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

Abstract :
Aims: Deep learning technology, especially YOLOv4, is being tested and put to use in the CCTV systems of the Federal Polytechnic Ile-Oluji Library to help with real-time threat detection, proactive monitoring, and resource optimization.

Sample The subject of study was CCTV footage from the Federal Polytechnic Ile-Oluji Library, with an eye towards library activity patterns, peak usage hours, and possible security issues.

Study Design: A case study utilizing YOLOv4 deep learning algorithms for real-time anomaly detection, activity tracking, and optimized resource management within the library premises.

Place and Duration of Study: The study was conducted at the Federal Polytechnic Ile-Oluji Library over an eight-month period, from September 2024 to May 2025.

Methodology: To handle real-time footage, the CCTV system was linked with YOLOv4 deep learning algorithms. System performance was assessed with reference to accuracy, recall, and F1 scores. Heat maps shown areas of maximum activity and library use. To improve dependability and efficiency, the paper tackled issues including environmental adaption, real-time processing, and privacy compliance assurance.

Results: With accuracy, recall, and F1 scores above 90%, the system proved to be very outstanding. Emphasizing resource allocation flexibility, heat maps found peak activity times—10 AM–2 PM. Effective detection of anomalies and dubious behavior helped to increase operational efficiency and library security.

Conclusion: Deep learning's inclusion into the CCTV systems at Federal Polytechnic Ile-Oluji Library improved security, monitoring, and human error minimization. More effective learning environments in educational institutions and a strong answer for building safer and secured libraries and their resources this scalable system presented.

***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 university library CCTV system's deep learning application. For maintaining safety in various settings, including hospitals, businesses, aeroplanes, and universities, 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 university 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. 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.

1.1 LITERATURE REVIEW

Bhatti et al., (2021) used state-of- the-modern deep learning algorithms for real-time weapon detection in CCTV footage in order to tackle surveillance challenges including angle variation, occlusions, and limited datasets Testing systems including VGG16, YOLOv3, and YOLOv4 produced a custom dataset including weapon databases, YouTube videos, and internet images. Under a focus on accuracy and recall, binary classification and region suggestion techniques were applied. Showing a 91% F1-score and 91.73% mean average precision, results revealed YOLOv4's superiority The article argues that including creative models like YOLOv4 improves automated threat identification, hence lowering reliance on human supervision and guaranteeing proactive security management.

Focusing on crowd analysis and violence identification in surveillance footage, Sreenu et al., (2019) thoroughly reviewed intelligent video surveillance utilizing deep learning approaches. Understanding the difficulties in organizing unstructured huge data from CCTV streams, the work investigates deep learning uses for crowd behavior analysis, anomaly detection, object and action recognition, It draws attention to the challenges of real-time processing and violence detection in crowded settings resulting from factors including group activities and weather conditions. Emphasising improved real-time capabilities, the study contrasts many deep learning models, points out shortcomings in present approaches, and suggests future directions for overcoming these problems.

Gupta et al. 2020, in their paper CCTV as an effective surveillance system? The usage and possibilities of CCTV surveillance systems in improving the security of library materials in academic institutions all throughout India were investigated by an evaluation including 24 academic libraries of India. Gathering a 100% response rate, the study employed a standardized questionnaire sent to librarians from 24 academic libraries. The results showed that while increasing service efficiency, CCTV systems were judged to be rather helpful in managing theft, unethical losses, damage, and securing rare items. Over half of the institutions had written CCTV policies, even if many lacked necessary training policies or regular updates. The study underscored the need of updated rules to improve security and enhance library services since the late implementation of CCTV equipment in Indian libraries compared to rich countries clearly shows a difference. The study also recommended cost reducing for more universal acceptance of CCTV systems, which might lead to safer and more efficient library environments in developing countries. This study underlines the significance of include ethical concerns and continuous staff training with every implementation of monitoring technologies.

In their 2020 paper 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 might be coupled with face recognition (FR) technology for uses connected to attendance. The approach addresses real-world setting related problems such motion blur, camera quality, and fluctuating illumination that can greatly impact facial recognition system performance. The authors provide the design, implementation, and empirical comparison of machine learning libraries for producing an Attendance Taking Support System (ATSS) applied at FPT Polytechnic College. The system demonstrated scalability and adaptability by tracking the attendance of 120 students over five classes, hence usable not only in schools but also in many other situations needing attendance monitoring. Emphasising the efficiency of CCTV-based FR systems in attendance control, the results revealed that the accuracy of the system was suitable for various scenarios.

Karvande, et al. (2021) in their article on the growing need of continuous surveillance in public areas to guarantee individual security addresses It draws attention to the shortcomings of hand-operated camera systems and suggests an intelligent video surveillance system using several CCTV cameras and deep learning techniques for activity monitoring and identification. The Parallel Deep Learning Framework the authors provide combines deep learning methods including YOLO, Convolutional Neural Networks, and backbones including VGG16, MobileNet, and ResNet101 to detect weapons and humans. Furthermore, included in the article is a Dynamic Selection Algorithm dynamically selecting between object detection backbones to maximize system stability and performance. To improve system performance also a logistic regression filter is applied.

