Digital Twin for Temple Security and Crowd Management

Ayush Bisen[1], Madhura Rewatkar[2], Sai Konde[3], Monali Gulhane[4], Nitin Rakesh[5], Akhil Gupta[6]

{ayushbisen321@gmail.com[1], madhurarewatkar1076@gmail.com[2], saikonde03@gmail.com[3], monali.gulhane4@gmail.com[4], nitin.rakesh@gmail.com[5], akhilgupta112001@gmail.com[6]}

Symbiosis Institute of Technology, Nagpur Campus, Symbiosis International University, Pune, Maharashtra, India[1, 2,3,4,5,6]

Abstract. The article describes a system based on Digital Twin technology that resolves security and crowd management issues in temple zones, where large gatherings usually result in such problems. The system's main feature is the real-time tracking of individuals through YOLO-based object detection, along with the detection of left-behind objects to quickly identify possible security threats. In addition to the detection of objects, the system also provides information about the crowd density. This data is then rendered as a 3D interactive spatial map that always clearly shows the distribution of the crowd. After the detection of either a case of overcrowding or a suspicious object, the system automatically activates alarms that enable the prompt response of security staff. The entire system is connected to a monitoring dashboard for visualization and control in real-time, which is developed on Dash and further incorporated within a Wix-based web interface to guarantee easy access and operational flexibility. The suggested methodology results in a strong Digital Twin framework for the improvement of situational awareness, the elimination of manual monitoring, and the betterment of safety management in sensitive and heavily populated areas like temples.

Keywords: Digital Twin, YOLO, Object Detection, Crowd Management, Temple Security, Real-time Alerts, Dash UI, Wix Integration

1 Introduction

The protection of public places and the security of such places as temples, which are mostly visited due to their religious and cultural significance, are among the major concerns of the modern world. These places are prone to massive gatherings during festivals and rituals as well as the peak visiting times, hence, the risk of overcrowding, unattended luggage and unauthorized access, all of which have serious safety implications, are raised. As detailed in Table I, traditional security methods in such environments heavily rely on human observation through personnel, which is prone to errors, fatigue, and slow response times especially during stressful moments [1].

To eliminate the above drawbacks an all-encompassing AI-based security and crowd control system that is specifically designed for temple grounds has been proposed in this paper. The system applies YOLO-based object detection for real-time human tracking and also detects possibly left behind or suspicious items in different areas of the temple site. The system is able

IICTDS 2025, December 19-20, Chandigarh, India
Copyright © 2026 EAI
DOI 10.4108/eai.19-12-2025.2362624

to automatically alert and minimize the human participation whenever it notices incompatible activity patterns or security threats, thus security staff can respond faster [2].

The system is complemented by a user-friendly Dash-based interface and a Wix-powered web dashboard. This platform provides the administrators with the facilities to access live surveillance feeds, get notified, see the crowd density in real time, and change the system parameters all from a single centralized location.

The system not only boosts real-time situational awareness by automating some of the vital aspects of monitoring and detection but also, to a large extent, eliminates the risks caused by human delay or neglect. It offers a number of benefits to the temple managers such as the efficient and secure way of handling big crowds, no security incidents happening, and the staff and worshippers being in a safe environment. The document describes in detail the design of the system, its technical implementation, and the actual possibility of changing the security management in temples through planting.

1.1 Motivation

Temples and other worship places are frequented by huge crowds, making them open to various security risks such as lost property, overcrowding, and illegal activities. Conventional monitoring systems depend a great deal on human supervision, which is not efficient, prone to mistakes, and demands constant human monitoring. It is predicted that there will be an AI-powered security system able to detect and respond to security threats by itself in real-time. Eventually, critical issues mentioned below have to be solved which are the main reasons why project should be undertaken:

• Prevention of Possible Harms: Unattended items in public places are a potential risk. Automatic detection of unattended items not only prevents the risks but also allows the authorities to take necessary measures beforehand for example, using unattended items as a source of theft or fire through burning accidents, etc.

• Better Management of Crowds: The negative consequences of overcrowding are stampedes and mismanagement. The management can take preventive and even corrective steps in a timely manner by monitoring the crowd density in real-time.

• Lessening of Human Dependence: The security personnel will be able to focus on more critical areas as a result of the surveillance automation since the monitoring will be done automatically. This will not only minimize human errors but also maximize operational efficiency.

• Improving Security at the Temple: AI not only learns from past incidents but also reports them immediately therefore, in case the authorities are alerted by AI, they can respond to the incident promptly and even prevent it from happening.

1.2 Contribution

Our work is crucial for attaining more sophisticated security and crowd management systems that are AI-powered for such delicate public venues as temples and religious centers. These areas normally carry the issues regarding their cultural value, as well as there are many visitors and very little physical security. The system that we suggest is based on traditional supervision,

but AI, computer vision, and visualization in real-time are also taking part in this execution as being more creative and asserting in handling the safety and security of the crowd.

