Intelligent Surveillance Framework for Crowd Detection and Alerting

Aman Kumar[1*], Hina Hashmi[2], Priya Singh[3] and Danish Raza Rizvi[4]

[1]*Dept. of CSE, VGI, Greater Noida, UP, India**
[2]*FOECS, Teerthanker Mahaveer University, Moradabad, UP, India*
[3]*Dept. of ACSE, GLBTIM, Greater Noida, UP, India*
[4]*Dept of Comp. Engg, FET, Jamia Millia Islamia, New Delhi, India*

Abstract

Declaring some regions to be unobstructed and crowd-free occasionally becomes a sudden necessity, and in instances like national constitutional applications and the pandemic COVID-19, monitoring a large area for the aforementioned rule and compliances becomes a major issue. Although drone surveillance technologies have been used by our security agencies recently, manual monitoring and procedures are still quite popular in nations like India. The automation that has been implemented, however, is neither intelligent nor self-alerting. There is always a path of action for security problems. Therefore, a significant democratic risk could result from a delay in the main course detection, such as when there is a crowd assembling in India or sensitive or secure regions for a while or for the reasons mentioned above. Although the governments are deploying CCTVs to monitor events, they are not yet sophisticated enough for the reasons already outlined. Even in this day and age of increasing automation, warnings must be triggered by themselves. The age we live in now is one of superior artificial intelligence systems, bringing new dynamics to society and ushering in a new era known as society 5.0. One further application of the AL & ML-based self-alerting system for secure and red alerting zones with low cost and less computationally intensive development has been put into practice and tested in this study. In order to comply with law enforcement's view of crowd gatherings, the tested system will serve as a framework for intelligent technologies that will eventually self-alert in secure and sensitive zones.

Keywords

Surveillance, Crowd Detection, Self-alerting System, Secure Zones, Artificial Intelligence (AI) and Machine Learning (ML)

1. Introduction

Smart and safe zones are defined as technically sophisticated and contemporary urban or rural locations equipped with voice activation, video surveillance, and data-collection sensors, among other electronic techniques [1]. Effective management of different laws, resources, assets, or other services can be achieved with the information retrieved from the recorded or gathered data. This type of strategy may prove beneficial in areas designated as prohibited for group gatherings by government regulations such as Section 144 [2] (which, in India, restricts gatherings involving more than four individuals).

Proceedings of SNSFAIT 2024: International Symposium on Securing Next-Generation Systems using Future Artificial Intelligence Technologies, Delhi, India, August 08-09th, 2024
Corresponding Author : aman.kec@gmail.com (A. Kumar)

By getting the information from video surveillance, zonal officials can directly interact and communicate with the community and also support monitoring and controlling some unauthorized activities that occur in the restricted zones. To collect the data required to monitor the restricted area safety and security, visual insights can be captured in the form of videos and images by the devices like drone cameras, and network cameras which can be further analyzed by some computer vision techniques to extract useful insights.

Some special cameras having additional features can be used for this purpose like IP cameras that can record live videos and send the digital data to a particular server for further processing. These cameras can be installed virtually or with a supporting device and they are capable of recording exterior and interior details of the location [3]. They can be used as a smart alerting system for security officers for unusual activity in any restricted zone. Video Surveillance can be applied to serve as an alert system for prohibited gathering zones, remote video monitoring, facility protection, loss prevention, facial recognition public safety, outdoor perimeter security, and many more [4].

In this study, we proposed an innovative and Intelligent Surveillance Framework for Crowd Detection and Alerting that supports security systems to enforce safety and security in red and secure zones. Referring, to secure and sensitive zones means the area secured in some special conditions viz. provide Covid-19 pandemic or Section-144 (in India) applied zones. The proposed system will be the adhesive to the various pre-installed surveillance systems, giving live feed to various controlling/observational units and recording live videos from the prohibited zones through drones or stable CCTVs. The live/recorded shots will be inputted into a computer vision-based deep learning model that will total the number of individuals gathered in a group by detecting their heads and testing them according to a decided threshold as per the situation. Finally, the system will notify and transmit an alert if more than the threshold count will be gathered at a place and under law enforcement rules.

