Review On Deep Learning-Based Approach to Intelligent Video Surveillance Systems

ETTEBONG STEPHEN JAMES[1], KINGSLEY M. UDOFIA[2], AKANINYENE B. OBOT[3]
[1,2,3]*Department of Electrical and Electronics Engineering, University of Uyo. Uyo. Akwa Ibom State*

Abstract- ***In recent years, the demand for more robust and Intelligent Video Surveillance Systems (IVSS) has grown due to the increasing need for public safety and security in both urban and remote environments. This study investigates the application of various techniques like Deep Learning (DL) and Machine Learning (ML) techniques in enhancing video surveillance systems considering anomaly detection, human behaviour recognition, violence detection and weapon identification. A comprehensive literature review was conducted for the assessment of performance, advantages and limitations of existing intelligent surveillance systems which highlighted that the capabilities of advanced models such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs) and hybrid deep learning architectures in automatically analysing video footage, learning complex patterns and detecting threats in dynamic environments significantly outperform traditional methods in terms of accuracy, adaptability and operational efficiency. The study concludes that the integration of DL and ML into surveillance systems presents a promising direction for modern security infrastructure, which not only reduces the burden on human operators but also enhance real-time threat detection and response, making them indispensable tools for future surveillance applications.***

Indexed Terms- ***Video Surveillance; Deep Learning (DL); Machine Learning (ML); CNN; RNN; Human Behaviour***

I. INTRODUCTION

In order to improve public safety and deter crime, surveillance cameras are now routinely deployed in a variety of locations, including residences, banks, businesses, and airports. Alternatively, by examining these films and trying to identify the offender, it is possible to determine the time and location of the crime and, more precisely, the culprit. In the meanwhile, someone is required to monitor the cameras from behind the scenes and identify any unusual activity. However, because anomalies are so uncommon, they cause fatigue, and when they do occur, they might often be unaware of it. He loses the anomalous, to put it another way (Sultani et al., 2018). Additionally, the anomaly-detection procedure is founded on human intuition, which is acquired over time. However, further issues with non-automated crime prediction and detection systems that rely on monitoring surveillance recordings include the skill level of the individual for recognising signals of crime happening and the expense of hiring him.

Machine learning and deep learning techniques must be used to extract certain visual cues in order to automate anomaly detection (James et al., 2020; Sochima et al., 2025; Mei and Zhang, 2017). Specific characteristics for various anomaly classes, such as vandalism (Ghazal et al., 2007), violence detection (Febin et al., 2020), and robbery (Lao et al., 2009), might be helpful for improving the performance of these algorithms (Yan et al., 2016). Reducing the damage by anticipating the crime's scene and timing. However, security personnel are also on time. For example, in an experiment produced in Santa Cruz, California, cops get daily crime estimates every morning. This forecasting covers certain areas. them to patrol Thirteen wrongdoers have been stopped in the designated areas over the first six months of the program, according to a Santa Cruz spokeswoman (Ferguson, 2012). Hence, this paper presents a comprehensive review on the application of various techniques, both manual and automated surveillance of an environment. The survey is aimed to determine the performance, limitations, strengths and existing

gaps in the area of environmental surveillance. These gaps will serve as a guide to propose the most effective technique which can be adopted for future surveillance technologies for a more effective, reliable and proactive security and threat mitigation system.

II. LITERATURE REVIEW

Gervais (2023) presented a smart surveillance system for the detection of anomalies at home through the use of machine learning and computer vision technique. This study employs a number of machine learning computational methods for anomaly detection, such as Autoencoders (AE) neural networks, which are used to reconstruct input data but frequently fall short, and Support Vector Machines (SVMs), which find the hyperplane that best separates normal data from anomalies. Using the Home the Assistant platform, a self-configured laboratory was constructed to gather testing event data for the system's experimental implementation. Also used was data from online device simulations based on the CASAS HH114 dataset. The result of implementing the study showed that it achieved a Root Mean Square Error (RMSE) value of 0.0023. While, Pooja and Rajkumar (2024) used Recurrent Neural Network (RNN) technique for intelligent real-time surveillance of a Closed-Circuit Television (CCTV) camera system. A real-time video record containing both regular and unusual occurrences is used in the proposed strategy. There are twenty-one instructional videos and sixteen testing videos. The RNN technique is used to identify events with a threshold value of 0.006 after the capture data is transformed to frames using a 3D Convolution Network and a Spatiotemporal AE. The accuracy result was 95% when the obtained data was analysed using the RNN algorithm.