Emphasising its critical relevance in modern public safety, Mounika et al. (2022) investigated how deep learning might be included into intelligent video surveillance systems. These systems use cutting-edge technologies for real-time monitoring and proactive threat identification including motion sensors, night vision, and high-definition imaging. Deep learning models allow tracking of moving targets, identification of questionable behaviour, and start of automated responses—including web-communication alerting systems. The work emphasizes developments in data processing and computer vision to maximize security in several environments. This paper emphasizes how transforming deep learning can be in improving surveillance capacity and handling changing safety concerns.

In 2022 Dharmik et al. looked explored deep learning-based methods for people identification and missing object detection in smart CCTV systems. More particularly, the work aimed on successful object detection using deep convolutional neural networks—more especially, the YOLO (You Only Look Once) method—processing real-time CCTV data. Using these methods, the system demonstrated a 10% sparsity improvement over current solutions in missing item detection—a crucial activity in surveillance systems. Faster response actions made feasible by this development highlight how much deep learning might perhaps boost security and efficiency in real-time monitoring systems.

Specifically with regard to vehicle classification, Designed and put into use by Deep learning-based Smart Surveillance System Djula et al. (2023) Unlike conventional systems driven by centralized processing, this smart CCTV system runs independently creating textual data from the analysis. Driven on a Jetson Nano running Python, the model runs object detection utilizing the YOLOv7 algorithm across development including labelling, dataset preparation, training, and testing. Tests of durability exposed the dependability and efficiency of the system under several environmental settings. For jobs like traffic monitoring and security, the system displayed suitable for CPU use, constant RAM consumption, and good performance.

Lee and Kang (2024) presented a three-stage deep learning method for effective anomaly recognition in CCTV images, hence improving video monitoring. With a pre-trained convolutional neural network (CNN) to extract features from video frames, the architecture begins with targeting dynamic backgrounds and prolonged video sequences. Under bidirectional long short-term memory (BiLSTM) and multi-head attention in the second stage, these features are transformed into time series data under investigation. Finally, relative spatial embeddings and a proprietary Transformer encoder catch long-term correlations to identify anomalies. Tested on difficult datasets, this method significantly advances automated video analysis for security applications since it dramatically improves accuracy and efficiency.

2. MATERIAL AND METHODS

Modern deep learning algorithms enhance surveillance and threat detection capabilities when included in present CCTV systems. Architecture consists of high-end CCTV cameras, a processor, deep learning models, and a user interface. After cameras capture extensive video, it is delivered to a local server or cloud-based platform where object detection and classification using the YOLOv4 algorithm is done. 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. Good handling of large video streams produces intelligent outcomes and excellent response times. Testing the technology, the Federal Polytechnic Ile-Oluji Library found how effectively it detected security concerns and enhanced campus safety. This architecture provides a scalable and effective answer to many security issues, highlighting the function deep learning performs in contemporary surveillance.

2.1 Integrating Deep Learning in Federal Polytechnic Ile-Oluji Library CCTV System

1. System Design and Architecture

The proposed system incorporates deep learning algorithms into the current CCTV infrastructure. The following components comprise the architecture:

A. **CCTV Cameras**: High-resolution cameras capable of capturing detailed video footage.
B. **Deep Learning Models**: YOLOv4 is employed for object detection and classification.
C. **Processing Unit**: A local server or cloud-based platform for running deep learning models.
D. **User Interface**: A dashboard for displaying real-time alerts and accessing recorded data.

2. Algorithm Training and Testing

The deep learning model is trained using a dataset containing images of weapons, suspicious behaviors, and other security-related scenarios. Training involves:

A. Annotating images to label 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. 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.

3. RESULTS AND DISCUSSION

Three key challenges exist when deep learning is included into CCTV systems to increase security in academic libraries. Forty percent of the environment adaptability is accounted for by ensuring the system functions consistently under various circumstances including crowd dynamics, camera location, and illumination. Contributing thirty percent to privacy compliance calls for careful adherence to data protection rules balancing efficient surveillance. Real-time processing accounts another thirty percent since computing needs must be satisfied to enable instant threat identification. Reaching goals related to security and privacy depends on overcoming these obstacles. Figure 1 thus highlights the difficulties in deep learning CCTV systems' deployment.