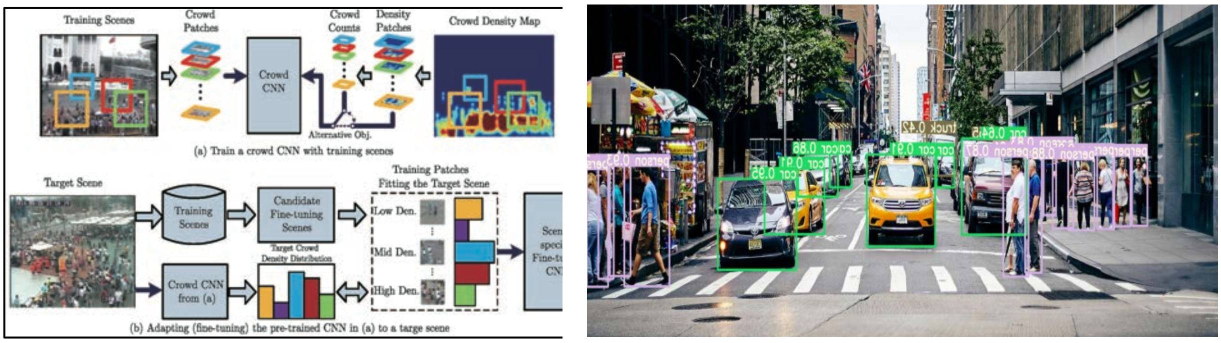
1.1Smart Vision Computing

Computer Vision has the ability to replicate human vision through computers. Computer vision is a sub-field of Artificial Intelligence that have the processing techniques to define various attributes for the data collected from various sources in the form of digital videos, and images [5]. The process of mimicking human visuals through computers involves the process of image acquisition, analysis, detection, and identification. This complete process helps a computer to learn and understand visual data and interpret it accordingly. It develops the ability of computers to understand symbolic information from digital data. With the continuous growth in Computer Vision technology, various application areas are getting served by these techniques like object detection, video tracking, pose and motion estimation, scene reconstruction, image restoration, etc [6].

Deep learning and convolutional neural network (CNN) have been proven as the most demanding and attractive technologies used by many researchers in the field of AI. In computer vision, deep learning was triggered by the influential work of reporting performance improvisations on object recognition on hard images by Krizhevsky et al. through CNN [7]. Deep learning techniques are the most precisely used technique that can work over a huge dataset by applying complex calculations through developed models [8]. Recently, many scientists got the benefits of deep learning techniques to evaluate and analyze complex calculations like the reconstruction of brain circuits of humans [9], analyzing DNA mutations [10], and many other cutting-edge scientific solutions provided by deep learning algorithms. Deep learning architecture follows the development and management of powerful neural networks to handle deep and large numbers of convolutional layers.

Object detection and counting are one of the application areas covered with deep learning techniques. It refers to a process of multiple tasks to identify various objects in digital images [27]. Deep learning object detection methods can be classified into two categories; one and two-stage methods [11]. In the one-stage object detection method, the model can directly predict and interpret the location of the object in the image by its related pixels and can directly classify the detected model effectively. Two-stage object detection methods perform the detection tasks in two steps by extracting ROI in an image according to the probability of the object existing in that image then, recognition and detection are completed with the help of candidate regions. One-stage methods perform detection tasks in less time than two-stage methods but show less accuracy in comparison to two-stage methods [12].

Object counting is one of the facets of advanced computer vision techniques. Figure 1 shows as the ability to count persons, vehicles, pedestrians, buildings, etc from digital images or live videos through deep learning algorithms [13][14]. The opportunities in this field can be merged with other techniques that convert raw data into a kind of structured data to support various services like video surveillance, traffic control systems, etc.



a. Object Detection & Counting using CNN[13]   b. Computer Vision Smart City Applications[14]

Figure 1: Development, Implementation & Application Examples

A thorough procedure for scene-specific crowd counting is shown in Figure 1 (a). First, a target scene is selected, and then a variety of training settings are put together. When crowd estimating is needed, the process for scene-specific crowd counting starts with carefully choosing a target scene that best matches the environment in which the model will be used. In addition, a wide range of training scenarios covering different types of crowd settings is assembled to facilitate effective model training. A comprehensive crowd CNN database is used to select one or more pre-trained CNN models as candidates for fine-tuning. This selection is made after taking into consideration the models' performance and their suitability for scenes that are comparable. By applying these candidate models to the training pictures, it becomes easier to create density distributions that show how the crowd densities are arranged spatially within each scene. Subsequently, the training pictures are carefully examined to extract patches that capture different aspects of crowd dispersion, including low, medium, and high densities. These patches are essential training data that help the chosen CNN models be refined. By means of an iterative process of fine-tuning, the parameters of the selected model are carefully modified in order to more closely correspond with the particular features and details of the target scene. A scene-specific CNN model that is precisely calibrated for precise crowd counting in the intended setting is ultimately produced as a result of this fine-tuning, which guarantees the model's adaptability to changing crowd densities and distributions. This customized method greatly improves the accuracy and dependability of crowd counting systems, enabling them to analyze and understand crowd dynamics in practical situations with greater effectiveness.

The remainder of this paper is organized as follows, section II is the Related Works and summary of the literature review for having various previous works done in hand so that the adhesiveness that we want to do toward making it a smart alerting system would be justifiable, next section III is demonstrating the design & description of the main contribution, section IV describing the main implementation aspects and results, the paper has been concluded with future scope in section V.

2. Related Works

This section presents a comprehensive literature evaluation of the most recent AI and Ml-based solutions that can be improved upon or used as the foundation for the intelligent alert and surveillance system. Finally, a very detailed summary of the results is provided in Table 1 of this section.

In 2022, Elharrouss et al. [15] proposed a CNN-based model for the number of persons in a crowd and generated a density map for the crowd. They have developed their model for crowd counting in for a football stadium. They have used the FSC-Set (Football supporters crowd) dataset that contains 7000 images of a variety of scenes of thousands of people. Their proposed model is applicable for other applications also as face recognition, object detection, etc. The obtained results show satisfactory performance outcomes.