Kamble et al., (2022) presents a study on the recognition of anomaly in an environment through smart surveillance system based on deep neural network framework. The neural network architecture proposed in this paper consists of three modules: an object identification module, an object discriminator and tracking module, and an abnormal activity detection module based on recurrent neural networks. Users may apply online for a variety of security services, such as motion, fall, and anomaly detection, using the system. These services may be used to monitor a variety of locations, such as residences, streets, offices, schools, retail establishments, and other interior spaces. The behaviours considered in this study are represented by three different types of video data from different sources: motion detection, fall detection, and anomaly detection. The person subject in the video is identified using the object detection framework You Only Look Once version 4 (YOLOv4). The Visual Geometry Group (VGG16) approach is used to extract features in the first stage, and the RNN model is used to analyse the dynamics of anomalous behaviour. The UCF Crime dataset, a brand-new, extensive dataset with 128 hours of uncut real-world surveillance footage recorded on CCTV cameras, is used to train and test the suggested anomaly detection method. It contains 13 realistic anomalies, such as arrests, shoplifting, arson, burglaries, assaults, explosions, vandalism, fights, traffic accidents, robberies, abuse, shootings, and thefts. Following model training and analysis, it was shown that the proposed system can identify both normal and abnormal occurrences with an accuracy of 80.43%, 80% precision, and 80% recall.

Vijeikis et al., (2022) adopted machine learning and computer vision techniques for the detection of violence from a video surveillance camera system. The proposed technique extracts spatial characteristics using MobileNet V2 as the encoder in a U-Net-like network. An LSTM model is then used to extract and classify temporal characteristics from the gathered information. The datasets used to build the model include the RWF-2000 dataset for violence detection, the movie fight dataset, and the hockey dataset combined. The RWF-2000 dataset contains 1600 videos, including 1000 recordings of hockey matches and 200 films of combat sequences. With a total of 4,074,435 parameters, the dataset is fast and computationally light. The results of the system implementation reported that the proposed model achieved an accuracy of 82.3% is the RWF-2000 dataset, then 97.1% the hockey fight dataset and finally 99.5% accuracy in the movie fight scene dataset. Then, with the use of the same datasets, Huillcen-Baca et al., (2024) researched on the recognition of human violence in real-time over a surveillance camera. However, the Global Temporal

Extractor (GTE), Short Temporal Extractor (STE), and Spatial Motion Extractor (SME) modules were the three modular patterns utilised in this work. SME extracts regions of interest from frames, STE extracts temporal characteristics, and GTE extracts long-lasting temporal features. GTE additionally refines the proposed 2D CNN model used in the study. Based on the model's performance, the proposed approach identified violent scenarios with 90.71% accuracy and identified non-violent settings with 89.29% accuracy.

Chunchwar et al., (2024) applied YOLOv8 and alert mechanism in video tracking device for detection of weapons in real-time through surveillance video analysis in different situations. The system collected a set of photographs and related annotations specifically created for weapon detection. Three weapon kinds are considered in the training data: handguns, knives, and artillery. This solution utilises Streamlit for the user interface and an email alert mechanism for timely alerting after carefully selecting datasets and training examples on bespoke datasets. Notably, security staff were notified instantly when firearms were identified in the live camera stream thanks to an expansion of the email alert mechanism for system users. The model's Mean Average Precision (MAP), as determined by the study's implementation, was 0.78. However, in another study by Mukto et al., (2024), YOLOv5 and MobileNetv2 models were applied for monitoring of crime in real-time for an environment. This study presents a suggested criminal Monitoring System (CMS) that uses a variety of deep learning and image processing techniques in conjunction with the processes and capabilities of CCTV cameras to detect criminal situations. The CMS operation first looks for weapons, then it looks for violent incidents that have occurred there, and lastly it uses face recognition to identify the individuals involved. While the MobileNetv2 model was used to identify violence in the scene, the YOLOv5 model was utilised to detect weapons. The face detection and identification model were then created using the Local Binary Pattern Histogram (LBPH). The results of the study reported that the weapon detection model detected weapons with about 80% accuracy, then the violence detection model was also 95% accurate. Finally, the face recognition model had a 97% accuracy rate.

