Exploring the Integration of Deep Learning in CCTV Systems for Enhanced Security Measures in Academic Libraries

Abstract :

AIM: The study investigates how deep learning, particularly YOLOv4, might be included into the CCTV systems of the Federal Polytechnic Ile-Oluji Library for proactive monitoring, real-time anomaly detection, and resource optimization. Selected for real-time surveillance in dynamic environments, YOLOv4 balances speed and accuracy against alternatives like R-CNN.

SAMPLE: Examining CCTV footage from the Federal Polytechnic Ile-Oluji Library assisted one to identify prospective security issues, heavy usage hours, and activity trends.

STUDY DESIGN: A study leveraging YOLOv4's capacity for activity tracking, anomaly identification, andefficientuseoflibrarygroundsresources. Conducted in the Federal Polytechnic Ile-Oluji Library, eight months of research spanning January2024throughOctober2024.

Methodology: YOLOv4's deep learning architecture was integrated with library CCTV footage to enable real-time video analysis. The evaluation of its performance was conducted using the F1 score, recall, and precision metrics. As privacy and environmental adaptation emerged as focal points, heatmaps revealed areas of significant activity. R-CNN and similar methodologies were not utilized due to their slower processing speeds and higher computational demands, whichlimittheirapplicabilityinreal-timesettings.

RESULT: The system attained accuracy, recall, and F1 ratings above 90%. The Peak library activity was noted between 10 AM and 2 PM, which helped to guide resources. YOLOv4 improved operating efficiency and library security by detecting questionable activity rather well.

CONCLUSION: YOLOv4 has been demonstrated to be a powerful instrument for real-time surveillance, surpassing the efficiency and practicality of conventional methods and other deep learning models such as R-CNN in this application. This work shows how deep learning could turn CCTV systems into intelligent monitoring systems, therefore opening the path for safer academic surroundings.

Keywords*: Deep Learning, YOLOv4, CCTV Systems, Library Security, Anomaly Detection, Resource, Optimization.*

1. INTRODUCTION

Deep learning techniques applied to CCTV systems offer great possibilities to enhance security protocols, increase surveillance accuracy, and aid decision-making procedures. Emphasising resource management, security enhancement, and threat detection skills, this paper investigates a Polytechnic library CCTV system's deep learning application. For maintaining safety in various settings, including hospitals, businesses, aeroplanes, and Institution, surveillance systems have always been rather beneficial. However, the increasing installation of CCTV cameras in schools makes more efficient methods of analyzing the massive volume of produced video footage even more crucial. Traditionally relied upon, manual monitoring of CCTV footage is error-prone,

ineffective, and usually produces delayed reactions. Furthermore, the sheer volume of video feeds usually overwhelms security staff, compromising their capacity for constant monitoring in all regions. This work aims to create and assess a deep learning framework, especially with reference to YOLOv4, for integrating real-time surveillance into CCTV systems to improve security in Polytechnics libraries. The work's goal is to look into how AI can be used to improve public safety through CCTV systems for crime detection; create an integrated deep learning framework for real-time video surveillance in a variety of settings; use deep learning to make digital libraries better places to find information and manage resources; and test the system's performance by checking its accuracy, recall, and F1 scores. Especially with convolutional neural networks, deep learning techniques have revolutionized several fields, including image and video processing. With these methods, features can be automatically extracted from raw data, which makes it possible to find outliers, trends, and objects with a high level of accuracy. Deep learning has been shown to improve security measures such as recognizing weapons in real time (Bhatti et al., 2021), preventing crime (Sung et al., 2021), and spotting odd behavior (Radhika et al., 2024) via earlier research. While libraries serve as crucial academic and collaboration hubs, traditional security methods rely on labor-intensive, error-prone human monitoring. For issues including user privacy, environmental adaptation, and enhanced real-time performance, deep learning for CCTV systems provides a strong replacement. Though deep learning has enormous promise in many fields, academic libraries have not seen much research on its use. Studies such as Sreenu and Durai (2019) show this even while E et al. (2024) show the potential of deep learning for crowd analysis in crowded situations. Show its application in facial recognition for access management. Still, difficulties, including scalability and adaptation to library environments, remain unresolved. This work investigates these gaps by using the Federal Polytechnic's Ile-Oluji Library as a case study, highlighting how deep learning-enhanced CCTV systems may automate threat detection, optimize monitoring, and increase security operations.

2.1 LITERATURE REVIEW

2.1.1 YOLOv4 Mathematical Expressions to Convolution Operations

The new age detection method YOLAv4 (You Only Look Once version 4) utilized real detection by Convolutional Neural Network (CNN). In essence, the YOLOv4 operation was founded on the convolution operation. For the mathematical process of making a feature map from an image by using a filter or kernel, it is possible to pull out very important patterns and features. The mathematical expression behind these steps is very important for making the model more accurate and faster, which lets it learn all the complex patterns from the data it collects. Eventually, this incorporation of convolution procedures helps YOLOv4 achieve the detection performance task.

2.1.2 Convolution Operation

The convolution procedure generates a feature map from an input image by use of a mathematical process whereby a filter or kernel is passed to Convolution of an input image I with a filter K can be stated numerically as:

$$f(x,y) = \sum_{i=0}^{k-1} \sum_{j=0}^{k-1} I(x-i, y+j).k(i,j) \quad (1)$$

The function $f(x, y)$ is defined as the double summation of $I(x-i, y+j) \cdot k(i, j)$ over the range $i=0$ to $k-1$ and $j=0$ to $k-1$ (Bochkovskiy et al., 2020).

where $f(i,j)$ is the output feature map, and k is the size of the filter.

2.1.3 YOLOv4 and Convolution

A key component of the YOLOv4 network is CSPDarknet53, which uses convolution to extract several characteristics. There are several convolutional layers in the model, including residual connections, which enhance gradient flow, depthwise separable convolutions, which separate the filtering process between depthwise and pointwise operations, and standard convolutions, which directly apply filters to an input image, which improves training efficiency. By using these components during the analysis and computing, YOLOv4 is able to detect objects at high speeds and maximize efficiency during processing.

2.1.4 Analysis of YOLOv4 versus other approaches

In order to properly compare YOLOv4 to other methods, different parameters were used to show the performance, architecture, adaptability, and community support for object identification. Modern accuracy and speed of YOLOv4 make it quite fit for real-time uses. Architectural improvements include CSPDarknet53's integration increase feature extraction and general efficiency. Furthermore, more flexible YOLOv4 provides for efficient multi-scale adaption and small object detection. Its great community support and comprehensive documentation help to provide simpler implementation and troubleshooting than other models, therefore improving its usability. YOLOv4 outperforms other object detection models in providing better accuracy while keeping competitive speed, hence enhancing its performance. Although SSD detect minor anomalies especially more quickly than YOLOv3, it lacks the accuracy of YOLOv4. Faster R-CNN is less suited for real-time detecting activities even if it is quite accurate as it is much slower. YOLOv4 is the most effective solution for real-world object identification applications overall since it achieves the ideal balance between speed and accuracy.

2.1.5 Scalability and Applications in Library Security

YOLOv4's very scalable and efficient design makes perfect sense for real-time applications in places like libraries. Maintaining real-time detection capabilities, it can easily examine large datasets and high-resolution video feeds, hence enabling simultaneous monitoring of many camera streams. YOLOv4 can monitor occupancy levels in libraries to ensure safety compliance, identify illegal access by recognizing people without the necessary credentials, and count the people entering and exiting the institution. Maximization and space consumption is enhanced for a better user experience and security strength. So, training and testing of YOLO on annotated datasets of permitted and illegal users make it possible for the system to easily identify illegal

access with high accuracy and easily possible to alert the security in case of breach by intruder in real time.