In 2021,Celik et al. [16] proposed a model to observe and analyze the social distancing among gatherings. They have used the pedestrian detection method to find out the distance between people so that they can't violate social distancing rules in a sensitive place. In their work, they have also compared crowd-counting methods with pedestrian detection.

In 2021, Kumar et al. [25] proposed an artificial Intelligence (AI) and Machine Learning (ML) powered automated surveillance system that has great potential for implicit monitoring, enabling discrete classification, decision-making, and alerting procedures. They shows the capabilities of computer vision have been transformed by the incorporation of AI & ML approaches, especially in object detection and identification, enabling smarter surveillance systems. In this research, they offer a framework for a live video-based intelligent surveillance system. Using real-time object detection, this system seeks to enable automatic detection and notification of anomalous events. The platform provides a mechanism to build advanced surveillance systems that can detect anomalies in real-time streaming settings by utilizing AI and ML.

In 2021, Ahmed et al. [17] illustrated the pros and cons of various existing crowd detection and counting methods. They have used multiple sensitive areas where generally crowds can be seen like shopping malls, airports, religious or educational events, etc. Various automated video surveillance methods are analyzed and compared here. They have also proposed their own method to count and analyze the persons in a crowd. In the crowd analysis, they detected the age and gender attributes of the persons.

In 2021, Rajendran et al. [18] developed a smart AI-based crowd-controlling system that is necessarily required for various religious events like KumbhMela organized in Prayagraj. To minimize the dangerous impacts of uncontrolled gathering, an automated system is built that ensures a limited number of people gather in a limited area. They have used deep learning methods as the baseline models. Their proposed system can maintain the safety of people.

In 2020, Amine et al. [19] proposed a novel approach of 3D- CNN for feature extraction from Doppler images for the purpose of counting people in a closed environment. They have used the movements of persons to detect a human. They have developed WiFi radar that uses a 6 GHz-frequency band for detection of people. Experimental results show the model is achieving 89% accuracy in counting the crowd.

In 2020, Huang et al. [20] proposed a novel attention-based contextual CNN model that contains two basic components. Their model works as a dense dilated model that makes front and back layers tightly connected to hold the information in a scale-changing environment. Model feasibility is verified and compared against various existing models.

In 2020, Das et al. [21] developed an innovative deep-learning model to analyze the physical gathering of people in the crowd. They have completed the whole process in three phases. Initially, they have captured the local information by generating density maps, and then the counting feature is applied to those density maps. Finally, they have split crowd density maps to standing and sitting density maps. Results are showing the model is achieving excellent results in counting the gathering.

In 2020, Liu et al. [22] proposed CRNet (cross-stage refinement network) to refine density maps. Multiple fully connected layers are used in their CNN model to produce density output for correct predictions of crowd zones at different stages. Various data augmented methods are used to resolve multiple challenges of unconstrained scenes.

In 2019, Huynh et al. [23] proposed a novel and alternative way to count the crowd by handling scale variation issues. They have used various depth and crowd datasets to train their model to generate density and depth map estimations. They have reduced the labeling cost by using their proposed approach. Extensive experiments are performed to show the superiority of the proposed model over various existing models.

In 2018, Bailas et al. [24] explored and utilized video processing capabilities to show various image processing techniques in real-time crowd detection through low-powered devices. They have analyzed various deep learning algorithms for crowd counting and summarized the results to show the prediction is based on inputted scale video frames. They have implemented their framework using the CNN model to achieve state-of-the-art results in this domain.

Table 1: Literature Survey Summary

Author/ Year	Approach	Summary / Findings &Research Gap(s)
Elharrouss et al.,/ [2022]	CNN-based People Detection	• Detection is based on an early approach • Density map for the crowd • Low accuracy for highly crowded scenes
Celik et al.,/ [2021]	Deep learning-based smart Crowd Counting Method	• Model to observe and analyse the social distancing • Used the pedestrian detection method for distance measurement • Directly learn to regress the count • Less Interpretable
Ahmed et al.,/ [2021]	CNN-based Unique People Count Analysis	• High object framing • Lacks location information
Rajendran et al.,/ [2021]	L & T Smart World AI-based Crowd Management System	• Deep learning methods as the baseline models • Low accuracy in low crowd scenes
Amine et al.,/ [2020]	3D-Convolutional Neural Network (3D-CNN)	• Feature extraction from Doppler images for crowd count • Counting the people in a closed environment • Low image level count
Huang et al.,/ [2020]	Attention Based Contextual CNN Model (ACCNN)	• Hold the information in a scale-changing environment • Medium head indication
Das et al., / [2020]	Attention-based Deep Learning Framework	• Crowd density map to standing and sitting density maps • Less accurate for a highly crowded area
Liu et al.,/ [2020]	Cross-Stage Refinement Network (CRNet)	• Claimed refined density maps • Unconstrained scenes handing using data augmentation • Class imbalance problem
Huynh et al.,/ [2019]	Alternative Novel Way for Crowd Counting	• Handled scale variation issues • Reduced the labeling cost • Limited for an area
Bailas et al.,/ [2018]	CNN-based model for Crowd Detection	• Utilize video processing capabilities • The prediction is based on inputted scale video frames • Low image-level count
Kumar et al.,/ [2021]	SSE: A Smart Framework for Live Video Streaming based	• Real-time object detection and identification • Efficient anomaly detection without manual intervention, bolstering surveillance effectiveness. • Enhances security measures in monitored areas

