoring radio frequency signals, the proposed framework performs image-based detection, enabling reliable identification of mobile phones regardless of whether the device is connected to a communication network, operating in airplane mode, or disconnected from cellular services. By integrating YOLOv5, YOLOv9, Raspberry Pi 5, and the Twilio API, the framework simultaneously detects mobile phone usage, recognizes the identity of the individual involved, and automatically generates real-time notifications to support immediate policy enforcement.

The framework begins with the video acquisition stage, where a high-resolution USB camera continuously captures live video streams from the designated no-mobile zone. The surveillance area may include examination halls, hospitals, government offices, conference rooms, libraries, research laboratories, or any other location where mobile phone usage is prohibited. The captured video frames are transmitted directly to the Raspberry Pi 5, which functions as the central processing unit of the proposed surveillance system. The Raspberry Pi performs all image processing and deep learning inference locally, eliminating the dependency on external cloud-based computation and enabling real-time monitoring within a compact and energy-efficient IoT environment.

Once the video frames are received, the proposed system performs image preprocessing to prepare the captured frames for deep learning analysis. Each frame is resized according to the input requirements of the object detection model while preserving image quality and essential visual features. Standard preprocessing operations such as normalization and image scaling are applied to improve computational efficiency and maintain consistent model performance under different illumination conditions. These preprocessing operations ensure that the subsequent object detection stage receives standardized image inputs capable of producing reliable detection results.

Following preprocessing, the framework executes the mobile phone detection module using the YOLOv5 object detection model. YOLOv5 is selected because of its excellent balance between computational efficiency and detection accuracy, making it highly suitable for deployment on embedded hardware platforms such as the Raspberry Pi 5. The model is specifically fine-tuned to classify two object categories: person with phone and person without phone. During inference, YOLOv5 simultaneously performs object localization and classification in a single forward pass through the neural

International Journal for Novel Research in Economics , Finance and Management
www.ijnrefm.com
ISSN:3048-7722
Volume 4, Issue 4, July-Aug-2026, PP: 01-10

network, enabling real-time identification of individuals carrying or actively using mobile phones. When the model detects a person with phone, the corresponding bounding box coordinates are generated and forwarded to the subsequent face recognition module for identity verification.

The proposed framework then activates the face recognition module, which utilizes the YOLOv9 deep learning model for identifying individuals associated with the detected mobile phone. Unlike the object detection stage, which identifies the presence of a mobile device, YOLOv9 focuses specifically on recognizing the face of the detected individual. The facial region corresponding to the bounding box generated by YOLOv5 is extracted and processed by the fine-tuned YOLOv9 model. The face recognition model has been trained to identify authorized individuals within the monitored environment and associates each recognized face with the corresponding personal information stored in the system database. This dual-stage detection strategy significantly improves surveillance accuracy by ensuring that notifications are generated only after both mobile phone detection and facial identification have been successfully completed.

After successful face recognition, the identified individual is matched with a pre-registered database containing user information and corresponding mobile contact details. The database maintains the mapping between recognized facial identities and authorized contact numbers, allowing the framework to retrieve the appropriate recipient information automatically. This identity mapping process enables personalized notification rather than generic violation reporting, thereby improving the effectiveness of policy enforcement while reducing manual administrative effort.

Once both mobile phone usage and offender identity have been confirmed, the proposed framework enters the alert generation module. The system automatically prepares a violation report containing the detected individual's identity, timestamp, and event information. The report is transmitted through the Twilio Application Programming Interface (API), which immediately generates an SMS notification directed to the registered mobile number associated with the identified individual. Automated notification enables rapid response by security personnel or administrative authorities without requiring continuous manual surveillance. The real-time alert mechanism further strengthens policy compliance by informing offenders immediately after unauthorized mobile phone usage is detected.