Raksha et al., (2025) applied machine learning technique for real-time surveillance system for the detection of anomaly. The system developed a efficacious anomaly detection algorithm combines three interrelated steps: the application of Convolutional Neural Networks (CNNs), the addition of mask recurrent convolutional neural networks, and the improvement of spatial awareness via semantic segmentation. The UCF-Crime dataset, which includes real-world surveillance footage of a variety of criminal actions such theft, robbery, burglary, and vandalism in congested urban settings, was utilised to carry out the study. With a 98.5% accuracy rate, the method's output tackles issues like outliers and video noise. With GPU and FPGA acceleration, locality in anomaly detection is investigated, demonstrating scalability and resilience for bigger datasets.

Rashvand et al., (2025) presents a pose-based anomaly detection system for improving the security of a retail store. The study addressed issues including data scarcity, privacy problems, and model biases by introducing PoseLift, a privacy-preserving dataset created especially for shoplifting detection. Bounding box, person ID, and human posture annotations are among the anonymised data made available by the PoseLift dataset. The study extracted the annotations using a similar method. To locate and identify people in each video frame, the YOLOv8 object detection model was used. Bounding boxes are created by YOLOv8 around detected individuals, showing where they are in the scene. The study have used the Byte Track technique for person ID annotations, which guarantees that people can be tracked even in congested scenarios and enables the system to give IDs to each individual. The results of the system implementation presented that the system achieved an Area Under the Receiver Operating Characteristic Curve (AUC-ROC) of 67.46%, AUC- Precision-Recall Curve (AUC-PRC) of 84.06% and Equal Error Rate (EER) of 0.39.

Dhumal et al., (2024) presents a deep learning-based anomaly detection system through environmental surveillance of a crowded place. The proposed approach uses a 63-layer deep CNN model named "L4-BranchedActionNet" to watch public and private areas including banks, retail malls, train stations, and airports in order to spot odd activity linked to crimes

like theft, damage, and other suspicious acts. First, the CIFAR-100 dataset is used to pre-train the framework for object detection using SoftMax. Entropy coding and an Ant Colony Optimisation System (ACOS) are used to further optimise the features that were taken from the dataset using the CNN model. Multiple classifiers, including SVM and KNN models, are then used to classify the optimised features. The implementation results of the study reported that SVM proved to achieve the highest accuracy score at 99.24%. Validation on the Weizmann action dataset achieved an accuracy score of 97.96%.

Jeon et al., (2024) presents a system called PASS-CCTV which is an anomaly surveillance system through CCTV footage analysis considering adverse environmental conditions. Strong human feature extraction and sophisticated object filtering techniques are used in the suggested human tracking method. Furthermore, the framework performs well in identifying human-related abnormalities such arson, loitering, desertion, and intrusion. A prompt-based detection method that allows for active user engagement in recognising anomalous situations was then further added in the study. The suggested method has shown notable performance improvements in extensive tests utilising the Korea Internet & Security Agency (KISA) CCTV datasets, especially in difficult weather situations. The system is then further verified using the ABODA and FireNet datasets. According to the system's performance, the method used produced 100% recall, 97.12% specificity, and 99.08% accuracy.

Afreen et al., (2023) researched on the development of a Smart Surveillance (SS) system for monitoring of High Security Area (HSA) through Internet of Things (IoT) technology. The Gravity Microwave Sensor (GMS) used in this system is very effective since it can pass through non-metallic obstacles. Detecting suspected things behind walls is made much more effective by combining GMS with Arduino UNO. The GMS is an IoT-based solution as it can be connected to mobile communications. The SS-HSA system analyses system performance using machine learning algorithms such gradient boosting, naïve bayes, random forest, decision tree SVM, and KNN classifiers running at a GMS frequency. Since the Arduino UNO is not impacted by the surrounding environment, the dataset was collected in several locations during the experiment. 1600 executions were carried out when it was utilised in a residence. There were 1150 trials when it was utilised in a retail mall and 1200 executions when it was used in a garage. As per the study's implementation results, the approaches implemented here achieved accuracies of 94% for gradient boosting, 95% for random forest, 93% for naïve bayes, 97% for decision trees, KNN, and 96% for SVM.