	Alerting System	

3. Contribution – Smart Vision Adhesiveness toward an Intelligent Alerting System

From industrial operations to cyber-security, intelligent alerting systems are at the forefront of proactive risk mitigation and decision support. But these systems' ability to adhere to operational routines and organizational goals is just as important to their efficacy as their technical prowess. This study clarifies the ways in which smart vision principles are integrated into intelligent alerting systems to enhance their stickiness and facilitate their easy assimilation into organizational structures. The idea behind the Intelligent Surveillance Framework for Crowd Detection and Alerting is well-defined, serving as an application of object detection using deep learning models within the context of smart IoT-enabled security systems, particularly for future smart city scenarios. The framework's architecture is outlined as a three-layered structure, enhancing modularity and process handling efficiency. In Figure 2, the first layer, termed the user-interface layer, serves as the interface through which users interact with the system. This layer is responsible for receiving inputs, displaying relevant information, and facilitating user interactions. The second layer of the system encompasses AI and ML-based modeling. Here, sophisticated algorithms and machine learning models are employed to process incoming data, analyze patterns, and make intelligent decisions. This layer plays a crucial role in detecting crowd anomalies, identifying objects, and predicting potential security threats. Subsequently, the third layer represents the output layer, which translates the insights generated by the AI and ML models into actionable alerts. Specifically, in this context, the output layer focuses on object identification-based anomaly representation. When deviations from normal crowd behavior or the presence of suspicious objects are detected, an alert is triggered, signaling potential security risks. By employing this three-layer architecture, the Intelligent Surveillance Framework achieves enhanced modularity, allowing for efficient system management and scalability. Each layer is responsible for distinct functions, enabling streamlined processing of data and facilitating rapid response to security incidents. Moreover, the integration of deep learning-based object detection capabilities adds a layer of sophistication to the system, enabling it to adapt to evolving security threats and scenarios.

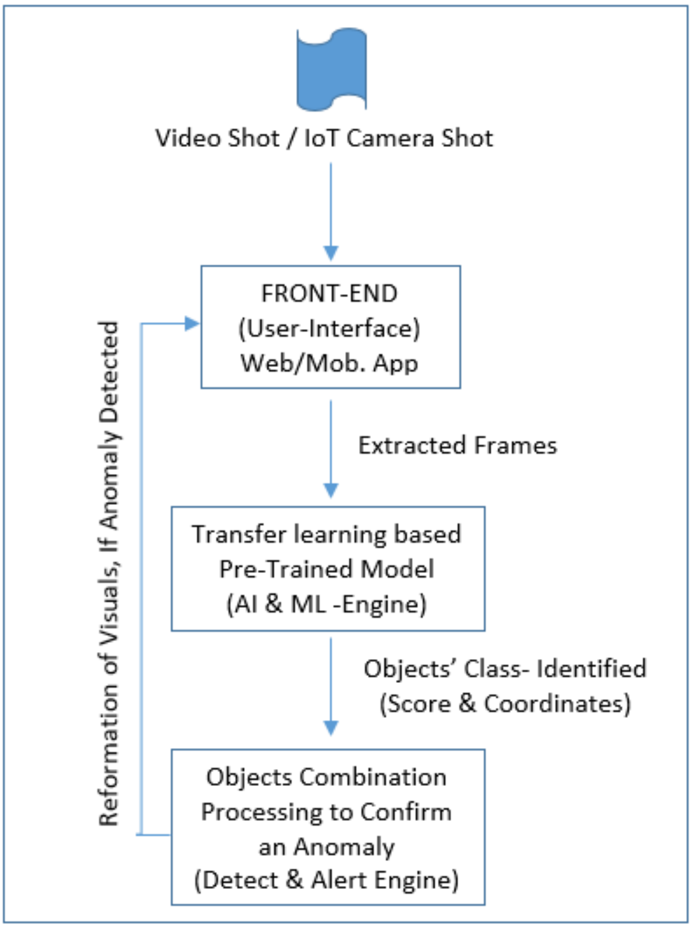


Fig. 2 High-Level View Diagram

The system's high-level schema has already been identified as the "black box" view that provides a holistic picture of the system. This high-level design schema also shows system modelling as straightforward three-layer architecture with a modular implementation strategy.