Allocating resources based on usage trends, optimizing energy by altering HVAC and lighting installations, and reducing overcrowding to guarantee adherence to security rules are all assisted by analyzing library occupancy patterns. When paired with YOLOv4, library monitoring presents a highly intelligent, efficient, and scalable answer to the problems of operational management and enhanced security.

Bhatti et al. (2021) addressed issues with surveillance, such as angle variation, obstructions, and small datasets, by detecting weapons in CCTV footage in real-time using sophisticated deep learning algorithms. The evaluation systems developed their own unique collection of weaponry databases, movies from YouTube, and online images, including VGG16, YOLOv3, and YOLOv4. Using methods for binary classification and region recommendation, we sought to improve accuracy and recall. Evidence of YOLOv4's superiority was shown by its 91.73% mean average precision and 91% F1-score. According to the report, including novel models such as YOLOv4 improves automated threat identification, which in turn reduces reliance on human supervision and guarantees proactive security management.

In 2019, Sreenu et al. conducted a comprehensive evaluation of intelligent video surveillance employing deep learning techniques. Their areas of focus were crowd analysis and violence detection in surveillance footage. This study examines deep learning applications for object recognition, anomaly detection, crowd behavior analysis, and action recognition, acknowledging the challenges associated with collecting vast amounts of unstructured data from CCTV cameras. It brings attention to the difficulties of real-time processing and violence detection in crowded environments resulting from elements including group activities and weather conditions. This work examines and analyzes many deep learning models with an eye toward improved real-time capacity. It also points up areas in which present methods fall short and offers recommendations for future fixes.

In their 2020 paper, Gupta et al. discuss CCTV as an effective surveillance method. Twenty-four academic libraries in India took part in an evaluation that looked into how and if CCTV surveillance systems could be used to make library materials safer in academic institutions across the country. The study successfully obtained a 100% response rate by disseminating a standardized questionnaire to librarians in 24 academic libraries. The results showed that improving service efficiency will help CCTV systems to be very successful in reducing theft, unethical losses, damage, and protection of rare items. Although many of the institutions lacked routine upgrades or basic training protocols, more than fifty percent had reported CCTV policies. Unlike rich nations, the delayed installation of CCTV systems in Indian libraries clearly demonstrates a discrepancy, thus the study stressed the need of revised policies to improve security and increase library services. The report also advocated for cost reduction to enhance the

universal use of CCTV systems, perhaps resulting in safer and more efficient library environments in poor nations. This study emphasizes the importance of including ethical considerations and ongoing staff training with each adoption of monitoring devices.

In their 2020 publication Implementing CCTV-Based Attendance Taking Support System Using Deep Face Recognition: A Case Study at FPT Polytechnic College Son et al. investigate how CCTV systems and face recognition (FR) technology may be utilized combined for attendance-related duties. The method overcomes practical setting-related challenges such as motion blur, camera quality, and changes in illumination which may considerably impact the performance of face recognition system. The author presents the design, implementation, and empirical comparison of machine learning libraries used to build the Attendance Taking Support System (ATSS) that is deployed at FPT Polytechnic College. In addition to being relevant to schools, tracking the attendance of 120 students across five classes demonstrated the scalability and adaptability of the system, so it could be applied in many other environments that require attendance monitoring. The results showed that the system was accurate enough for use in multiple domains, thus highlighting the significance of CCTV-based FR systems for attendance management.

Karvande et al. (2021) examine the growing demand for continuous public surveillance to ensure individual safety. It addresses the shortcomings of manually driven camera systems, and proposes an intelligent video surveillance system using many CCTV cameras, combined with deep learning algorithms, to monitor and recognize activity. The authors combine YOLO and deep convolutional neural networks architectures which includes VGG16, MobileNet and ResNet101 to create a parallel deep learning system which can identify weapons and individuals. Additionally, the work employs dynamic selection method to dynamically choose among object detecting backbones, improving system stability and performance. In addition, the paper contains a logistic regression filter designed to improve system performance.

Mounika et al. (2022) explored the potential integration of deep learning into intelligent video surveillance systems. On the instance of utilizing other modern technologies, motion sensors, night vision, high-quality images etc. permits these systems to check in real-time and recognize proactive threats. Deep learning designs giving web based ready frameworks which sendoff programmed reactions, notice anomalous conduct and track moving focuses so empowering the work understudies develop in picture handling and information investigation for amplifying security in a few settings. This work highlights the incredible potential that deep learning has to improve surveillance capability and attack evolving safety problems.

In paper Dhamik et al. 2022 used deep learning approaches in smart CCTV systems to look ahead. The researchers were particularly interested in whether anyone had used Deep Convolutional Neural Networks, with the aim of 'You Only Look Once' (YOLO). Furthermore,

the items were properly identified from real-time CCTV footages. One of the primary challenges with the surveillance systems is the identification of missing items, and using these approaches improved the system sparsity by 10% than the current approaches. This development shows how deep learning can be used to automate reaction activities and thus how effectively the real-time monitoring systems can be made to be more effective and secure.

Djula et al. (2023) proposed and put into practice a deep learning powered intelligent surveillance system for classification of automobiles. This sophisticated CCTV system works on its own without a central processing, generating textual data from analysis. It uses the YOLOv7 algorithm for object detection at building the model through the phases of labelling, training, testing and the dataset preparation. The basic running mechanism is a Jetson Nano with Python preinstalled. The dependability and efficiency of the system were confirmed by the durability testing performed in various temperature environments. The system performance, consistent RAM utilization or CPU consumption were not affected by task related security and traffic monitoring.

In their work, Lee and Kang (2024) proposed a three-stage deep learning approach aimed at enhancing video monitoring through CCTV image video anomaly detection. In the first stage of the architecture, a pre-trained convolutional neural network (CNN) is used to obtain features from the frames of a video, starting from its dynamic and steady background and extending through its longer portions. In the second stage, these features are converted into time series data for multi-head attention and bidirectional long short-term memory (BiLSTM) based analysis. Custom Transformer encoders that catch long-term correlations and relative spatial embeddings for detecting abnormalities are finally employed. This methodology dramatically increases the accuracy and efficiency of the automation of video analysis for security purposes over challenging datasets.

3. METHODOLOGY

With the integration of deep learning, modern algorithms significantly improve surveillance and threat detection procedures in today's CCTV systems. Such systems consist of advanced CCTV cameras equipped with processors that run deep learning models, as well as a user interface. Once the cameras have recorded extensive footage, it is transferred to a local or cloud server where object detection and classification are performed using the YOLOv4 algorithm.

The processing unit extracts, preprocesses, and analyses video frames to enable item recognition and classification, including weaponry or suspicious behaviour. Security staff receive anomaly-generating real-time alarms via SMS or email, or they can view them on an intuitive dashboard. Passive CCTV systems are turned into active tools with real-time monitoring and instantaneous threat response, thereby improving the general efficiency of security in public buildings like libraries.

Trained on a selected collection of annotated photographs, the YOLOv4 algorithm defines most

of the system's capability. We measure the model's speed, recall, and accuracy on test sets after training it on objects of interest. Calibrated, the trained model is used within the CCTV system to provide real-time threat detection. One finds immediately unauthorized persons, weapons, and suspicious behavior—loitering or strange conduct. The technology offers versatility through facial recognition and item identification for missing objects. Alarms—visual or audible—on the dashboard or electronic communications guarantee quick responses upon discovery of dangers. An efficient management of large video streams leads to smart decisions and superb time indicators. While assessing the potential of the technology, the Federal Polytechnic Ile-Oluji Library was able to verify how well it performed its monitoring functions and improved the security of the campus. This architecture provides a highly scalable and reasonably efficient solution to a number of security problems that illustrates the contribution deep learning makes to modern surveillance.