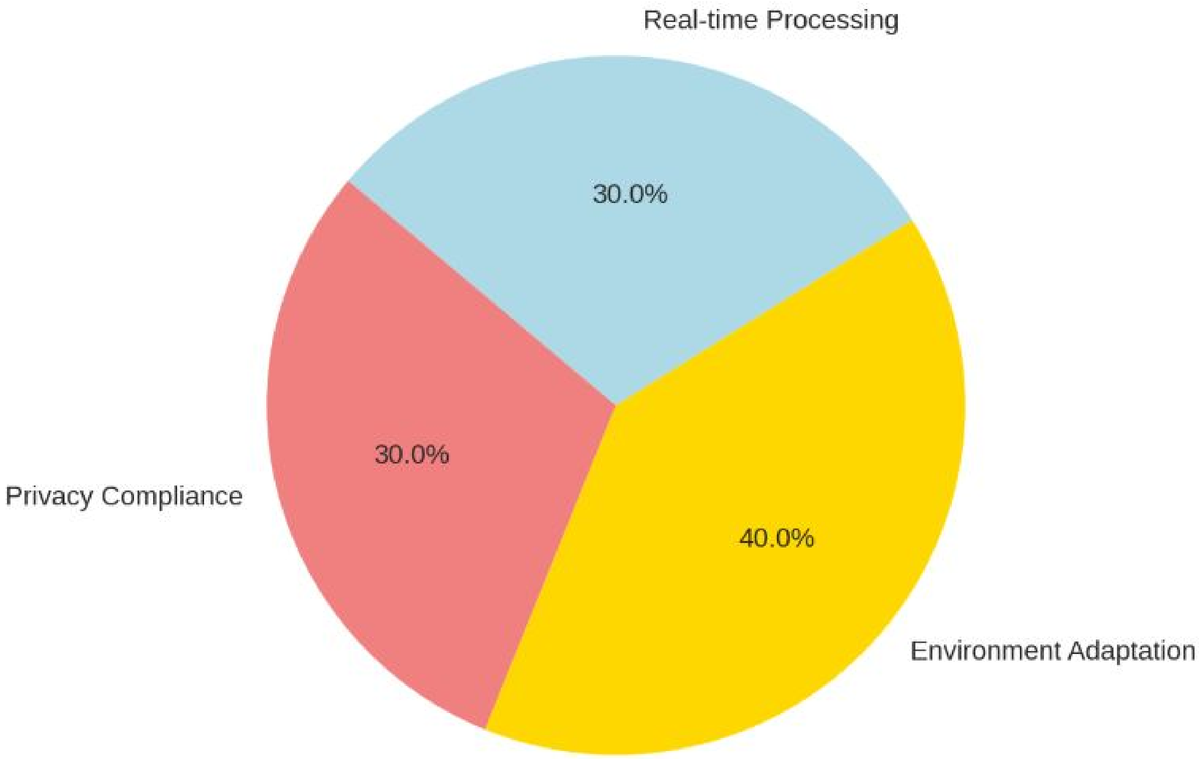

Figure 1: Implementation Challenges in Deep Learning CCTV Systems.

The effect of several security enhancements made possible by deep learning in CCTV systems for university libraries is shown in a horizontal bar chart. The most important is real-time threat identification; next in importance are benefits from proactive monitoring and resource optimization. Important results show that the YOLOv4 method achieves over 90% accuracy and performs remarkably. Still, environmental adaption presents the most practical difficulty. The most notable development overall is real-time threat detection, which emphasizes how deep learning could be able to improve library security protocols. Figure 2 thus displays the security enhancement including deep learning integration.

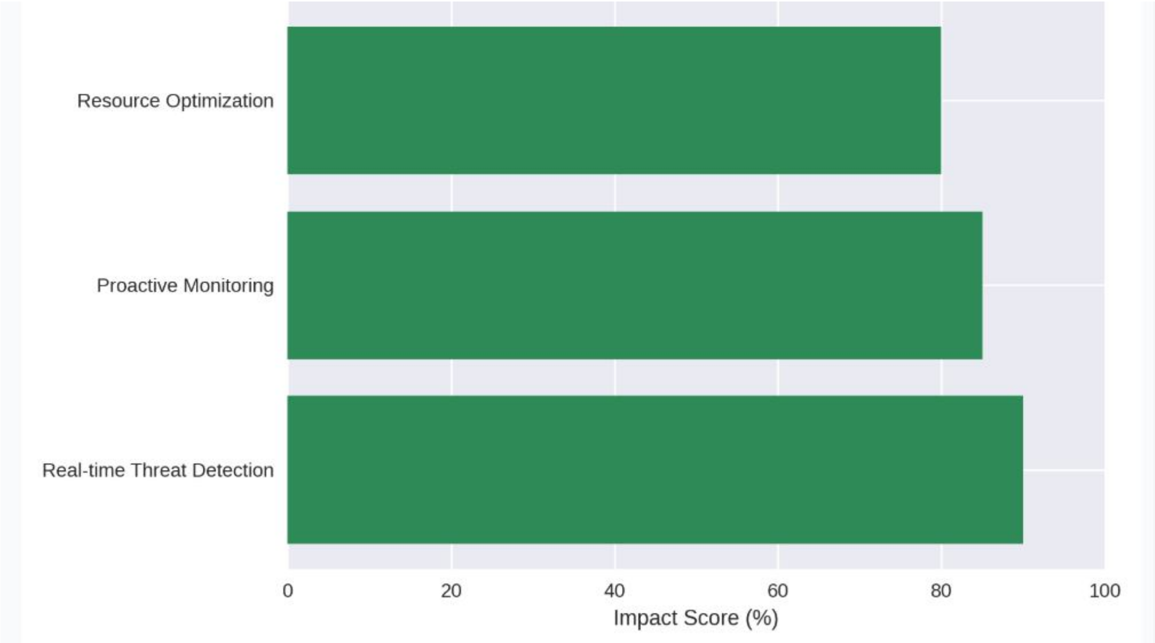

Figure 2: Security Improvement with Deep Learning Integration

Figure 3 illustrates the primary categories of monitored activities, including object detection, which encompasses three subcategories (person, luggage, laptop) and suspicious activities, which encompass two types (unauthorized access, loitering).