Figure 3 shows architectural paradigm that represents how AI and ML engine support is displayed at the next level of system construction and design while also emulating our earlier models [25]. The system architecture demonstrates a comprehensive three-tiered approach to intelligent surveillance and anomaly detection. Positioned at the initial tier is the input interface, serving as the ingress point for data acquisition. Here, live images are meticulously captured via an IP camera, ensuring a continuous stream of visual data for subsequent analysis. Transitioning to the middle tier, the AI and ML engine assumes a central role, constituting the computational backbone of the system. Equipped with meticulously trained models, this tier meticulously scrutinizes the extracted frames, employing sophisticated algorithms to discern odd classes or anomalies amidst the visual data. Leveraging cutting-edge deep learning techniques, the engine meticulously analyzes patterns and deviations from typical behavior, thus fostering a robust foundation for anomaly detection. Finally, at the output generation tier, the system meticulously processes the insights garnered from the AI and ML engine.

Here, a video is meticulously generated, meticulously highlighting the detected anomalies marked within the frames. This meticulously crafted video serves as a comprehensive visual representation of potential security threats, meticulously providing actionable insights for stakeholders. Presented seamlessly through the frontend interface or overlaid onto the initial interface, this detailed visualization empowers end-users with the information needed for timely responses and interventions. The system triggers alert notifications upon detecting unusual scenes or anomalies. These notifications are transmitted through a Django mobile app, providing real-time updates to designated users or security personnel. This ensures timely communication of potential security threats, empowering stakeholders to take prompt action as needed. In essence, this meticulously designed three-tiered architecture epitomizes efficiency and effectiveness, seamlessly integrating AI and ML capabilities to enhance anomaly detection and visualization within smart city surveillance frameworks.

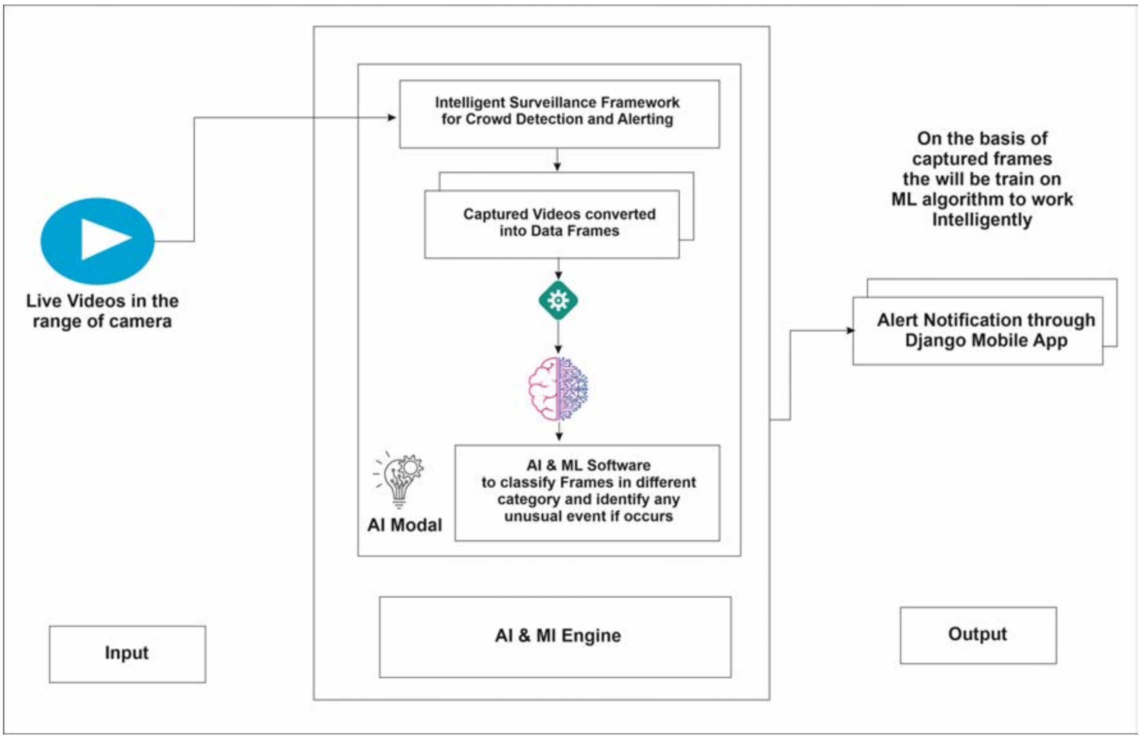


Fig. 3. Architectural Paradigm Proposed by Us[25]