To ensure continuous operation, the proposed framework performs real-time monitoring throughout the surveillance period. The USB camera continuously captures video frames while the Raspberry Pi repeatedly executes object detection, face recognition, identity verification, and alert generation for every incoming frame. This continuous processing allows the system to monitor multiple individuals dynamically while maintaining low processing latency suitable for practical deployment in restricted environments. Since the complete processing pipeline executes locally on the Raspberry Pi, the framework minimizes communication overhead while preserving user privacy and improving operational reliability.

The proposed framework additionally emphasizes edge-based computing, which represents an important advantage over cloud-dependent surveillance systems. Performing inference directly on the Raspberry Pi reduces network latency, eliminates dependence on high-bandwidth internet connectivity, and enables uninterrupted surveillance even when external communication services experience temporary interruptions. The compact hardware architecture also simplifies installation within existing surveillance infrastructures while reducing deployment costs and overall power consumption.

Overall, the proposed Phone Patrol framework establishes an intelligent surveillance architecture by integrating YOLOv5-based mobile phone detection, YOLOv9-based face recognition, Raspberry Pi 5 edge computing, USB camera-based video acquisition, and Twilio API-based SMS notification into a unified IoT system. The coordinated interaction among these components enables accurate detection of unauthorized mobile phone usage, reliable identification of offenders, and immediate automated alert generation without relying on mobile network signals. Consequently, the proposed framework provides a practical, scalable, and efficient solution for enforcing mobile-free policies while improving security, confidentiality, operational discipline, and administrative efficiency across various restricted environments.

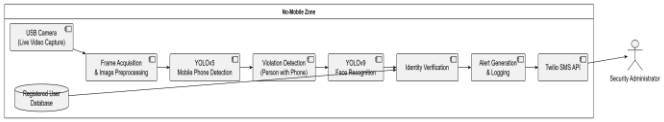

Fig. 1. Proposed System Architecture of Phone Patrol for Intelligent Mobile-Free Environment Surveillance

IV. SYSTEM ARCHITECTURE

The proposed Phone Patrol architecture is designed as an intelligent edge-based surveillance framework that integrates computer vision, deep learning, face recognition, and Internet of Things (IoT) technologies to detect unauthorized mobile phone usage in restricted environments. The architecture enables continuous monitoring of no-mobile zones by combining real-time video acquisition, deep learning-based object detection, facial recognition, identity verification, and automated notification into a unified processing pipeline. Each functional module performs a specialized task while interacting seamlessly with other components to provide accurate, reliable, and real-time surveillance. The modular design also supports scalability, low computational

overhead, and future enhancement without modifying the overall operational workflow.

The architecture begins with the Video Acquisition Module, where a high-resolution USB camera continuously captures live video streams from the designated no-mobile environment. The surveillance camera monitors locations such as examination halls, hospitals, government offices, libraries, research laboratories, conference rooms, and other secure facilities where mobile phone usage is prohibited. The captured video frames are transmitted directly to the Raspberry Pi 5, which serves as the central processing unit responsible for executing the complete surveillance pipeline.

Once the video stream is received, the captured frames are forwarded to the Image Preprocessing Module. Since real-time surveillance requires standardized image inputs, preprocessing operations are performed to improve computational efficiency before deep learning inference. Each frame is resized according to the input dimensions required by the detection models while maintaining image quality and preserving important visual information. Image normalization and scaling are also performed to improve detection consistency under varying lighting conditions. These preprocessing operations reduce computational complexity while ensuring stable object detection performance.

Following preprocessing, the standardized image enters the Mobile Phone Detection Module, where the YOLOv5 object detection model performs real-time analysis. YOLOv5 has been selected because of its lightweight architecture, fast inference capability, and excellent detection accuracy on embedded hardware platforms such as the Raspberry Pi 5. The model is specifically trained to classify two categories: person with phone and person without phone. During inference, YOLOv5 simultaneously performs object localization and classification in a single forward pass, allowing rapid identification of unauthorized mobile phone usage. Whenever the system detects an individual carrying or actively using a mobile phone, the corresponding bounding box coordinates are generated and transferred to the subsequent facial recognition module for identity verification.