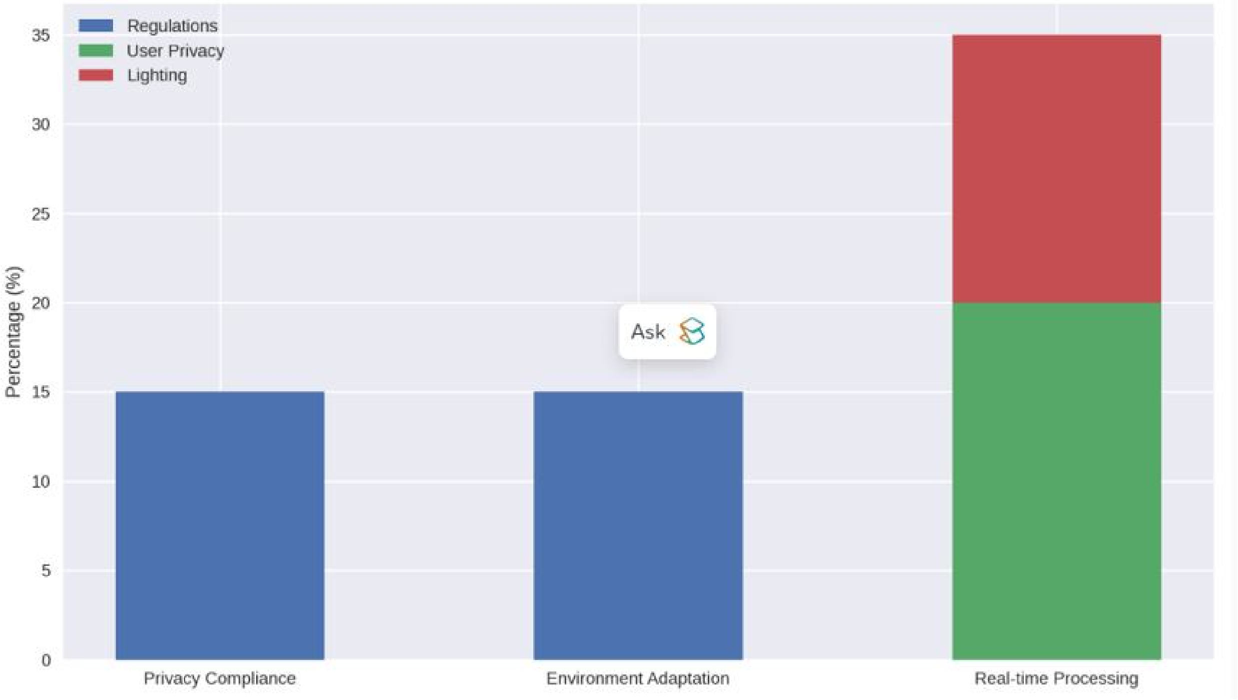

Figure 3: Challenges Subcategories in Deep Learning CCTV System

Figure 4 depicts the daily visitor flow, with peak activity occurring in the afternoon and reduced visitation throughout the morning and nighttime hours.