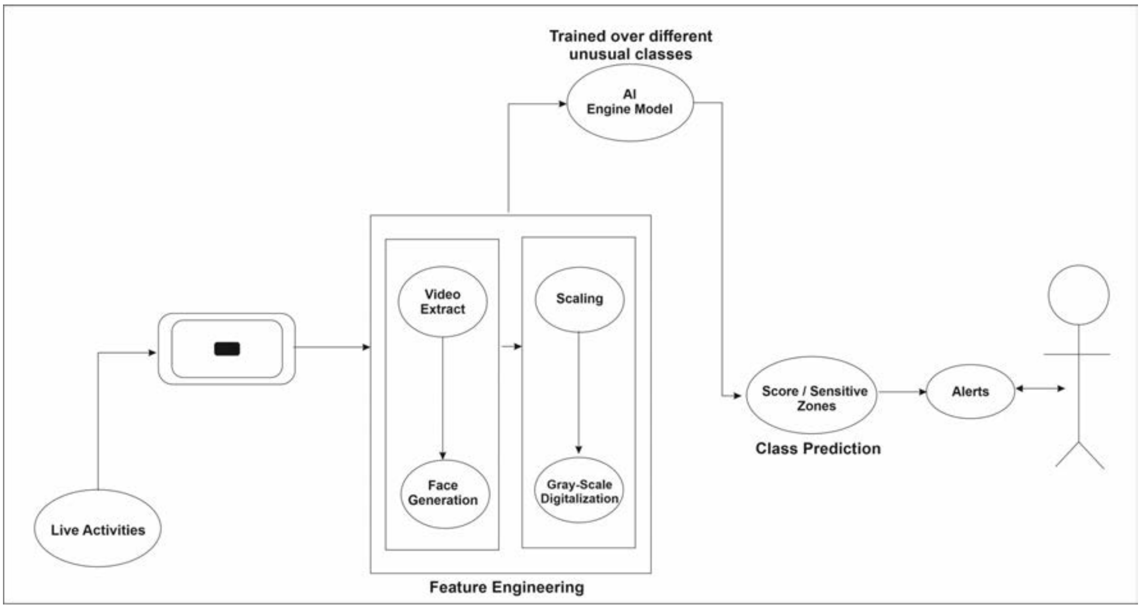


Fig 4. System Workflow

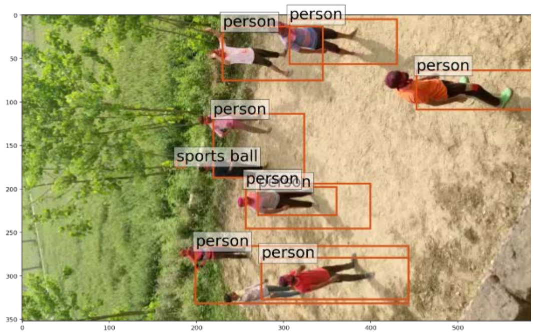
The use case for the modeled system is shown in Figure 4. In the use-case diagram depicting the operational workflow of an intelligent surveillance and anomaly detection system, users interact with the system to monitor security events and receive alert notifications. The system captures live activity through cameras installed in the monitored area, and this data is then analyzed by the AI Engine Models. Within these models, feature engineering is performed, involving the extraction of frames from video extracts and grayscale digitalization through scaling to facilitate efficient analysis. Additionally, a dedicated predictor class within the AI Engine Models is designed to identify anomalies within secure and sensitive zones, utilizing advanced algorithms and machine learning techniques.

Upon detecting anomalies or suspicious activities, the AI Engine Models generate alert notifications, which are promptly displayed to users through a user interface. Users can monitor these alerts in real-time and take appropriate actions as necessary to address potential security breaches or anomalies. In essence, this use-case diagram illustrates the seamless integration of live activity capture, AI-driven analysis, and alert notification generation within an intelligent surveillance system, enhancing security monitoring capabilities in sensitive environments.

4. Implementation Aspects- Experiments and Results

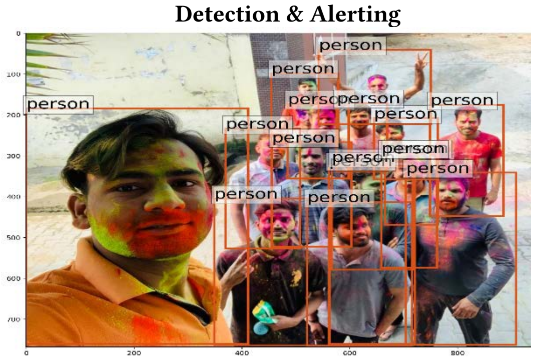
As AI & ML-based modelling is a data and computing-intensive process. The framework initially experimented with pre-trained models using YOLO v3[26]. Pre-trained modelling appears to be used for compressive AI & ML-based modelling employing the transfer-learning concept for hardware-constrained machines and to make smart computing accessible with less computational equipment as well for the creation of intelligent/smart systems. As far as results are concerned, current implementation and experiments are good enough to release the idea of a smart alerts system as the first document, so the paper is. Below Figure 5(a,b),5(c,d) and 5(e,f) are the figures taken from the current results, more specifically, Figure 5(a, c and e) are the frames from a small test video before processing, after processing the same frames (in Figure 5(b, d and f) showing the marking of some persons, ball, colored faces, firemore than a pre-set threshold value, in a single frame, the AI engine that is in back-end have detected it as one kind of anomaly of secure and sensitive zone violation and mapped to a formatted alert also propagate to the concerned. Based on the Smart-Vision and Object Detection system itself generation and propagating alerts to the concerned authorities or a control center.