Notable aspects of our project include:

1) Person Detection with AI

The YOLO (You Only Look Once) deep learning model is utilized by the system for rapid detection of people in the video surveillance feeds. This feature, shown in Fig. 1, enables continuous guarding of people qualitatively, and thus allows the if and then workflows of the security staff to be enriched with quick and timely actions by spotting the areas of foreseen jam or illicit entry and then intervening before the issue spreads.

2) Detection of Abandoned Objects

The implementation of the object-tracking and anomaly-detection techniques make it possible for the system to instantly spot post fader or left objects through the project by the operation of these techniques. This functionality is depicted in Fig. 5, and it is very critical in the prevention of security threats like lost bags, suspicious packages, or sometimes even people's stuff that might get unnoticed in a very crowded temple environment due to their being unnoticed.

3) Real-Time Visualization of Crowd Density

The system adopts grid-based 3D visualization techniques that convert raw detection data into clear and separated visual feedback. As a result, security personnel can not only see how many people there are but also how they are distributed spatially over different regions in the surveillance area, which helps in quicker locating of overcrowded areas or bottlenecks in different camera streams. This is illustrated in Fig. 2.

4) Feed Integration with Multiprocessing

The architecture of our system employs multiprocessing techniques to monitor several surveillance cameras at the same time, thus, enabling monitoring of various angles and locations at no performance cost, as depicted in [3]. Consequently, it becomes a system capable of scaling and hence suitable for the wide surveillance coverage of large temple premises.

5) Automated Security Alerts

The detection of potential security threats caused by, for instance, people gathering or seeing trash has made the system issue real-time notifications automatically. According to Fig. 3, Telegram is the communication channel that is used here, as administrators of the temple and security personnel are notified immediately via it, so they can react right away, thus not needing to constantly watch [4].

6) Intuitive Monitoring Dashboard

An uncomplicated, engaging interface that works based on Dash has been developed and incorporated with a Wix site. The supervisors of the temple can monitor the surveillance live feed, receive alerts, determine the amount of people, and access old records using any gadget that is connected to the internet. The system has been designed in such a way that it is very easy to use and even the non-tech users can access it without much difficulty [5].

1.3 Organization

Below is a description of our project:

• Introduction – Provides an overview of the project, its significance, and the problem it addresses.

• Motivation – Explains the reasons for choosing this topic, emphasizing the need for AI-driven security in religious spaces.

• Related Work – Discusses existing crowd monitoring and security surveillance systems, highlighting their limitations and how our approach differs.

• Implementation Details – Describes the technologies and methodologies used, including:

o YOLO for Object Detection

o Telegram API for Alerts

o Dash & Wix for the User Interface

o Multiprocessing for Efficient Video Processing

2 Background Survey and Theory

A thorough background survey was conducted to prepare and implement this project.

2.1 Guide to Crowd Monitoring

Monitoring of crowds has become an important area of research and practice especially in the field of public safety at large public places like religious sites, concerts, festivals, sports events, and political rallies and also in transportation hubs. The growing concentration of people in urban and semi-suburban areas leads to new safety concerns. Among these, there are risks of overcrowding, stampedes, unauthorized entry, and security violations.

In the past, the monitoring of crowds was carried out primarily by human observation. This was done through the installation of CCTV and the use of stationary threshold-based alarm systems. The number of people in a given area or a specific pattern of behavior must exceed a pre-set level before an alarm is triggered either manually or automatically. Despite the fact that these traditional methods have been efficient, they still have certain limitations that are inherent in their nature:

• Scalability: Human observers can only monitor a limited number of screens or sections simultaneously; hence, large-scale surveillance is inefficient.

• Fatigue and Human Error: The longer the observation period, the less alertness there is and the more the key events are likely to be missed.

• Latency: Detection and response time is generally procrastinated, especially when human judgment or simple rule-based systems are involved.

The previously mentioned restrictions were at the same time the cause of the emergence of the technologies like Computer Vision, Deep Learning, and IoT which have completely

transformed the crowd monitoring perception. The aforementioned technologies make it possible to develop the next-generation automated systems which would have the capability to, in a completely automatic manner, detect and follow individuals in real-time, and conduct the analysis of their behaviour even at the different locations without any manual supervision.

Among all innovations, one that stands out the most is the usage of YOLO (You Only Look Once). A highly precise object detection algorithm, YOLO, is capable to not only count but also recognize people in video streams even if the lighting and crowd conditions are difficult. When combined with machine learning algorithms for predicting behavior and supported by IoT-capable sensors for environment monitoring (temperature, noise levels, motion), these systems can provide real-time information on:

• Distribution of crowd density

• Anomalous patterns of movement

• Detection of abandoned objects

• Risk forecasting for congestion or overpopulation

Moreover, today's systems are sophisticated enough to perform various tasks simultaneously, such as alerting security staff to potential threats, capturing important information for post-incident analysis, and connecting with data visualization tools (like Tableau or Dash-based dashboards) to improve awareness and decisions during the crisis.