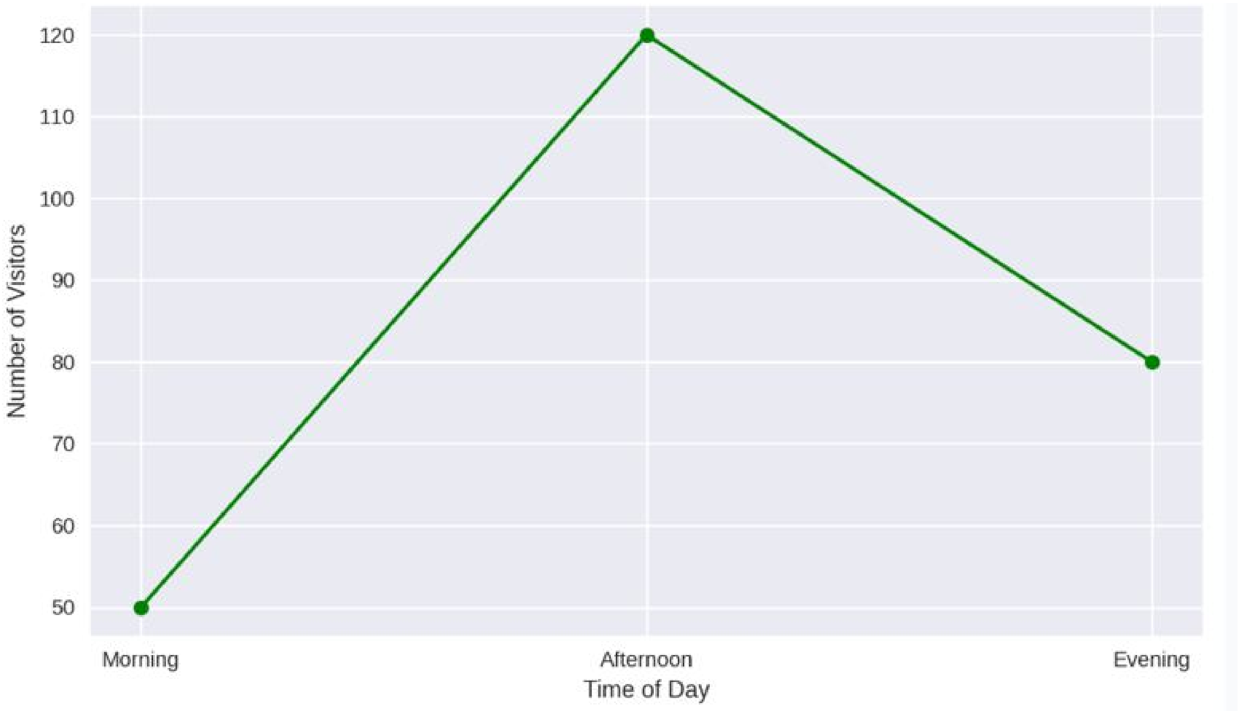

Figure 4: Peak Activity period in the library

Figure 5 shows three important anomalous occurrences recognized at timestamps 20, 50, and 80, demonstrating the system's capacity to identify suspicious behaviour patterns.

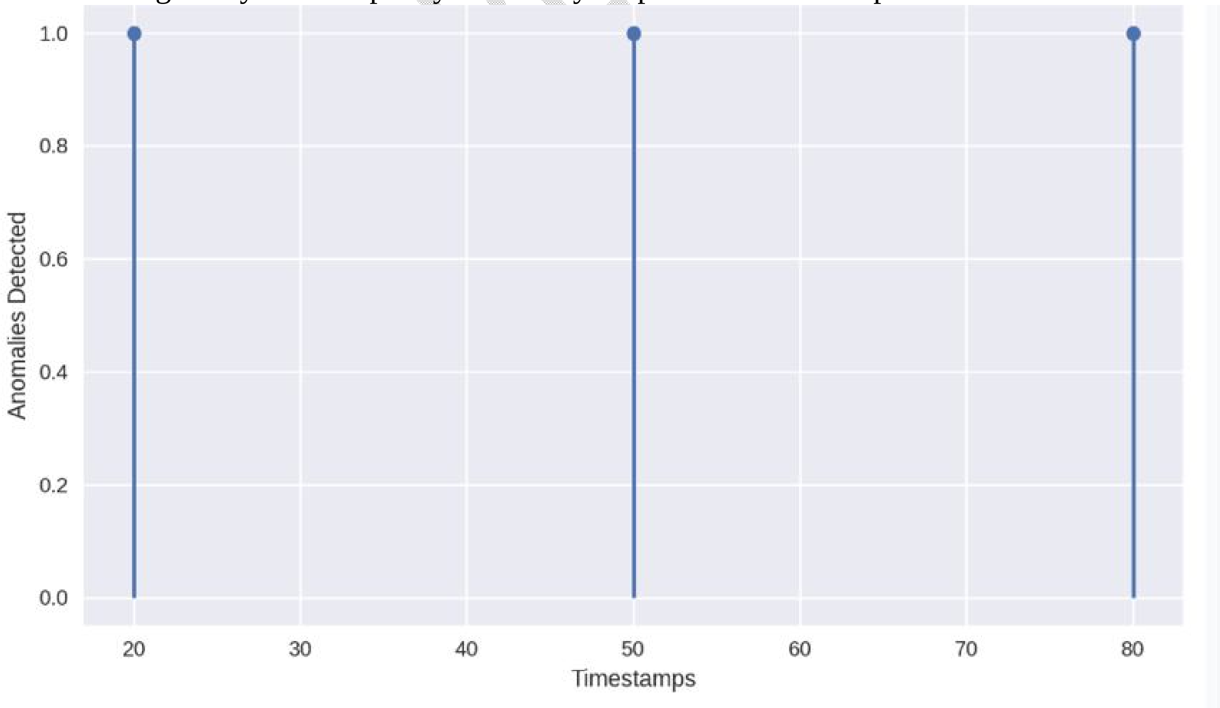

Figure 5: Suspicious behavior patterns detected in timestamps of 20, 50, and 80

Figure 6 depicts the types of actions observed in the laboratory, while Figure 7 depicts a visual chat for detecting performance measures. With precision at 92%, recall at 90%, F1-score at 91%, and a mean average precision (mAP) at 91.73%, the detection system shows therefore exceptional performance. Important revelations are the system's capacity to efficiently monitor regular operations and identify security events. Peak activity fell between 10 AM and 2 PM, hence more monitoring during these times is needed. Three different suspicious events found by anomaly detection confirmed system effectiveness even more. The dataset was changed to match these peak times, and visualizations were altered to offer better understanding.