The detected violation is then processed by the Face Recognition Module, which employs the YOLOv9 deep learning model to identify the individual responsible for the policy violation. Unlike YOLOv5, which focuses on detecting mobile phone usage, YOLOv9 specializes in facial recognition by analyzing the cropped facial region extracted from the previously detected bounding box. The model has been fine-tuned to recognize registered individuals within the surveillance environment, enabling reliable identification under different viewing angles and illumination conditions. This two-stage detection strategy significantly improves surveillance reliability by ensuring that alerts are generated only after successful mobile phone detection and face recognition have both been completed.

Following facial recognition, the framework activates the Identity Verification Module, where the recognized face is compared with information stored in the registered user database. The database maintains records containing user identities and corresponding contact information associated with authorized personnel or registered individuals. Once a successful facial match is obtained, the system retrieves the relevant contact details required for automated notification. This identity verification process eliminates manual identification while enabling personalized alert generation based on the recognized individual.

The verified information is subsequently transferred to the Alert Generation Module, where a complete violation report is automatically prepared. The generated report includes the recognized person's identity, violation timestamp, detection details, and surveillance event information. The prepared notification is forwarded to the Twilio Application Programming Interface (API), which functions as the communication gateway responsible for transmitting SMS alerts. Using the stored contact information obtained during identity verification, Twilio immediately sends a notification message to the corresponding registered mobile number, thereby enabling rapid policy enforcement and administrative response. Automated SMS generation significantly reduces manual intervention while improving response time whenever unauthorized mobile phone usage is detected.

The proposed architecture continuously repeats the complete surveillance cycle throughout system operation. Incoming video frames are captured, preprocessed, analyzed using YOLOv5, verified using YOLOv9, matched with the registered database, and processed through automated SMS generation in real time. Since the complete inference pipeline executes locally on the Raspberry Pi 5, the system minimizes communication latency while maintaining continuous surveillance independent of cloud-based computation. Edge-based deployment also enhances data privacy, reduces bandwidth requirements, and enables reliable operation even in environments with limited internet connectivity.

An additional strength of the proposed architecture is its network-independent detection capability. Conventional mobile phone detection systems rely on electromagnetic or radio frequency signals, making them ineffective when devices operate in airplane mode or remain disconnected from communication networks. In contrast, the proposed image-based framework analyzes visual information directly from surveillance footage, enabling reliable detection regardless of network status or communication protocol. This characteristic significantly expands the applicability of the framework across diverse restricted

environments where traditional signal-based detection techniques are ineffective.

Overall, the proposed system architecture successfully integrates USB camera-based video acquisition, YOLOv5 mobile phone detection, YOLOv9 face recognition, Raspberry Pi 5 edge computing, identity verification, Twilio-based SMS notification, and real-time surveillance into a unified intelligent monitoring framework. The coordinated interaction among these modules enables accurate detection of unauthorized mobile phone usage, reliable offender identification, immediate automated notification, and efficient enforcement of mobile-free policies. Consequently, the proposed architecture provides a practical, scalable, and cost-effective surveillance solution that improves security, confidentiality, operational discipline, and regulatory compliance within modern restricted-access environments.

V. FEATURE ENHANCEMENT

The proposed Phone Patrol framework incorporates several enhancements that significantly improve the efficiency, reliability, and practicality of detecting unauthorized mobile phone usage in restricted environments. By integrating YOLOv5, YOLOv9, Raspberry Pi 5, computer vision, edge computing, and the Twilio API into a unified surveillance platform, the system provides accurate real-time monitoring without relying on conventional radio frequency detection techniques. These enhancements enable continuous automated surveillance while improving security, reducing manual intervention, and strengthening policy enforcement across various mobile-free environments.