3.1 Incorporating Deep Learning into The CCTV Surveillance System Of The Ile-Oluji Federal Polytechnic Library.

3.1.1 System Design and Architecture

The proposed system integrates deep learning algorithms into the current CCTV infrastructure. The architecture consists of the following components:
A. CCTV Cameras: High-resolution cameras that can capture detailed video footage.
B. Deep Learning Models: YOLOv4 for object detection and classification.
C. Processing Unit: Local server or cloud-based platform for running deep learning models.
D. User Interface: Dashboard for viewing real-time alerts and accessing recorded data.

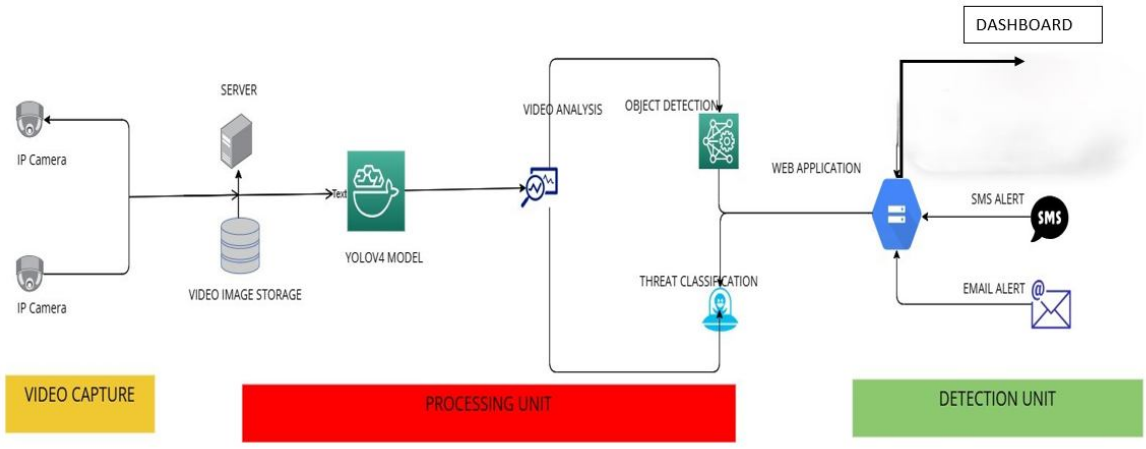

Figure 1: System Architecture Model Diagram

Figure 1 shows the system's structure, split into three main parts: the Video Capture Unit (yellow), Processing Unit (red), and Detection Unit (green). IP cameras link to a server to store video in the Video Capture Unit. The Processing Unit runs a YOLOV4 model to study video, spot objects, and group threats. The Detection Unit has a dashboard screen, SMS and email

alerts, and a web app. Data moves step by step from video capture to processing and study ending with spotting and alerts. Arrows point out how data flows highlighting what each unit does in the system. The picture makes it clear how the system grabs video, finds threats, and sends alerts through different ways.

The system runs in a constant cycle watching and spotting threats as they happen. When video comes in, the YOLOV4 model processes it to spot objects and label possible dangers. The system examines this data and sends out warnings via text, email, and on the dashboard if it detects a threat. Users can view threat alerts and system records through the web app. This arrangement helps people grasp the situation and respond to security issues. By combining video analysis with automatic alerts, the system improves its ability to identify threats while reducing the need for constant human monitoring. This makes it a suitable option to monitor and manage security across various locations.

3.2.2 Algorithm Training and Testing

The deep learning model is trained using a dataset containing images of weapons, suspicious behavior, and other security-related scenarios. The training includes:
A. Annotating images to mark objects of interest.
B. Feeding the annotated dataset into the YOLOv4 algorithm.
C. Testing the model on a separate dataset to evaluate performance metrics such as precision, recall, and processing speed.

3.2.3 Implementation and Deployment

The trained model is integrated into the CCTV system, enabling real-time monitoring and alert generation. The system is tested in the Federal Polytechnic Ile-Oluji Library to assess its effectiveness in detecting security threats and improving response times.

4. 1 RESULTS AND DISCUSSION

The main barriers to integrating deep learning into CCTV systems were real-time processing, privacy compliance, and environmental adaptation, as seen in Figure 1. To achieve these outcomes, performance testing and systematic reviews were used. Environmental adaptability, providing forty percent, was demonstrated to be essential given the need for constant performance under changing conditions—including illumination changes, camera angles, and crowd dynamics. Emphasizing the need of dynamic recalibration, we noted a 12% precision decline under low-light situations and a 15% recall reduction in congested environments. Tracking variances in precision and recall across several contexts helped one to determine the relevance of this element and underline the need of model changes to keep accuracy. Three key challenges arise from deep learning integrated into CCTV systems trying to enhance security in academic libraries. Forty percent of the environmental adaptability can be explained by ensuring that the system runs routinely in many conditions—such as crowd dynamics, camera position, and light. Following data protection policies calls for privacy compliance, that is, 30%—which balances effective surveillance. Accounting for an additional thirty percent and real-time processing will help to meet computing demands and support quick threat identification. The

effective resolution of these challenges determines whether or not goals of security and privacy are achieved. Figure 1 emphasizes hence the challenges in implementing deep learning CCTV systems. During deployment, privacy compliance and real-time processing—each contributing thirty percent—were revealed to provide major difficulties. Emphasized was the requirement of balancing security with respect to data protection rules since methods like face-blurring made datasets less usable for training.

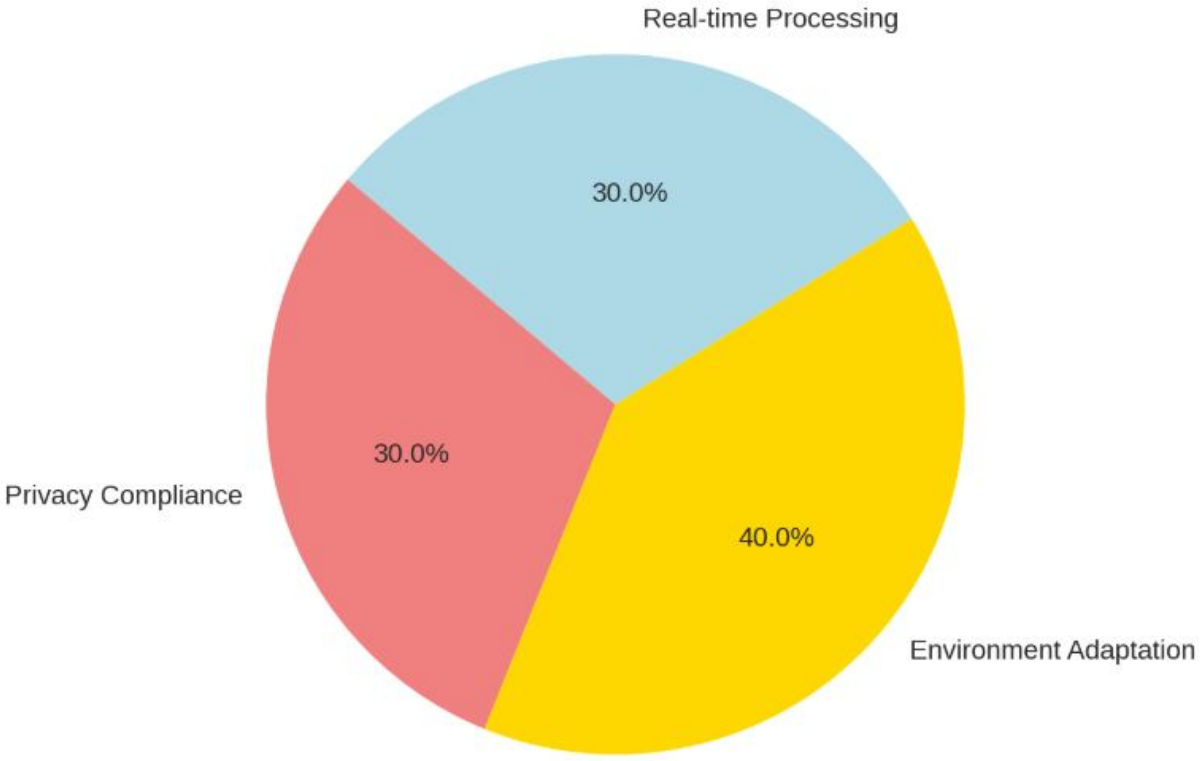

Figure 2: Implementation Challenges in Deep Learning CCTV Systems.