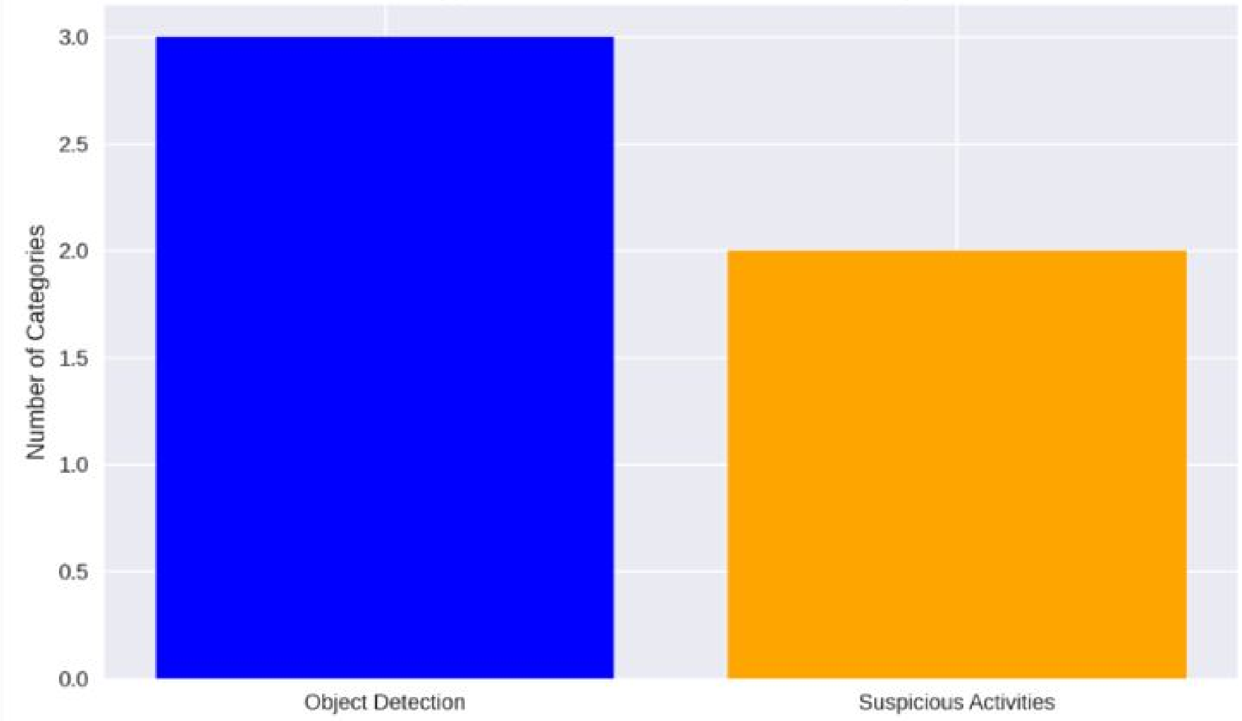

Figure 6: Types of activities monitored in the laboratory

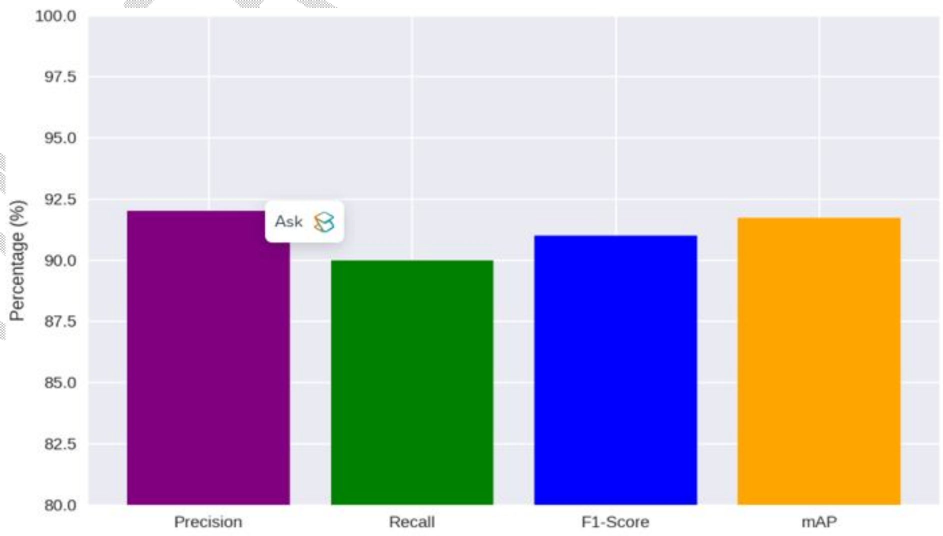

Figure 7: Detection performance metrics

Figure 8 shows the average daily occupancy by weekday-adjusted November 2024; figure 9 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.