One of the major enhancements introduced by the proposed framework is the adoption of image-based mobile phone detection instead of conventional signal-based detection techniques. Traditional mobile phone detectors primarily identify devices through radio frequency emissions or network communication signals, making them ineffective when smartphones operate in airplane mode, remain disconnected from cellular networks, or suppress wireless communication. The proposed framework overcomes these limitations by analyzing visual information captured from live surveillance video. Since detection depends entirely on computer vision rather than communication signals, the system can accurately identify mobile phones regardless of their network connectivity or operating mode. This enhancement substantially improves the reliability and applicability of the surveillance system within highly secure environments.

Another significant enhancement is the implementation of YOLOv5 as the primary object detection model. YOLOv5 provides an effective balance between computational efficiency and detection accuracy while maintaining real-time processing capability. Unlike conventional object detection algorithms that perform localization and classification separately, YOLOv5 simultaneously identifies object locations and categories within a single neural network inference. This unified detection mechanism reduces computational latency while enabling continuous monitoring of live video streams. The lightweight architecture of YOLOv5 also makes it highly suitable for deployment on embedded hardware platforms with limited computational resources, ensuring stable performance without requiring high-end computing infrastructure.

The proposed system further enhances surveillance performance by integrating YOLOv9 for face recognition. While the mobile phone detection module identifies the presence of unauthorized devices, the face recognition module accurately determines the identity of the individual responsible for the policy violation. This two-stage detection strategy significantly improves surveillance intelligence compared with systems that perform only object detection. By associating detected mobile phones with specific individuals, the framework enables personalized policy enforcement while reducing false notifications and improving administrative accountability.

A notable enhancement of the framework is the implementation of edge computing using the Raspberry Pi 5. Rather than transmitting surveillance footage to cloud servers for processing, the proposed framework performs all deep learning inference locally on the embedded device. Local processing substantially reduces communication latency, minimizes bandwidth consumption, and improves system responsiveness during real-time surveillance. Furthermore, edge-based computation enhances data privacy by preventing continuous transmission of surveillance images across external communication networks. The compact size, low power consumption, and affordable cost of the Raspberry Pi 5 additionally simplify deployment across multiple restricted environments without requiring specialized computing infrastructure.

The proposed framework also introduces real-time automated alert generation through integration with the Twilio Application Programming Interface (API). Following successful detection of unauthorized mobile phone usage and identification of the corresponding individual, the system automatically prepares and transmits an SMS notification to the registered contact number. This automated notification mechanism significantly reduces response time by immediately informing administrators or responsible personnel whenever a violation occurs. Unlike conventional surveillance systems that depend on continuous human observation, the proposed framework performs intelligent monitoring and notification autonomously, thereby improving operational efficiency while minimizing manual supervision.

Another important enhancement involves the implementation of a registered user database for identity verification. The framework maintains facial information and contact details of authorized users within a centralized database. Following successful face recognition, the detected facial identity is automatically matched with the stored database records to retrieve the corresponding user information. This automated identity management process enables personalized alert generation while eliminating manual identification procedures. Consequently, the system improves surveillance accuracy and simplifies administrative management within restricted facilities.

The proposed methodology further enhances detection reliability through continuous real-time surveillance. Live video streams are continuously acquired from the USB camera and processed frame by frame using the deep learning models. This uninterrupted monitoring enables rapid detection of mobile phone usage immediately after it occurs, allowing security personnel to respond without unnecessary delay. Continuous surveillance also improves the probability of detecting short-duration violations that may be overlooked during periodic manual inspections.

An additional enhancement of the framework is its modular system architecture, which improves flexibility and future scalability. Individual functional components including video acquisition, preprocessing, mobile phone detection, face recognition, identity verification, and SMS notification operate as independent modules connected through well-defined interfaces. Such modular organization simplifies future integration of improved deep learning models, additional surveillance cameras, cloud synchronization services, biometric authentication methods, or advanced notification platforms without requiring substantial modifications to the overall system architecture.