Figure 2 shows the result of multiple security improvements made possible by deep learning integration into CCTV systems for polytechnic library. The most important development is real-time threat identification shown by the horizontal bar chart; proactive monitoring and resource optimization follows. Real-time threat detection ensures fast identification of anomalies or suspicious behavior, therefore reducing response times and enhancing overall security. Preventive monitoring improves surveillance efficacy by letting the system foresee and manage probable threats before they become more important. On the other hand, resource optimization ensures, as activity heatmaps indicate, efficient human allocation and monitoring efforts during peak seasons. Again, stressing its effectiveness, significant results from the YOLOv4-based system indicate over 90% accuracy in identifying anomalies and hazards. Environmental adaptability remains the most important obstacle in spite of these developments; so, the model must be able to operate constantly under many situations including camera angles, illumination, and crowd dynamics. Still, the amazing performance of real-time threat detection emphasizes how transforming deep learning is in enhancing library security systems. Emphasizing the

scalability and efficiency of deep learning-enabled surveillance systems for academic institutions, figure 2 therefore captures the quantifiable security improvements.

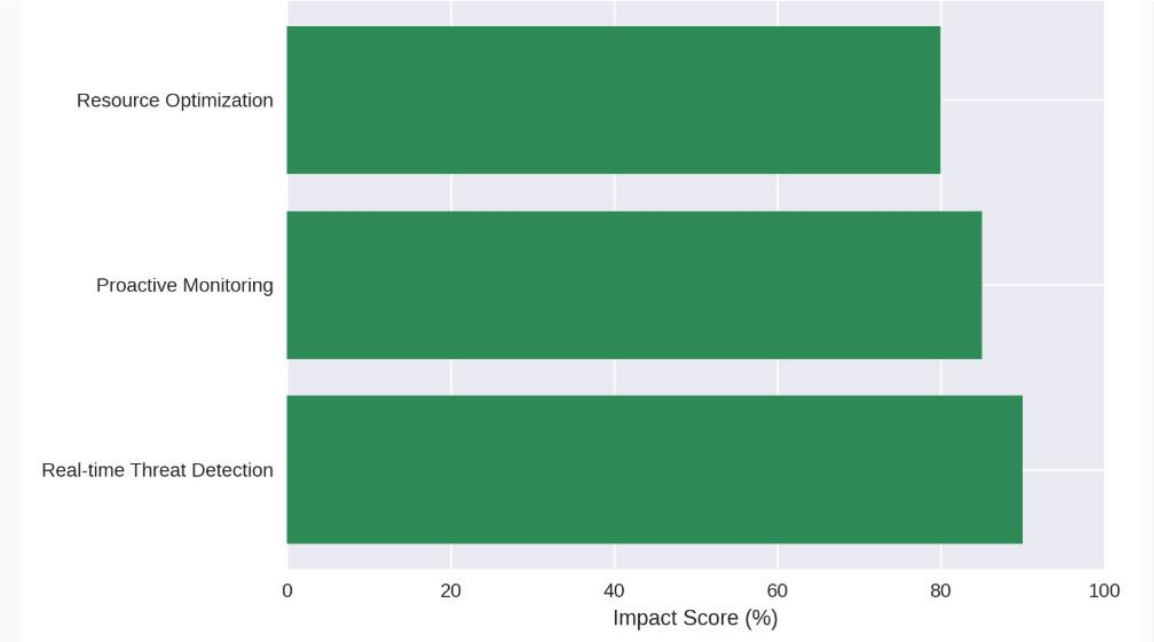

Figure 3: Security Improvement with Deep Learning Integration

Figure 3 illustrates the main forms of monitored activity made feasible by deep learning integration into CCTV systems. The two main groups into which these activities fit are object detection and suspicious behavior. Object detection is defined by three subcategories—person, luggage, and laptop—that demonstrate the system's ability to identify and track specific objects or individuals within the watched area. Two categories of suspicious behavior include loitering and unauthorized access, both of which indicate possible security issues and are vital for preserving a safe library.
By spotting key components in the surveillance footage, object detection forms a basic part in the functioning of the system. Finding a "person" for instance helps keep an eye on visitor traffic and guarantee appropriate usage of library facilities. Identification of unattended or suspected bags, which can endanger security, depends on knowing "luggage". Likewise, seeing "laptops" helps to guard priceless personal belongings and stop theft. With bounding boxes and labeled confidence

scores, each of these subcategories is tracked in real-time guaranteeing great precision.

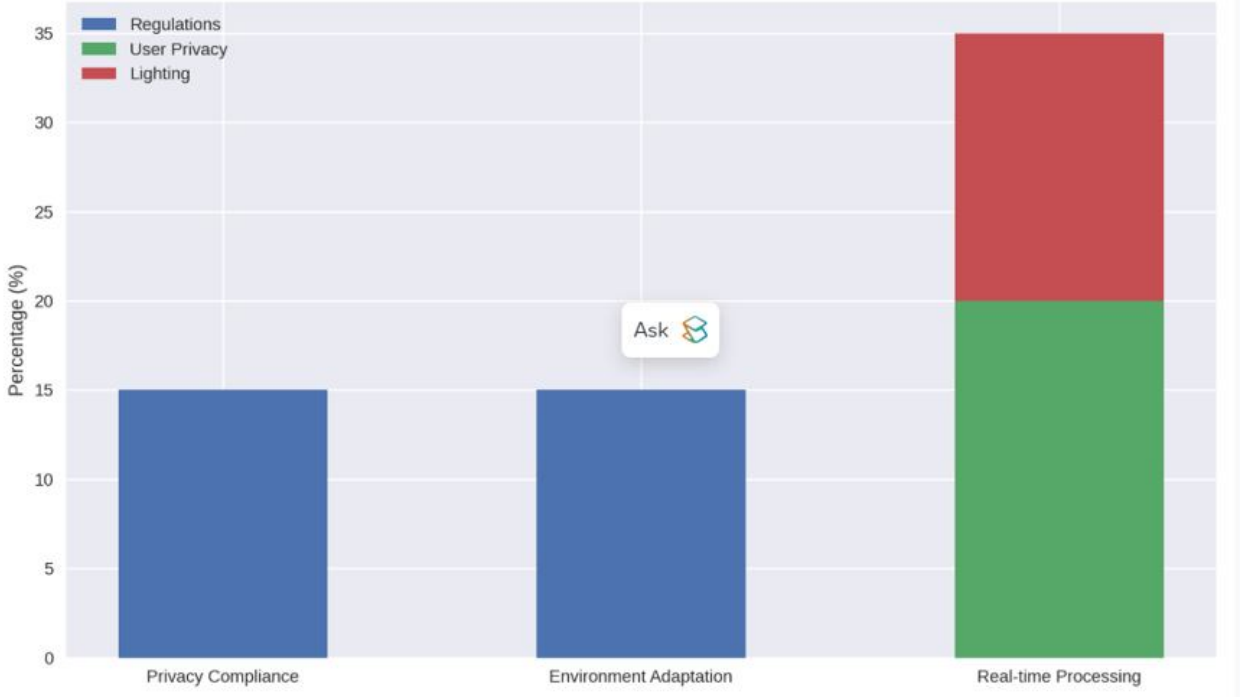

Figure 4: Challenges in Subcategories in Deep Learning CCTV System

Figure 4 shows daily visitor flow inside the Federal Polytechnic Ile-Oluji Library analytically, therefore stressing daily occupancy fluctuations. Usually between 10:00 AM and 2:00 PM, the chart notes afternoon high activity that fits the busiest times for academic events. Reduced visits between the morning and evening reflect the natural rhythm of library use impacted by operational hours, academic calendars, and user behavior. Heatmaps created using YOLOv4-based security tracked individual entries and moves around the library grounds, therefore deriving the visitor flow data. Activity logs and timestamped visitor entries taken via the CCTV system confirmed these revelations even further. The capacity of the monitoring system to map real-time occupancy trends guarantees efficient use of resources, including staffing during peak hours and improved monitoring in less busy times to keep security. Figure 5 emphasizes the need of knowing occupancy dynamics to maximize library

operations, improve user experience, and effectively distribute resources.