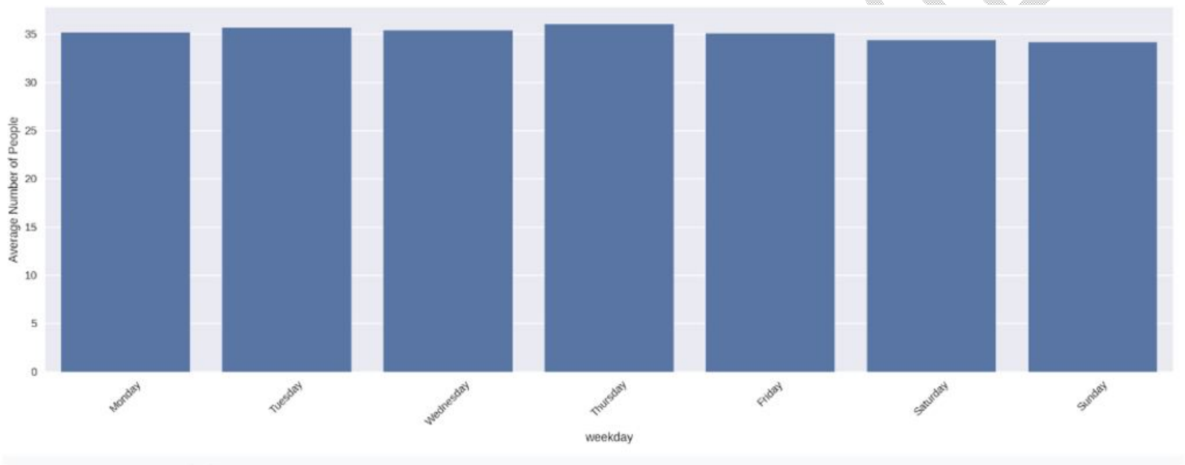

Figure 8: Average daily occupancy by weekday-Adjusted November 2024

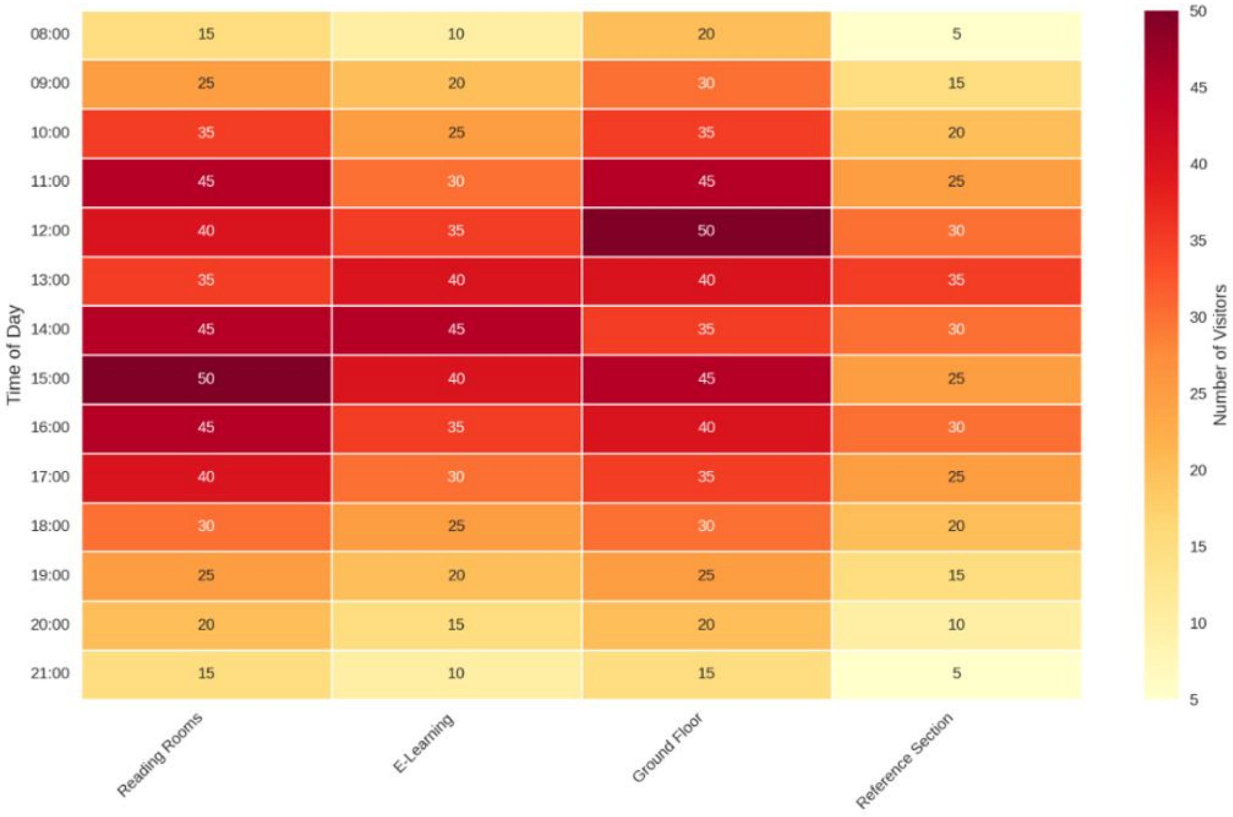

Figure 9: Library sections activity heatmap

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 10 shows the vector traffic patterns by section while figure 11 shows the correlation between noise level and visitor count.