The framework also improves computational efficiency through optimized image preprocessing and lightweight deep learning inference. Image normalization, resizing, and efficient frame processing reduce unnecessary computational overhead while maintaining detection accuracy. The combination of optimized preprocessing with lightweight YOLO architectures enables the Raspberry Pi 5 to perform real-time inference despite its limited hardware resources. Consequently, the proposed system achieves an effective compromise between computational complexity and surveillance performance suitable for practical IoT deployment.

Finally, the proposed framework significantly enhances security, operational discipline, and policy compliance across mobile-free environments. Automated detection and immediate notification discourage unauthorized mobile phone usage while supporting administrators in maintaining secure working environments. Educational institutions can reduce examination malpractice, government organizations can strengthen confidentiality, hospitals can minimize operational interference, and research laboratories can better protect sensitive information through continuous intelligent surveillance.

Overall, the proposed feature enhancements establish Phone Patrol as a comprehensive intelligent surveillance framework capable of accurately detecting unauthorized mobile phone usage while simultaneously identifying offenders and generating automated notifications. By integrating YOLOv5, YOLOv9, Raspberry Pi 5, computer vision, edge computing, and the Twilio API, the proposed system provides a reliable, scalable, and cost-effective solution for enforcing mobile-free policies. These enhancements substantially improve detection accuracy, operational efficiency, surveillance intelligence, administrative responsiveness, and security across modern restricted-access environments while demonstrating the practical benefits of deep learning-based edge surveillance systems.

VI. RESULTS AND DISCUSSION

The proposed Phone Patrol framework was experimentally evaluated to determine its effectiveness in detecting unauthorized mobile phone usage and recognizing individuals in restricted environments using real-time deep learning techniques. The experimental analysis focused on evaluating the performance of the YOLOv5 model for mobile phone detection and the YOLOv9 model for face recognition while validating the capability of the integrated Raspberry Pi 5 platform to execute the complete surveillance pipeline efficiently. Performance evaluation was conducted using standard object detection metrics, including Precision, Recall, mean Average Precision at IoU threshold 0.5 (mAP50), and mean Average Precision across IoU thresholds 0.5–0.95 (mAP50–95). The obtained results demonstrate that the proposed framework provides highly reliable detection performance while maintaining real-time operation suitable for deployment in mobile-free environments.

Prior to model training, the collected image dataset underwent extensive preprocessing and augmentation to improve model robustness under diverse environmental conditions. Images containing individuals with and without mobile phones were resized, normalized, and augmented using multiple transformation techniques to increase dataset diversity while preserving important visual characteristics. These preprocessing operations enabled the deep learning models to learn discriminative visual features more effectively, thereby improving detection accuracy under varying illumination conditions, viewing angles, and background complexity. The prepared dataset was subsequently divided into training, validation, and testing subsets to enable objective evaluation of the developed models.

The first stage of experimentation evaluated the YOLOv5 model for real-time mobile phone detection. During

training, the model demonstrated stable convergence while continuously improving localization and classification accuracy across successive epochs. The precision and recall curves indicated effective learning behavior with minimal overfitting, confirming the suitability of the selected network architecture for the mobile phone detection task. The final experimental evaluation reported a Precision of 0.921 and a Recall of 0.977, demonstrating that the model accurately identified unauthorized mobile phone usage while minimizing both false positive and false negative detections. High recall indicates that the majority of policy violations were successfully detected, whereas high precision confirms the reliability of the generated detection results.

The object detection capability of the proposed framework was further assessed using mean Average Precision (mAP). The experimental results showed that the YOLOv5 model achieved an impressive mAP50 value of 0.992, indicating highly accurate localization and classification of mobile phones within surveillance images. Furthermore, the model obtained a mAP50–95 value of 0.873, demonstrating consistent detection performance across multiple Intersection over Union (IoU) thresholds. These results confirm that YOLOv5 effectively recognizes mobile phones under diverse environmental conditions while maintaining excellent localization accuracy suitable for practical surveillance applications.