a. **Pre-processed frame from a gathering**



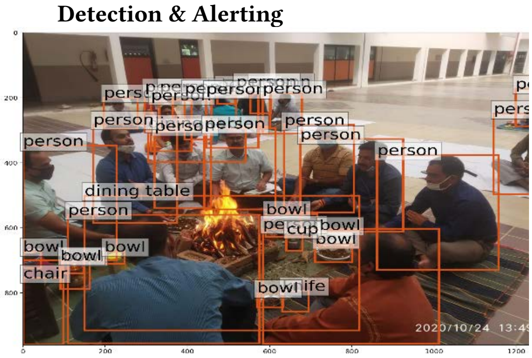
b. **Processed frame before a fight by gathering due to a ball Anomaly Detection & Alerting**

c. **Pre-Processed Colored Gathering Sample Picture**



d. **Colored / Hidden Crow Anomaly Detection & Alerting**

e. **Fire and Crowd Symbolic Sample Picture**



f. **Fire and Crowd Anomaly Detection & Alerting**

Figure 5 Experiments and Results' Images

5. SUMMARY& FUTURE WORK

This paper presents the implementation and testing of a new AI & ML-based self-alerting system designed for secure, red alert, or sensitive zones. The tested AI & ML-based Intelligent Surveillance Framework for Crowd Detection and Alerting represents a vital first step towards the creation of intelligent systems meant to ensure adherence to laws controlling crowd gatherings in areas with restricted access or during special events. Our platform offers a reliable method for identifying irregularities in crowds and promptly sending out alerts by utilizing cutting-edge artificial intelligence and machine learning algorithms. Furthermore, our implementation is part of a larger research and development project aimed at improving alerting and anomaly detection capabilities. While the logical outcomes of our current implementation seem promising, we recognize that more work needs to be done, especially to improve the system's analytical and logical reasoning skills. Subsequent research endeavors will center on enhancing the system's aptitude to identify intricate connections among identified objects and formulate well-informed judgments grounded in contextual data. Although there is still room for improvement, our system is still a useful tool for notifying relevant authorities when crowd sizes exceed pre-established thresholds, allowing for timely and appropriate reactions to possible security issues. Apart from security enforcement, our Intelligent Surveillance Framework has potential applications in multiple other sectors such as public safety, event management, and urban planning. Our goal is to provide a highly reliable and flexible solution by continuously improving our system's analytical skills and integrating feedback from real-world deployments. Furthermore, by laying the foundation for future developments in AI-driven surveillance systems, our framework opens the door to proactive approaches to danger detection and crowd management.

References

[1] Z. Ali, S. A. Chaudhry, M. S. Ramzan and F. Al-Turjman, "Securing Smart City Surveillance: A Lightweight Authentication Mechanism for Unmanned Vehicles," in IEEE Access, vol. 8, pp. 43711-43724, 2020, doi: 10.1109/ACCESS.2020.2977817.

[2] Code of Criminal Procedure, 1973. Section 144." Government of India.

[3] V. Babanne, N. S. Mahajan, R. L. Sharma, and P. P. Gargate, "Machine learning based Smart Surveillance System," 2019 Third International Conference on I-SMAC (IoT in Social, Mobile, Analytics, and Cloud) (I-SMAC), 2019, pp. 84-86, doi: 10.1109/I-SMAC47947.2019.9032428.

[4] R. Patrick and N. Bourbakis, "Surveillance Systems for Smart Homes: A Comparative Survey," 2009 21st IEEE International Conference on Tools with Artificial Intelligence, 2009, pp. 248-252, doi: 10.1109/ICTAI.2009.93.

[5] Y. Xing et al., "Advances in Vision-Based Lane Detection: Algorithms, Integration, Assessment, and Perspectives on ACP-Based Parallel Vision," in IEEE/CAA Journal of Automatica Sinica, vol. 5, no. 3, pp. 645-661, May 2018, doi: 10.1109/JAS.2018.7511063.

[6] B. Zhang, "Computer vision vs. human vision," 9th IEEE International Conference on Cognitive Informatics (ICCI'10), 2010, pp. 3-3, doi: 10.1109/COGINF.2010.5599750.

[7] A. Krizhevsky, I. Sutskever, and G. E. Hinton, ‘‘ImageNet classification with deep convolutional neural networks,’’ in Proc. Adv. Neural Inf. Process. Syst., 2012, pp. 1097–1105.

[8] Y. LeCun, Y. Bengio, and G. Hinton, ‘‘Deep learning,’’ Nature, vol. 521, pp. 436–444, May 2015.