III. THE CONCEPT OF VIDEO SURVEILLANCE SYSTEMS

Promising options for automated surveillance systems that view and monitor surroundings include digital cameras, as illustrated in Figure 1, surveillance monitoring, and control software frameworks. Individual behaviour, crowd behaviour, interpersonal interactions, motion detection, crowds, and their surrounding settings are all used to assess the scenarios and circumstances that were witnessed. Multiple tasks, such as detection, interpretation, comprehension, recording, and the creation of warnings as a result of the system analysis, are made possible by the design and implementation of these autonomous systems (Elharrouss et al., 2021).

Figure 1: Samples of Digital Cameras for Surveillance Systems (Source: Pass Security, 2025)

Significant advancements have been made in many sectors of the world during the last 20 years, which has made life more complex in many ways, including

human safety and security. As a result, monitoring and these elements are now required. In this sense, cameras and surveillance software frameworks placed in both public and private areas are suitable ways to guarantee security and comfort. Usually, humans are required to watch these cameras all day, every day, which is a time-consuming and costly duty. Thus, it is very desired to have an automated system that uses software frameworks and surveillance cameras to monitor and regulate events in real time under various conditions (Elharrouss et al., 2021).

A workable option is the development and deployment of software frameworks for video surveillance. An automated system's main purpose is to help security guards with a variety of responsibilities. The tasks that must be completed may pertain to several application areas, including homeland security, criminal prevention, motion detection, traffic management, and accident prediction. To monitor both interior and outdoor sceneries, such as parking lots, roads, shops, shopping centres, airports, train stations, and workplaces, more application areas can be added. In order to maintain public safety and control and avoid anomalous occurrences, especially in situational-awareness applications, software security and monitoring frameworks are being gradually implemented. As a result, a framework for security monitoring and control software that can automatically keep an eye on and manage human lives has to be created (Elharrouss et al., 2021; Porikli et al., 2013).

Deep Learning (DL) Application for Video Surveillance

In the moment, deep learning-based face recognition methods are quite successful in recognising individuals by their facial characteristics. The DL model is characterised by weights arranged in an array of values to get the necessary features (Ebere et al., 2025; Chidi et al., 2024). Its kernels are capable of detecting a borderline function or the contour of an image. In order to ascertain the control of the picture to be recognised, each DL model allocates a space (Hussain and Salim, 2019). Deep and broad neural networks have shown difficulty when used with datasets for general-purpose network applications. DLs, on the other hand, have shown to be more successful at detecting and recognising objects. Additionally, DLs have revolutionised the domains of audio processing and computer vision (Li et al., 2016).For example, smartphone devices are built with DL architecture-based AI-based object-recognition capabilities, allowing end users to access apps like digital fingerprints, voice commands, and object detection in photos (Pablo, 2020). Due to its ability to resolve some of the most challenging computer-vision issues, DLs have exceptional capabilities. Furthermore, DLs provide the special capacity to extract and categorise object properties by encoding spatial connections in datasets (Kekong et al., 2019).

DL is also used for automatic weapon identification in real-time video surveillance. This method tracks occurrences in real time and uses surveillance devices to monitor and categorise weaponry. Three processing units are specifically utilised. The object-detection module uses a DL in the first place, followed by a module for weapon classification in the second and a module for monitoring and alert functions in the third. The installed surveillance system monitors a designated area of interest and carries out fundamental monitoring and control tasks using a closed-circuit video system. The accuracy of two algorithms the shape and object-detection algorithms was assessed in terms of processing time and detection accuracy. According to the findings, the ALEXNET dataset's weapon and item categories, names, and forms may be detected with nearly optimum accuracy (Bhagyalakshmi et al., 2019).