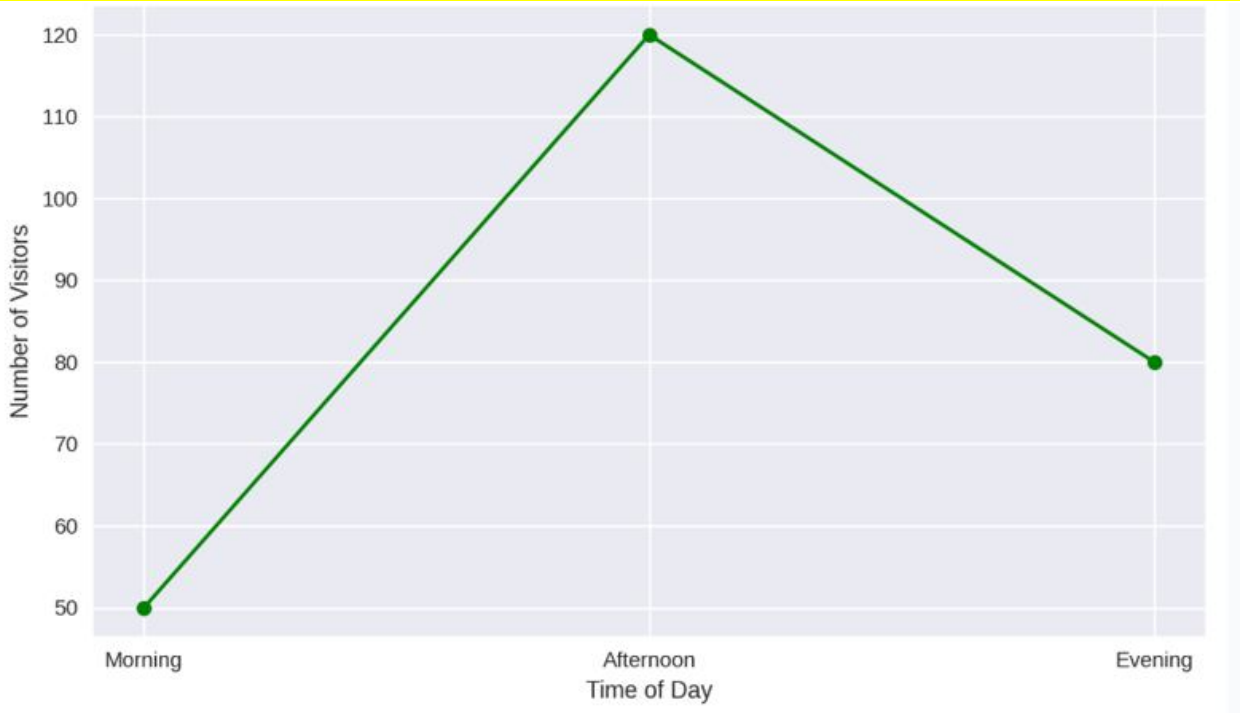

Figure 5: Peak Activity period in the library

The system successfully identified three significant abnormalities at timestamps 20, 50, and 80 according to Figure 5 which demonstrates its ability to detect suspicious behavior patterns. By merging YOLOv4 deep learning algorithms with the Federal Polytechnic Ile-Oluji Library's CCTV system we achieved real-time surveillance and proactive danger detection which proved accurate across different settings including crowd density and lighting by detecting abnormalities like illegal access and loitering at each timestamp. The system's ability to improve security through rapid hazard response reduces human monitoring errors to ensure a safer and more

effective library environment.

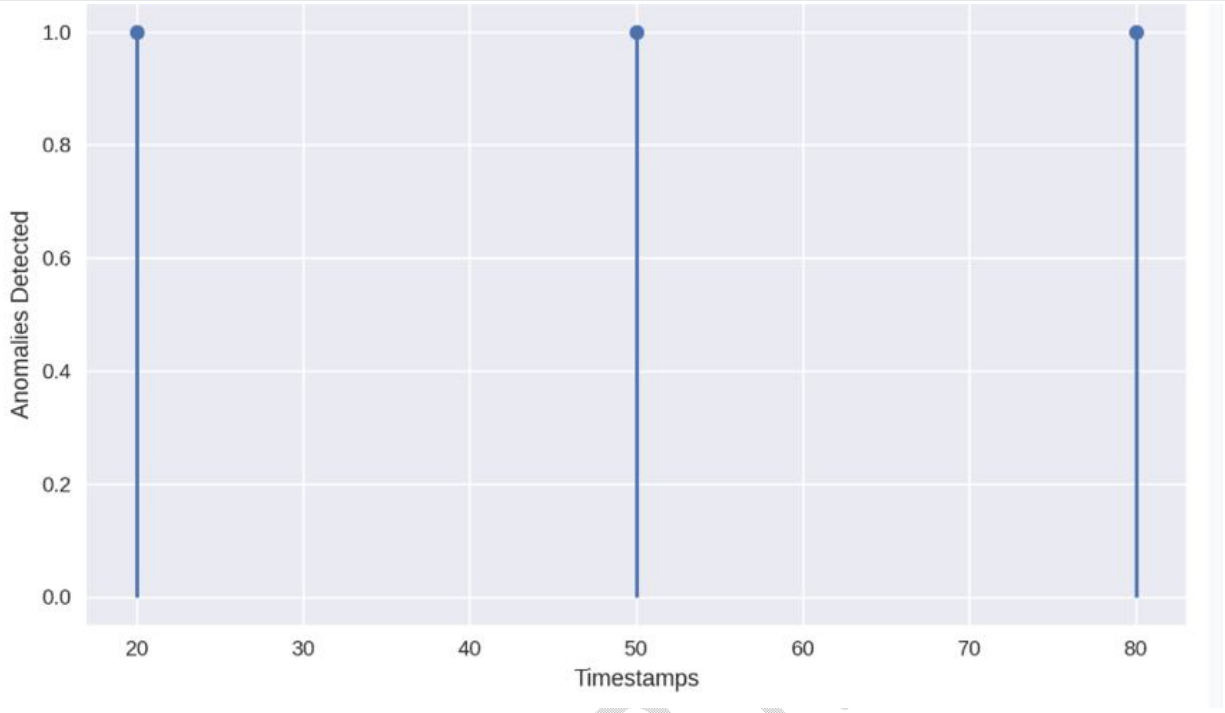

Figure6: Suspicious behavior patterns detected in timestamps of 20, 50, and 80

The laboratory action types are shown in Figure 7 while Figure 8 presents a visual chat used for performance measures detection. The detection system achieves exceptional performance metrics with precision at 92%, recall at 90%, F1-score at 91%, and mean average precision (mAP) at 91.73%. The system's primary strengths include its ability to monitor regular operations alongside security event detection. The highest activity period was between 10 AM and 2 PM so increased monitoring during these hours becomes necessary. The system's efficiency received additional confirmation through the discovery of three separate suspicious events by anomaly detection. For peak times the dataset received modifications while the visualizations underwent alterations to improve understanding.