The second phase of experimentation focused on evaluating the YOLOv9 face recognition model responsible for identifying individuals associated with detected mobile phone usage. Similar to the object detection model, YOLOv9 demonstrated stable learning behavior throughout the training process while effectively extracting discriminative facial features from the prepared dataset. Experimental analysis revealed that the face recognition model achieved a mAP50 of 0.994, confirming its exceptional capability for recognizing registered individuals within the monitored environment. Additionally, the model achieved a mAP50–95 value of 0.814, demonstrating reliable recognition performance across different detection thresholds and facial orientations. These findings indicate that YOLOv9 provides accurate facial identification suitable for intelligent surveillance applications requiring both identity verification and policy enforcement.

The integration of YOLOv5 and YOLOv9 within a unified surveillance framework significantly enhanced overall system intelligence. Rather than simply detecting the presence of a mobile phone, the proposed framework simultaneously identifies the individual responsible for the policy violation. This two-stage verification strategy reduces false notifications while improving administrative accountability. Experimental observations demonstrated that combining object detection with face recognition provides more meaningful surveillance information than performing either task independently.

The deployment of the proposed framework on the Raspberry Pi 5 further demonstrated the practicality of edge-based artificial intelligence for real-time surveillance. Despite the computational limitations of embedded hardware, the lightweight architectures of YOLOv5 and YOLOv9 enabled continuous processing of live video streams with satisfactory inference speed. Local execution eliminated dependence on cloud-based processing, thereby reducing communication latency, preserving user privacy, and enabling uninterrupted surveillance even under limited network connectivity. These characteristics make the proposed framework highly suitable for practical deployment in educational institutions, government organizations, hospitals, libraries, laboratories, and other secure facilities.

VII. CONCLUSION

This paper presented Phone Patrol, an intelligent deep learning-based surveillance framework developed to detect unauthorized mobile phone usage in restricted environments while automatically identifying the individuals responsible for policy violations. The proposed system integrates YOLOv5 for real-time mobile phone detection, YOLOv9 for facial recognition, Raspberry Pi 5 for edge-based processing, and the Twilio API for automated SMS notification, creating a unified surveillance solution capable of operating efficiently in mobile-free environments. Unlike conventional radio frequency-based detection methods, the proposed framework performs entirely image-based mobile phone detection, allowing reliable identification of mobile devices regardless of their communication status, airplane mode, or network connectivity.

The developed framework performs continuous monitoring by acquiring live video streams through a USB camera, preprocessing incoming frames, detecting mobile phones using YOLOv5, recognizing individuals through YOLOv9, verifying user identity using a registered database, and automatically generating SMS alerts whenever unauthorized mobile phone usage is identified. This integrated workflow enables immediate administrative response while reducing dependence on continuous manual surveillance. The deployment of the complete surveillance pipeline on the Raspberry Pi 5 further demonstrates the practicality of edge computing for real-time intelligent monitoring applications, providing low-cost implementation together with reduced communication latency and improved data privacy.

Experimental evaluation confirmed the effectiveness of the proposed framework for both object detection and facial recognition tasks. The YOLOv5 model achieved a Precision of 0.921, Recall of 0.977, mAP50 of 0.992, and mAP50–95 of 0.873, demonstrating highly accurate mobile phone detection under real-time operating conditions. Similarly, the YOLOv9 face recognition model produced a mAP50 of 0.994 and mAP50–95 of 0.814,

confirming reliable identification of individuals associated with detected mobile phone usage. These experimental results validate the capability of the proposed deep learning models to deliver accurate surveillance performance while maintaining efficient embedded deployment suitable for practical implementation.