A proposed approach for making the training of densely layered networks easier is deep residual learning for image identification. In this method, the input layers are considered as reference points, and the network layers are formulated as a learning residual function. This method offers thorough evidence that residual networks, using straightforward optimisation strategies, may attain excellent accuracy based on a significantly enhanced network depth. The residual networks were evaluated using the ImageNet dataset with a depth of 158 layers, which is 8 times deeper than the VGG nets with lower complexity. According to experimental

results based on a real-life application, encouraging outcomes were obtained (He et al., 2016). A system based on Convolutional Neural Networks (CNNs) that uses deep learning for real-time object recognition and tracking provides an additional option for this scenario. The foundation of this idea is a spatial-temporal process. Target interaction and occlusion biases are addressed by this method. A superior method for real-time object recognition, tracking, and counting in various datasets is provided by a software system that uses TensorFlow and YOLO's algorithm (Kusuma and Ashwini, 2023).

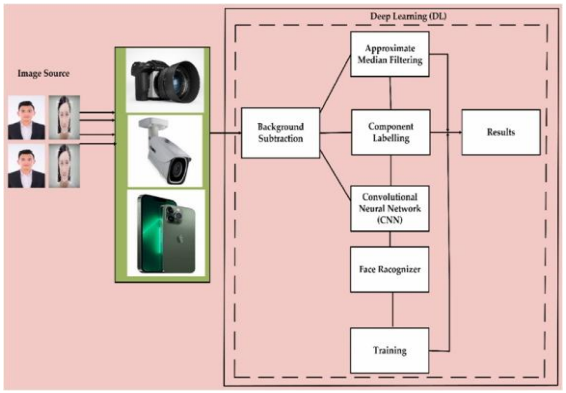

Figure 2: Block diagram of a DL technique (Abba et al., 2024)

A face recognition method based on deep learning is shown in Figure 2. A digital camera, a security camera, or an iPhone are the three devices that the algorithm uses to capture an item, as seen in Figure 2. Frames in the video sequence are identified by doing background subtraction. Component labelling for tracking, the face recogniser for identification, and the median filter for detection are the methods that the deepest learning algorithm layer categorises as appropriate for usage. Additionally, the algorithm keeps an eye on and manages the system's training procedure. The resultant object is the target object.

Figure 3 shows the DL framework's system logging flowchart. To protect the system from unauthorised users, the algorithm starts with user authentication. Once the authentication procedure is complete, a legitimate user can choose between activities like object monitoring, object detection, and CNN or DL recognition. The system's three primary picture sources were an iPhone, a digital camera, and a security camera, as seen in Figure 2. In order to detect an item, the algorithm subtracts the relevant background. The tracking method is the same, except the monitored items are placed inside a rectangular box. In order to recognise the items, training is used to identify their faces. Depending on the number of objects in the picture, rectangular boxes are used to identify each thing by its name and face.

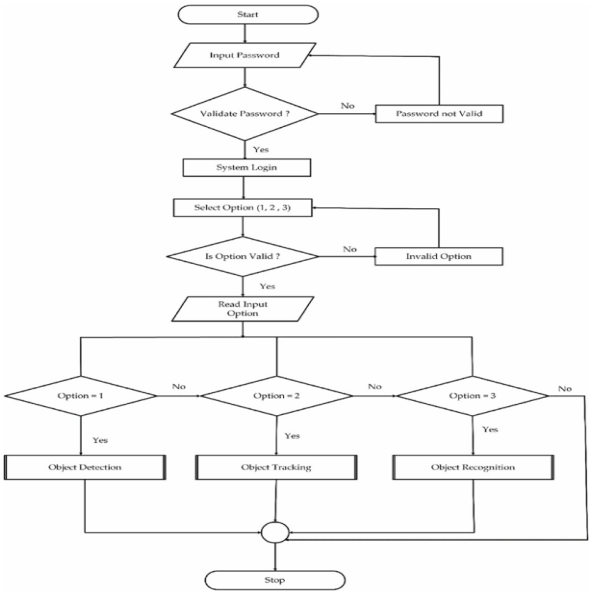

Figure 3: Flowchart of DL Approach for Video Surveillance and Analysis(Abba et al., 2024)