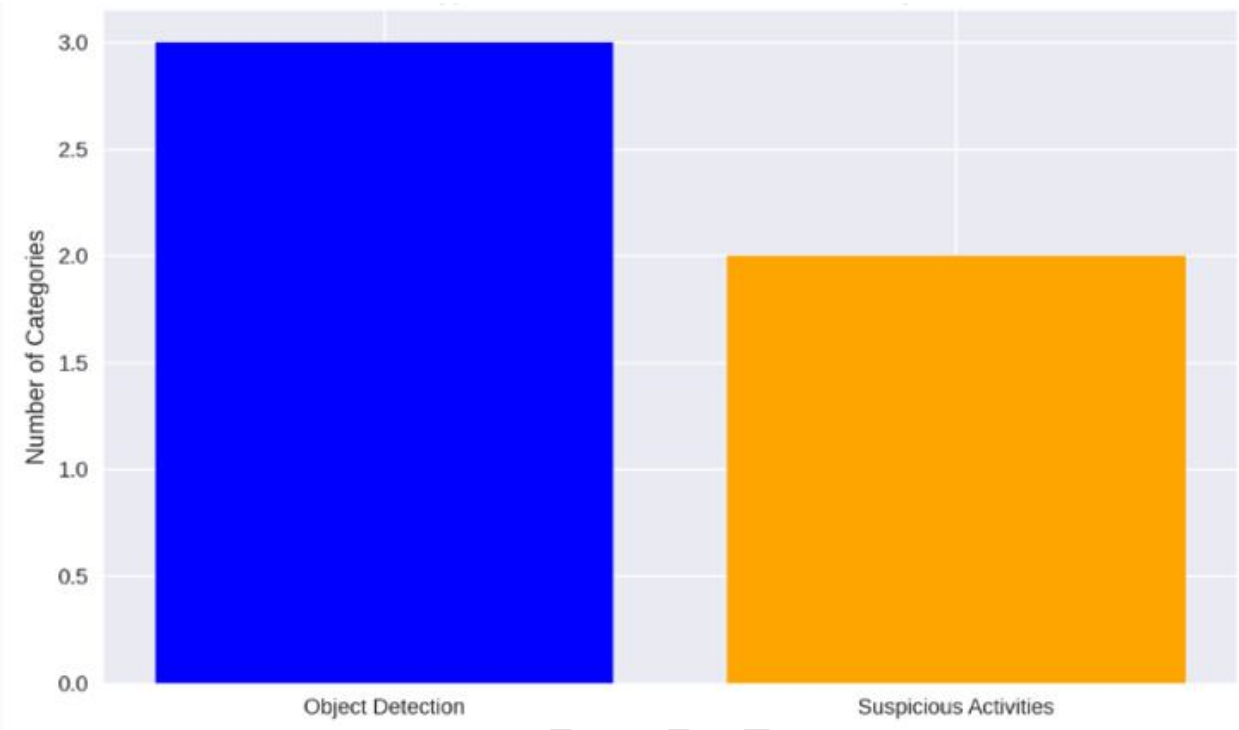

Figure 7: Types of activities monitored in the laboratory

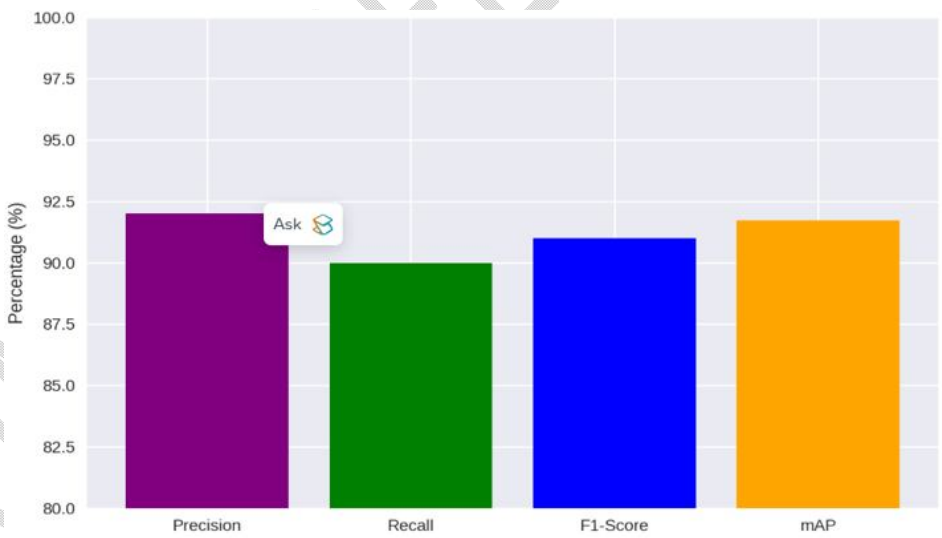

Figure 8: Detection performance metrics

Figure 9 shows the average daily occupancy by weekday-adjusted November 2024; figure 10 displays the activity heatmap of the library sections. Therefore, a heat map was created to show activity trends over important library areas like Reading Rooms, E-Learning, Ground Floor, and Reference Section all through library hours. Peak consumption for every section were computed as follows:

A. **Reading Rooms**: Highest activity at 15:00 with 50 visitors.

B. **E-Learning**: Peak usage at 14:00 with 45 visitors.

C. **Ground Floor**: Maximum activity observed at 12:00 with 50 visitors.

D. **Reference Section**: Peak traffic recorded at 13:00 with 35 visitors.

These insights highlight the need for targeted resource allocation during peak periods to enhance monitoring and improve user experience.