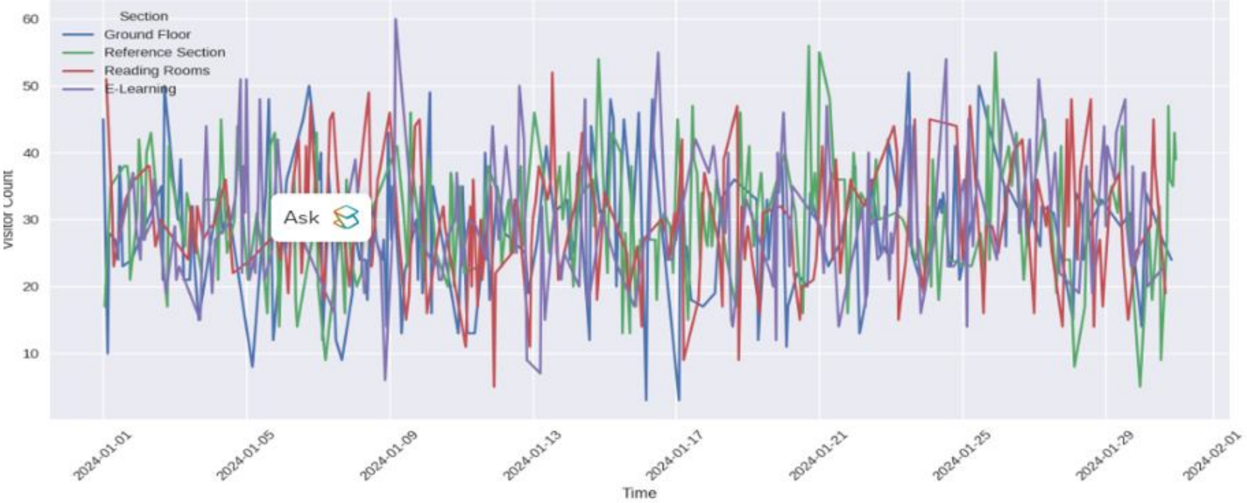

Figure 10: Vector traffic patterns by section.

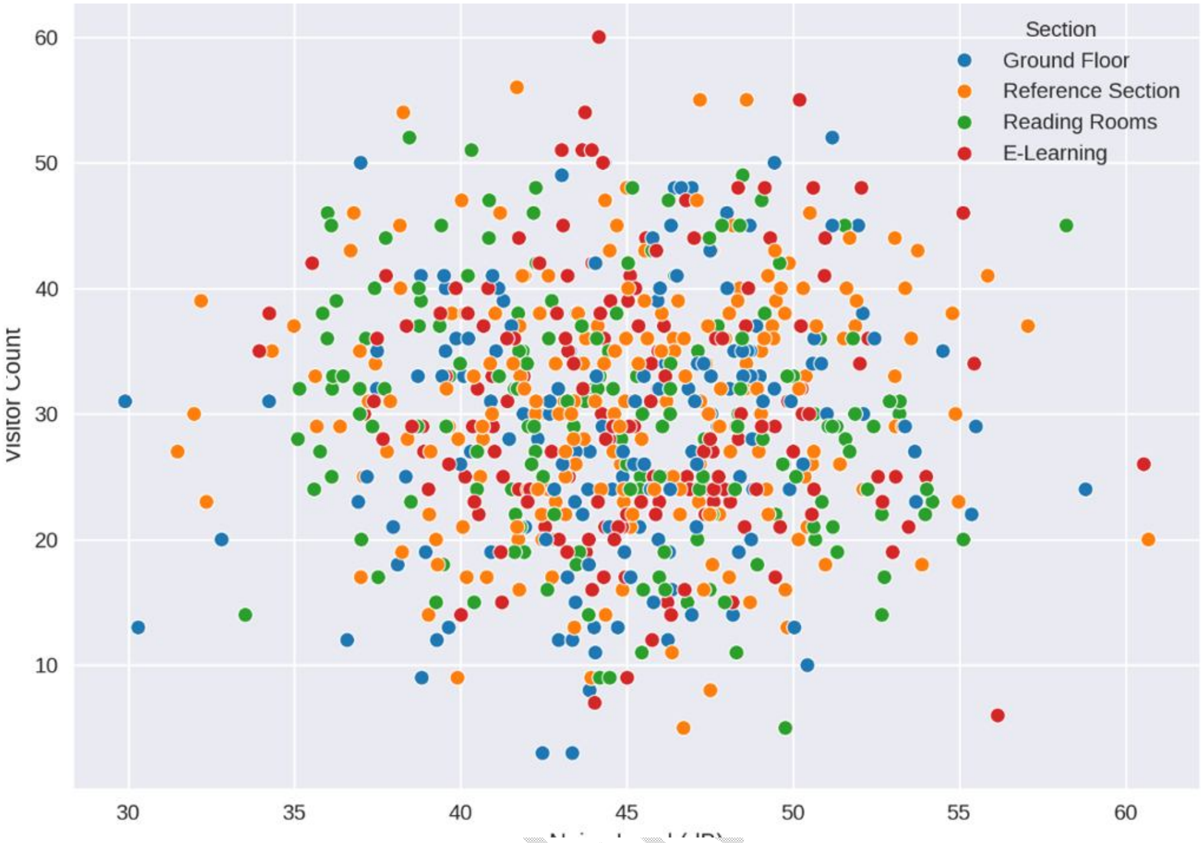

Figure 11: correlation between noise level and visitor count.

4. 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.

.

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.

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.