IV. OBSERVATIONS FROM THE STUDY

From the reviews conducted, it has been observed that the surveillance of environment is a very sensitive and time-reliant activity which requires pro-active efforts for the mitigation of the consequences. Some of the major observations from the review are:
Reliance on Manual Monitoring and Human Supervision

Traditional surveillance systems heavily depend on human operators to monitor live video feeds and detect a suspicious activity which introduces a significant margin for human error, particularly due to fatigue, distraction or information overload. According to multiple studies, it was identified that security personnel can only maintain consistent attention for a limited time often making them to miss events after prolonged monitoring periods.

High False Alarm Rates

One of the most common issues identified from the literature is the high occurrence of false positives in traditional surveillance systems which leads the system to typically rely on motion detection, line crossing or threshold-based triggers. The frequent occurrence of such false alarms can overwhelm monitoring staff and lead to desensitization, whereby genuinely suspicious activities are overlooked or not acted upon promptly. This undermines the overall effectiveness of the surveillance infrastructure.

Inability to Interpret Complex Human Behaviour

Traditional monitoring techniques lack the capacity to understand and analyze complex behavioural patterns as they may be able to detect motion or breaches of predefined zones, they are incapable of distinguishing between benign and malicious intent. Hence, this limitation makes it inadequate for tasks such as violence detection, crowd behaviour analysis or recognizing small indicators of suspicious activity capabilities that DL models can achieve with much higher precision.

Scalability Issues in Large-Scale Deployments

Traditional monitoring systems become increasingly inefficient as the number of surveillance points grows. This means that with the addition of each extra camera, it increases the burden on human operators, data storage systems and network bandwidth operating the system. Therefore, without the adoption of automation or intelligent analysis, scaling such systems to city-wide deployments results in operational inefficiencies and higher costs. However, literature on smart surveillance systems emphasizes on the need for intelligent video analytics for the management large volumes of data effectively which traditional systems are ill-equipped to handle.

CONCLUSION

This study explored the development and effectiveness of Intelligent Video Surveillance Systems (IVSS) using DL and ML algorithms for real-time monitoring, anomaly detection, human behaviour recognition, violence detection and weapon/threat identification in an environment. Through an extensive literature review, the research analysed the limitations of traditional surveillance systems and highlighted the growing need for intelligent systems capable of processing vast visual data streams accurately. From the review, it was discovered that traditional monitoring approaches which is reliant on manual observation or simple rule-based algorithms were found to be inefficient, prone to errors, and incapable of adapting to complex or dynamic environments. These limitations have driven the evolution toward intelligent systems, where deep learning models such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and hybrid approaches to play important roles in enhancing detection accuracy and system reliability.

The study reviewed recent applications of ML and DL techniques in intelligent surveillance, focusing on object tracking, behavioural analysis, threat detection and real-time alert generation. Observations from the literature show that integrating deep learning improves scalability, adaptability, and precision, reducing false alarms and enabling proactive threat identification. Furthermore, deep learning techniques empower surveillance systems with the ability to learn patterns over time, making them suitable for high-security applications such as public safety, transportation hubs, military installations, and smart cities.

The findings of this study confirm that intelligent surveillance systems powered by deep learning and machine learning represent a transformative advancement in security technology. Unlike the traditional method, DL algorithms enhance the system's ability to recognize complex behavioural patterns, detect anomalies in real-time and respond proactively to potential threats which results in more accurate, efficient and autonomous surveillance.

REFERENCES

[1] Abba, S., Bizi, A. M., Lee, J. A., Bakouri, S., & Crespo, M. L. (2024). Real-time object detection, tracking, and monitoring framework for security surveillance systems. Heliyon, 10, e34922. https://doi.org/10.1016/j.heliyon.2024.e34922

[2] Afreen H., Kashif M., Shaheen Q., Alfaifi Y.H., & Ayaz M., (2023) IoT-Based Smart

Surveillance System for High-Security Areas. Appl. Sci. 2023, 13, 8936. https://doi.org/10.3390/app13158936

[3] Bhagyalakshmi, P., Indhumathi, P., Lakshmi, R., &Bhavadharini, D. (2019). Real-time video surveillance for automated weapon detection. International Journal of Trend in Scientific Research and Development (IJTSRD), 465–470.