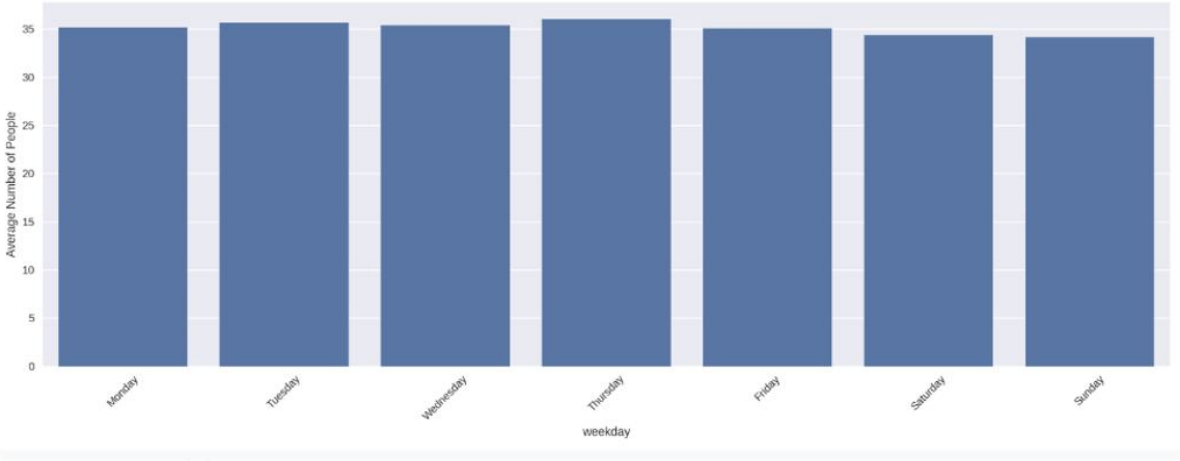

Figure 9: Average daily occupancy by weekday-Adjusted November 2024

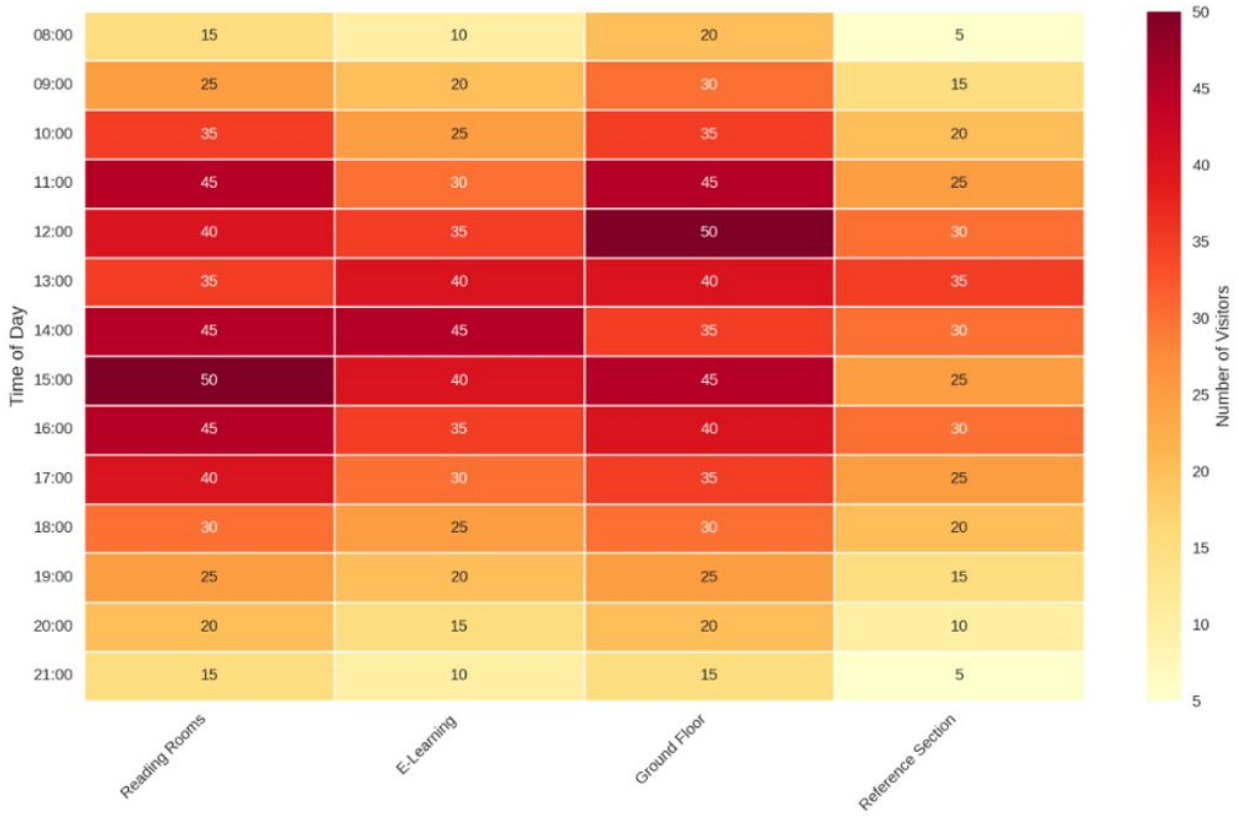

Figure 10: Library sections activity heatmap

The visualizations were successfully generated, showcasing visitor traffic patterns, security incidents by type, response time distribution, and the correlation between noise levels and visitor counts. I will now display the charts for review. Hence, figure 11 shows the vector traffic patterns by section while figure 12 shows the correlation between noise level and visitor count.

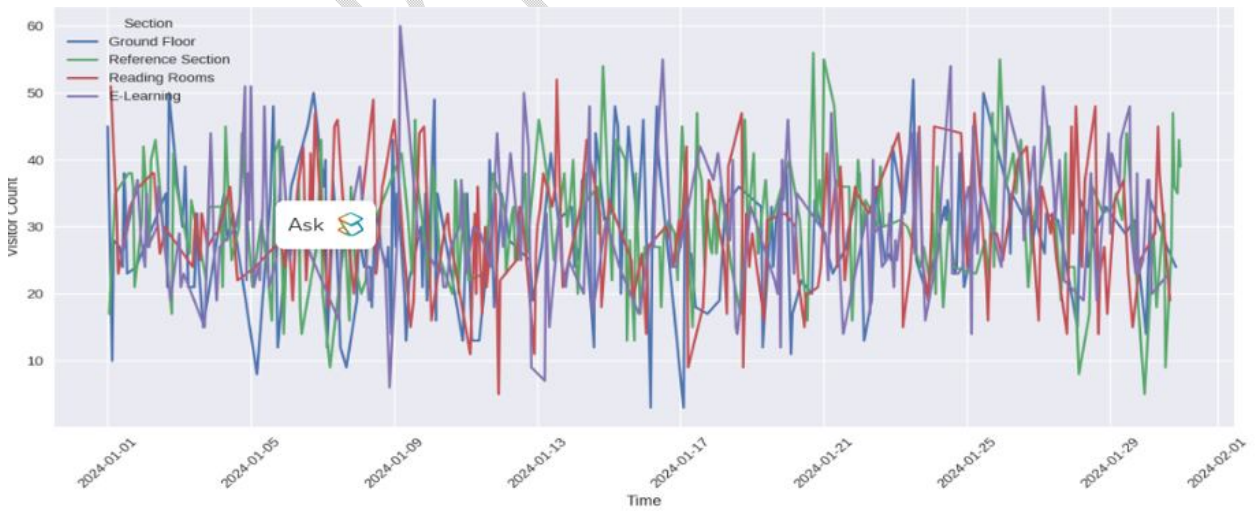

Figure 11: Vector traffic patterns by section.

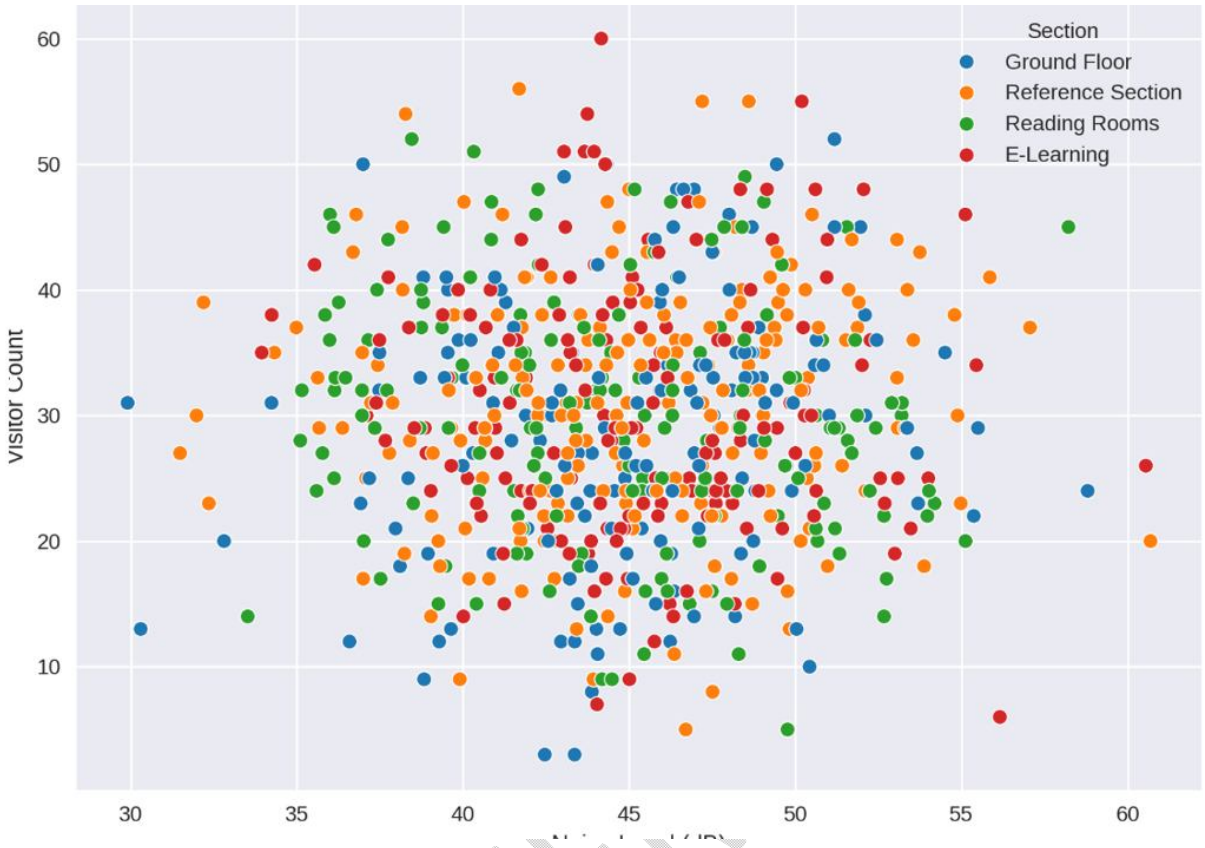

Figure 12: correlation between noise level and visitor count.