[9] M. Helmstaedter, K. L. Briggman, S. C. Turaga, V. Jain, H. S. Seung, and W. Denk, ‘‘Connectomic reconstruction of the inner plexiform layer in the mouse retina,’’ Nature, vol. 500, pp. 168–174, Aug. 2013.
[10] H. Y. Xiong, B. Alipanahi, L. J. Lee, H. Bretschneider, D. Merico, R. K. C. Yuen, Y. Hua, S. Gueroussov, H. S. Najafabadi, T. R. Hughes, Q. Morris, Y. Barash, A. R. Krainer, N. Jojic, S. W. Scherer, B. J. Blencowe, and B. J. Frey, ‘‘The human splicing code reveals new insights into the genetic determinants of disease,’’ Science, vol. 347, no. 6218, Jan. 2015.
[11] G. Li, Y. Yang, and X. Qu, ‘‘Deep learning approaches on pedestrian detection in hazy weather,’’ IEEE Trans. Ind. Electron., vol. 67, no. 10, pp. 8889–8899, Oct. 2020, doi: 10.1109/TIE.2019.2945295
[12] Z.-Q. Zhao, P. Zheng, S.-T. Xu, and X. Wu, ‘‘Object detection with deep learning: A review,’’ IEEE Trans. Neural Netw. Learn. Syst., vol. 30, no. 11, pp. 3212–3232, Nov. 2019
[13] Y. Yang and W. Gao, "A Method of Pedestrians Counting Based on Deep Learning," 2019 3rd International Conference on Electronic Information Technology and Computer Engineering (EITCE), 2019, pp. 2010-2013, doi: 10.1109/EITCE47263.2019.9094838.
[14] Viso.ai, "Computer Vision in Smart City Applications," Viso.ai. [Online]. Available: https://viso.ai/applications/computer-vision-in-smart-city-applications/. [Accessed: March, 2024].
[15] O. Elharrouss, N. Almaadeed, K. Abualsaud, S. Al-Maadeed, A. Al-Ali and A. Mohamed, "FSC-Set: Counting, Localization of Football Supporters Crowd in the Stadiums," in IEEE Access, vol. 10, pp. 10445-10459, 2022, doi: 10.1109/ACCESS.2022.3144607.
[16] M. Çelik and A. H. Örnek, "Comparing Crowd Counting and Pedestrian Detection Methods on Social Distance Detection," 2021 15th Turkish National Software Engineering Symposium (UYMS), 2021, pp. 1-5, doi: 10.1109/UYMS54260.2021.9659743.
[17] Ahmed, P. Bansal, A. Khan and N. Purohit, "Crowd Detection and Analysis for Surveillance Videos using Deep Learning," 2021 Second International Conference on Electronics and Sustainable Communication Systems (ICESC), 2021, pp. 1-7, doi: 10.1109/ICESC51422.2021.9532683.
[18] L. Rajendran and R. Shyam Shankaran, "Bigdata Enabled Realtime Crowd Surveillance Using Artificial Intelligence And Deep Learning," 2021 IEEE International Conference on Big Data and Smart Computing (BigComp), 2021, pp. 129-132, doi: 10.1109/BigComp51126.2021.00032.
[19] El Amine and V. Guillet, "Device-Free People Counting Using 5 GHz Wi-Fi Radar in Indoor Environment with Deep Learning," 2020 IEEE Globecom Workshops (GC Wkshps, 2020, pp. 1-6, doi: 10.1109/GCWkshps50303.2020.9367393.
[20] Y. Huang, A. Zhu, G. Duan, F. Hu and Y. Li, "ACCNet: Attention-based Contextual Convolutional Network for Crowd Counting," 2020 Chinese Automation Congress (CAC), 2020, pp. 1926-1931, doi: 10.1109/CAC51589.2020.9327438.
[21] S. S. Sarathi Das, S. M. Mukit Rashid and M. E. Ali, "CCCNet: An Attention Based Deep Learning Framework for Categorized Counting of Crowd in Different Body States," 2020 International Joint Conference on Neural Networks (IJCNN), 2020, pp. 1-8, doi: 10.1109/IJCNN48605.2020.9207370.

[22] Y. Liu et al., "Crowd Counting Via Cross-Stage Refinement Networks," in IEEE Transactions on Image Processing, vol. 29, pp. 6800-6812, 2020, doi: 10.1109/TIP.2020.2994410.
[23] V. Huynh, V. Tran and C. Huang, "DAnet: Depth-Aware Network for Crowd Counting," 2019 IEEE International Conference on Image Processing (ICIP), 2019, pp. 3001-3005, doi: 10.1109/ICIP.2019.8804413.
[24] C. Bailas, M. Marsden, D. Zhang, N. E. O'Connor and S. Little, "Performance of video processing at the edge for crowd-monitoring applications," 2018 IEEE 4th World Forum on Internet of Things (WF-IoT), 2018, pp. 482-487, doi: 10.1109/WF-IoT.2018.8355170.
[25] A. Kumar, H. Hashmi, S. A. Khan and S. Kazim Naqvi, "SSE: A Smart Framework for Live Video Streaming based Alerting System," 2021 10th International Conference on System Modeling & Advancement in Research Trends (SMART), 2021, pp. 193-197, doi: 10.1109/SMART52563.2021.9675306.
[26] Redmon, Joseph, and Ali Farhadi. "YOLOv3: An Incremental Improvement." arXiv preprint arXiv:1804.02767 (2018).
[27] H. Hashmi, R. K. Dwivedi and A. Kumar, "Identification of Objects using AI & ML Approaches: State-of-the-Art," 2021 10th International Conference on System Modeling & Advancement in Research Trends (SMART), MORADABAD, India, 2021, pp. 1-5, doi: 10.1109/SMART52563.2021.9676273.