[4] CHIDI, E. U., UDANOR, C. N., & ANOLIEFO, E. (2024). Exploring the Depths of Visual Understanding: A Comprehensive Review on Real-Time Object of Interest Detection Techniques. Preprints. https://doi.org/10.20944/preprints202402.0583.v1

[5] Chunchwar P., Shelare U., Nagpure A., Patil R., Dhole D., & Shete R.M., (2024) Real Time Weapon Detection using YOLOv8 and Alert Mechanism. International Journal for Research in Applied Science & Engineering Technology (IJRASET) https://doi.org/10.22214/ijraset.2024.60177

[6] Dhumal R., Chandgude P., Jamdade S., Pise M., & Kadam P.N., (2024) Deep Learning-Driven Surveillance System for Anomaly Detection in Crowded Environments. International Research Journal of Modernization in Engineering, Technology and Science https://www.doi.org/10.56726/IRJMETS63195

[7] Ebere Uzoka Chidi, E Anoliefo, C Udanor, AT Chijindu, LO Nwobodo (2025)" A Blind navigation guide model for obstacle avoidance using distance vision estimation based YOLO-V8n; Journal of the Nigerian Society of Physical Sciences, 2292-229; https://doi.org/10.46481/jnsps.2025.2292

[8] Elharrouss, O., Almaadeed, N., & Al-Maadeed, S. (2021). A review of video surveillance systems. Journal of Visual Communication and Image Representation, 77, 103116. https://doi.org/10.1016/j.jvcir.2021.103116

[9] Febin, I. P., Jayasree, K., & Joy, P. T. (2020). Violence detection in videos for an intelligent surveillance system using MoBSIFT and movement filtering algorithm. Pattern Analysis and Applications, 23(2), 611–623.

[10] Ferguson, A. G. (2012). Predictive policing and reasonable suspicion. Emory Law Journal, 62(2), 259.

[11] Gervais N., (2023) Smart Surveillance System With Anomaly Detection At Home. Faculty Of Computing And Information Sciences Masters Of Sciences In Information Technology. University of Lay Adventists of Kigali. Reg.No: M02141/2022

[12] Ghazal, M., Vazquez, C., & Amer, A. (2007). Real-time automatic detection of vandalism behavior in video sequences. In Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (pp. 1056–1060).

[13] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 770–778). https://doi.org/10.1109/CVPR.2016.90

[14] Huillcen-Baca H.A., Palomino-Valdivia F.d.L., & Gutierrez-Caceres J.C., (2024) Efficient Human Violence Recognition for Surveillance in Real Time. Sensors 2024, 24, 668. https://doi.org/10.3390/s24020668

[15] Hussain, S. A., & Salim, A. A. A. (2020). A real-time face emotion classification and recognition using a deep learning model. Journal of Physics: Conference Series, 1432, 012087. https://doi.org/10.1088/1742-6596/1432/1/012087

[16] James P. Suarez, J., & Naval, P. C. Jr. (2020). A survey on deep learning techniques for video anomaly detection. arXiv. https://arxiv.org/abs/2009.14146

[17] Jeon H., Kim H., Kim D., & Kim J., (2024) PASS-CCTV: Proactive Anomaly surveillance system for CCTV footage analysis in adverse environmental conditions. Expert Systems With Applications 254 (2024) 124391 https://doi.org/10.1016/j.eswa.2024.124391

[18] Kamble K., Jadhav P., Shanware A., &Chitte P., (2022) Smart Surveillance System for Anomaly Recognition. ITM Web of Conference 44, 02003 (2022) ICACC-2022 https://doi.org/10.1051/itmconf/20224402003

[19] Kekong P.E, Ajah I.A., Ebere U.C. (2019). Real-time drowsy driver monitoring and detection system using deep learning based behavioural approach. International Journal of Computer Sciences and Engineering 9 (1), 11-21