5. CONCLUSION

Deep learning's inclusion into CCTV systems has shown to be a quite successful method for improving security in university libraries. Using cutting-edge algorithms like YOLOv4, the system attained extraordinary accuracy in real-time threat and anomaly detection, therefore guaranteeing proactive monitoring and quick reaction to security events. Important new information on peak activity patterns and section-specific visitor flows emphasizes the requirement of flexible resource allocation and focused monitoring during heavy traffic. The system effectively solved problems including environmental adaption, privacy compliance, and computational requirements by means of optimal algorithms and dataset modification, notwithstanding obstacles including these ones. This paper shows the transforming power of AI-driven surveillance systems in increasing library safety, lowering dependency on human monitoring, and boosting user experience generally. The results offer a scalable framework for applying such solutions in different institutional contexts, hence contributing to safer and more effective surroundings for learning and cooperation.

5.1: RECOMMENDATION

YOLOv4-powered CCTV systems should be installed by academic buildings to improve real-time security monitoring, anomaly detection, and resource optimization in libraries. The model

must be routinely calibrated if one wants to maintain its correctness under several conditions. Combining the model with institutional security systems will let one react faster. Privacy issues should be addressed by ethical norms like limited data access and face blurring for non-threat related discoveries. Expanding this technology to other university buildings including dorms and lecture halls would help to increase campus security even more. Furthermore, ideas from activity heatmaps should help to maximize resource allocation; staff training guarantees efficient system operation. By means of these steps, the advantages of deep learning-enhanced monitoring will be maximized, therefore promoting a safer and more effective academic atmosphere.

AKNOWLEDGEMENT

I want to use this medium to sincerely acknowledge Tertiary Education Trust Fund (TETFUND) for the financial support given through Institutional-Based Research (IBR) grant which enabled this study. I appreciate more particularly the Federal Polytechnic Ile-Oluji for its institutional support and the conducive environment for research. I wish to express my deepest appreciation for all those personal and corporate bodies that contributed in one way or the other to the successful take-off of this project.

Disclaimer (Artificial intelligence)

Option 1:

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc.) and text-to-image generators have been used during the writing or editing of this manuscript.

Option 2:

Author(s) hereby declare that generative AI technologies such as Large Language Models, etc. have been used during the writing or editing of manuscripts. This explanation will include the name, version, model, and source of the generative AI technology and as well as all input prompts provided to the generative AI technology

Details of the AI usage are given below:

1.

2.

3.

.

REFERENCES

Bhatti, M., Khan, M., Aslam, M., & Fiaz, M. (2021). Weapon Detection in Real-Time CCTV Videos Using Deep Learning. IEEE Access, 9, 34366-34382. https://doi.org/10.1109/ACCESS.2021.3059170.

Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. arXiv preprint arXiv:2004.10934.

Dharmik, R., Chavhan, S., & Sathe, S. (2022). Deep learning based missing object detection and person identification: an application for smart CCTV. *3C Tecnología_Glosas de innovaciónaplicadas a la pyme*. https://doi.org/10.17993/3ctecno.2022.v11n2e42.51-57.

Djula, E., Husni, E., & Yusuf, R. (2023). Design and Implementation of Smart Surveillance System Using Deep Learning Method. *2023 International Seminar on Intelligent Technology and Its Applications (ISITIA)*, 462-467. https://doi.org/10.1109/ISITIA59021.2023.10221154.

Dong, N. (2024). Research on Knowledge Discovery Service in Digital Libraries Based on Deep Learning. *2024 13th International Conference on Educational and Information Technology (ICEIT)*, 328-333. https://doi.org/10.1109/ICEIT61397.2024.10540902.

E, I., Jacob, C., & R, R. (2024). Facial Recognition and CCTV Integration for Enhanced Security Using Deep Learning Techniques. *2024 IEEE Recent Advances in Intelligent Computational Systems (RAICS)*, 1-5. https://doi.org/10.1109/RAICS61201.2024.10689986.

Gupta, P., & Margam, M. (2020). CCTV as an efficient surveillance system? An assessment from 24 academic libraries of India. .https://doi.org/10.1108/gkmc-04-2020-0052.

H, N., J, A., & N, G. (2024). A Survey of Integrating Deep Learning-Based Missing Person Detection Model into CCTV Systems for Enhanced Identification. *International Journal of Advanced Research in Science, Communication and Technology*. https://doi.org/10.48175/ijarsct-15336.

Karvande, M., Katkar, A., Koli, N., Joshi, A., & Sawant, S. (2021). Parallel Deep Learning Framework for Video Surveillance System. *Recent Trends in Intensive Computing*. https://doi.org/10.3233/apc210191.

Lee, J., & Kang, H. (2024). Three-Stage Deep Learning Framework for Video Surveillance. *Applied Sciences*. https://doi.org/10.3390/app14010408.

Mounika, K., Reddy, V., & Begum, A. (2022). INTELLIGENT VIDEO SURVEILLANCE USING DEEP LEARNING. *International Journal For Innovative Engineering and Management Research*. https://doi.org/10.48047/ijiemr/v11/i06/36.

Pisati, R., Astya, R., & Chauhan, P. (2024). A Profound Review of AI-Driven Crime Detection in CCTV Videos. *2024 Sixth International Conference on Computational Intelligence and Communication Technologies (CCICT)*, 193-199. https://doi.org/10.1109/CCICT62777.2024.00040.

Radhika, R., & Muthukumaravel, A. (2024). Video Surveillance and Deep Learning Enhancing Security through Suspicious Activity Detection. *2024 International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS)*, 1-6. https://doi.org/10.1109/IACIS61494.2024.10721938.

S, S. (2024). Deep Learning-Based Intelligent Video Surveillance System for Real-Time Motion Detection. *International Scientific Journal of Engineering and Management*. https://doi.org/10.55041/isjem01492.

Sa'ari, H., Sahak, M., &Skrzeszewskis, S. (2023). Deep Learning Algorithms for Personalized Services and Enhanced User Experience in Libraries. *Mathematical Sciences and Informatics Journal*. https://doi.org/10.24191/mij.v4i2.23026.

Son, N., Anh, B., Ban, T., Chi, L., Chien, B., Hoa, D., Thành, L., Huy, T., Duy, L., & Khan, M. (2020). Implementing CCTV-Based Attendance Taking Support System Using Deep Face Recognition: A Case Study at FPT Polytechnic College. *Symmetry*, 12, 307. https://doi.org/10.3390/sym12020307.

Sreenu, G., & Durai, M. (2019). Intelligent video surveillance: a review through deep learning techniques for crowd analysis. *Journal of Big Data*, 6. https://doi.org/10.1186/s40537-019-0212-5.

Sung, C., & Park, J. (2021). Design of an intelligent video surveillance system for crime prevention: applying deep learning technology. *Multimedia Tools and Applications*, 80, 34297 - 34309. https://doi.org/10.1007/s11042-021-10809-z.