The proposed framework offers several practical advantages for secure environments. Educational institutions can effectively minimize examination malpractice by automatically detecting unauthorized smartphone usage. Government organizations and corporate offices can strengthen information security by preventing unauthorized mobile phone access to confidential workspaces. Hospitals, research laboratories, libraries, conference facilities, and other restricted environments can similarly benefit from continuous intelligent monitoring that improves operational discipline while reducing manual supervision. Furthermore, the automated SMS notification mechanism significantly enhances administrative responsiveness by providing immediate alerts whenever policy violations occur.

Overall, the proposed Phone Patrol framework establishes an efficient, scalable, and intelligent surveillance solution for enforcing mobile-free policies through advanced computer vision and deep learning techniques. The integration of YOLOv5, YOLOv9, Raspberry Pi 5, and Twilio API enables accurate detection, reliable facial identification, real-time edge processing, and automated alert generation within a unified architecture. Consequently, the proposed system contributes to improving security, strengthening policy compliance, reducing human intervention, and supporting intelligent surveillance across a wide variety of restricted-access environments.

Future research may focus on expanding the framework by incorporating multi-camera surveillance, multi-object tracking, advanced facial recognition models, cloud-assisted synchronization, and larger real-world datasets to improve scalability and robustness. The integration of emerging technologies such as edge AI optimization, IoT sensor networks, and transformer-based vision models may further enhance detection accuracy, processing efficiency, and autonomous surveillance capabilities, enabling broader deployment of intelligent mobile-free monitoring systems across complex real-world environments.

REFERENCES

1. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 779–788.
2. G. Jocher et al., "YOLOv5 by Ultralytics," GitHub Repository, 2023.
3. C. Wang, H. Liao, Y. Wu, P. Chen, J. Hsieh, and I. Yeh, "YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information," 2024.
4. J. Deng, W. Dong, R. Socher, L. Li, K. Li, and L. Fei-Fei, "ImageNet: A Large-Scale Hierarchical Image Database," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), Miami, FL, USA, 2009, pp. 248–255.
5. A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," Communications of the ACM, vol. 60, no. 6, pp. 84–90, 2017.
6. R. Girshick, "Fast R-CNN," in Proc. IEEE Int. Conf. Computer Vision (ICCV), Santiago, Chile, 2015, pp. 1440–1448.
7. 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, 2017.
8. K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, no. 2, pp. 386–397, 2020.
9. D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," International Journal of Computer Vision, vol. 60, no. 2, pp. 91–110, 2004.
10. P. Viola and M. Jones, "Rapid Object Detection Using a Boosted Cascade of Simple Features," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), Kauai, HI, USA, 2001, pp. I-511–I-518.
11. F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A Unified Embedding for Face Recognition and Clustering," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 2015, pp. 815–823.
12. O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep Face Recognition," in Proc. British Machine Vision Conf. (BMVC), Swansea, U.K., 2015.
13. S. Z. Li and A. K. Jain, Handbook of Face Recognition, 2nd ed. London, U.K.: Springer, 2011.
14. U. Ataro, K. Yadeta, and T. Wondimu, "Design of Mobile Phone Detector Circuit for Silent and Flight Modes," in Proc. Int. Conf. Electrical Engineering Research, 2022.
15. C. Mbaocha, "Design and Implementation of an Intelligent Mobile Phone Detector," International Journal of Engineering Research, vol. 8, no. 4, pp. 155–160, 2019.
16. A. Abdulhamid, M. Ismail, and H. Ibrahim, "Computer Vision-Based Embedded Surveillance Using Raspberry Pi," International Journal of Computer Applications, vol. 183, no. 15, pp. 20–27, 2021.
17. M. Owayjan, A. Kashour, N. Al Haddad, M. Fadel, and G. Al Souki, "Face Recognition Security System Using Artificial Intelligence," International Journal of Advanced Computer Science and Applications, vol. 13, no. 5, pp. 480–488, 2022.