[20] Kusuma, T., & Ashwini, K. (2023). Real-time object detection and tracking design using deep learning with spatial-temporal mechanisms for video surveillance applications. In H. S. Saini, R. Sayal, A. Govardhan, & R. Buyya (Eds.), Innovations in Computer Science and Engineering. ICICSE 2022 (Lecture Notes in Networks and Systems, Vol. 565). Springer. https://doi.org/10.1007/978-981-19-7455-7_56

[21] Lao, W., Han, J., & De With, P. (2009). Automatic video-based human motion analyzer for consumer surveillance system. IEEE Transactions on Consumer Electronics, 55(2), 591–598.

[22] Li, Y. D., Hao, Z. B., & Lei, H. (2016). Survey of convolutional neural networks. Journal of Computer Applications, 36(9), 2508–2515.

[23] Mei, T., & Zhang, C. (2017). Deep learning for intelligent video analysis. In Proceedings of the 25th ACM International Conference on Multimedia (pp. 1955–1956).

[24] Mukto M., Hasan M., al-Mahmud M., Haque I., Ahmed A., Jabid T., Rashid M., Islam M.M., & Islam M., (2024) Design of a real-time crime monitoring system using deep learning techniques. Intelligent Systems with Applications 21 (2024) 200311 https://doi.org/10.1016/j.iswa.2023.200311

[25] Munemma, D., & Uma-Maheswari, K. V. (2024). Predicting robbery behavior potential in indoor security cameras using propounding first AI approach. Journal of Engineering Sciences, 15(06).

[26] Pablo, R. (2020). Deep Learning for Beginners: A Beginner's Guide to Getting up and Running with Deep Learning from Scratch Using Python. Packt Publishing Ltd.

[27] Pass Security. (n.d.). Benefits of commercial video surveillance systems. Retrieved from https://www.passsecurity.com/benefits-of-commercial-video-surveillance-systems/

[28] Pooja B.R., Rajkumar N., (2024) Real-Time Intelligent Video Surveillance System using Recurrent Neural Network. International Conference on Machine Learning and Data Engineering (ICMLDE 2023) Procedia Computer Science 235 (2024) 1522–1531. 10.1016/j.procs.2024.04.143

[29] Porikli, F., Brémond, F., Dockstader, S. L., Ferryman, J., Hoogs, A., Lovell, B. C., Pankanti, S., Rinner, B., Tu, P., & Venetianer, P. L. (2013). Video surveillance: past, present, and now the future DSP forum. IEEE Signal Processing Magazine, 30(3), 190–198.

[30] Raksha., Ramyashree., Ganiga R., Nayak S.V., & Kini M., (2025) Machine learning Based Real Time Surveillance System for Anomaly Detection. Alvas Institute of Engineering and Technology

[31] Rashvand N., Noghre G.A., Pazho A.D., Yao S., &Tabkhi H., (2025) Exploring Pose-Based Anomaly Detection for Retail Security: A Real-World Shoplifting Dataset and Benchmark. arXiv:2501.06591v1 [cs.CV] 11 Jan 2025

[32] Sochima V.E. Asogwa T.C., Lois O.N. Onuigbo C.M., Frank E.O., Ozor G.O., Ebere U.C. (2025)"; Comparing multi-control algorithms for complex nonlinear system: An embedded programmable logic control applications;

[33] DOI: http://doi.org/10.11591/ijpeds.v16.i1.pp212-224

[34] Suarez, J. J. P., & Naval, P. C. (2020). A survey on deep learning techniques for video anomaly detection. arXiv preprint arXiv:2009.14146.

[35] Sultani, W., Chen, C., & Shah, M. (2018). Real-world anomaly detection in surveillance videos. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 6479–6488).

[36] Vijeikis R., Raudonis V., &Dervinis G., (2022) Efficient Violence Detection in Surveillance. Sensors 2022, 22, 2216. https://doi.org/10.3390/s22062216

[37] Yan, H., Liu, X., & Hong, R. (2016). Image classification via fusing the latent deep CNN feature. In Proceedings of the International Conference on Internet Multimedia Computing and Service (pp. 110–113).