In the process of public areas gradually getting larger and more intricate, AI-driven crowd surveillance systems are now regarded as a vital tool for the improvement of both security and efficiency for the operations. Likewise, the systems lessened the dependence on the staff presence, offered faster and more precise detection, and acted as a barrier to the emerging of incidents by taking risks down.

2.2 Basic Principles

The basic principle containing the algorithm, base models, and their use is listed.

• YOLO (You Only Look Once): A very fast and very precise real-time object detection algorithm. YOLOv8 is the latest and most perfect version, which is the best for detecting people in video streams in real-time.[1].

• Crowd Density Estimation: is defined as the number of individuals per unit area of a frame. It can be computed as:

$$\text{Crowd Density/ Area of Frame} = \text{Number of People}$$

• Three-Dimensional Visualization: Addition of a third dimension (Z) of density or elevation to the spatial graph makes it simple to determine congestion hotspots.

• Abandoned Object Detection: The method applies object tracking and temporal analysis for the detection of abandoned objects which in turn increases the surveillance application’s accuracy.

• UAV/Human Response Visualization: Threshold breaches trigger a virtual marker on the 3D map where human or drone intervention is necessary.

3) Theoretical Foundation

• Bounding Boxes → Spatial Density Grids: The locations of detected people are mapped to grids across the whole frame for the purpose of computing the density regionwise.

• Multi-feed Integration: Each video feed is mapped with a spatial offset in the 3D graph to show its area so that all parts of the crowd can be viewed at the same time.

• Real-Time Feedback Loop: A Flask API provides the interaction between the processing unit (main.py) and the Dash visualization dashboard resulting in the almost instantaneous updating of density visuals.

Table 1. Comparison of different methods

Feature	**Traditional CCTV+Manual**	**Static Threshold**	**Deep Learning**	**Our System**
Automation	Manual	Partial	Full	Full
Real-time Performance	Delayed	Moderate	High	Very High
Spatial Visualization	None	2D View	2D Heatmaps	3D Real-Time
Multi-feed Support	None	--	Partial	Full
Alert System	Human-based	Hard-coded	Configurable	Live-Dynamic
Abandoned Object Detection	None	--	Possible	Integrated
UAV/Human Indicator	Not Available		Experiential	Visualized via Black Dot
Scalability	Poor	Medium	High	Very High

2.3 Problem Description

The international trend of public meetings in all areas—religious assemblies, cultural festivals, concerts, political gatherings, or at transit points—continues to increase their frequency and size, and thus the necessity of practical, reliable, and smart crowd control and monitoring systems has grown immensely. Moreover, it is safe to say that public safety in crowded places is no longer just a matter of logistics; rather, security problems are coming up, mainly because of the increased occurrence of stampedes, crowd rushes, abandoned object attacks, and unauthorized access.

Traditionally, most public areas have been monitored with the help of outdated surveillance systems where security personnel in a central control room are continuously watching the CCTV feeds. While such arrangements have provided the basic security required, they have become insufficient and expensive recently, especially when dealing with huge crowds or mobile gatherings.

Few among the most significant disadvantages of the usual surveillance system include:

a) Human Tiredness and Slow Reaction:

Continuous human supervision causes unavoidable tiredness, increasing the likelihood of mistakes, taking longer to respond, and some incidents being overlooked, particularly during long shifts or at times of high activity.

b) No Real-Time Spatial Awareness:

The static camera monitoring system only presents an unconnected visual observation of the site, thus limiting the possibility of forecasting crowding or detecting crowding trends in larger or more complicated places.

Typically, older systems utilize straightforward, rule-based alerts (like motion or camera tampering) which lack the ability to spot crowd behavior anomalies like stampedes, quick dispersal, or group fights.

c) No Multi-Feed, Multi-Dimensional Visualization:

Operators may be limited to watching a variety of camera feeds on separate monitors that do not provide a common view, making it very difficult to discover the general patterns and to be aware of the overall situation during the whole process in real time.

The risk that large crowds pose — through, among others, stampedes and crowd crushes, as well as the threat of bombs hidden inside or under abandoned objects — is so great that an automated intelligent crowd monitoring system is needed immediately. Such a system should be able to use the data from the actual surveillance and process it in real-time, all the while providing results that can be acted upon.

The system mentioned above is able to resolve all the difficulties that have been raised through the following functionality:

a) Real Time Person Detection & Local Density Estimation:

The system estimation of local and global crowd density is based on the YOLO-based object detection being used to detect and count people uninterruptedly. This, in turn, allows for making the proactive crowd control decisions.

b) Anomaly Detection for Fights and Unusual Movement:

The system is capable of detecting conflicts or abnormal movements before they escalate by analyzing the movement of people and sudden changes in the flow of people.

c) Abandoned Object Detection:

An object is detected first and then its position is monitored to see whether it is in the pre-specified boundary or not, hence reducing the risk of security breach and also allowing fast response.

d) Concurrent Multi-Feed Video Handling:

The system works with the parallel input from various cameras, thus it provides wide-area monitoring without the quality being affected due to the use of multiprocessing techniques.

e) 3D Interactive Visualization of Crowd Data:

The real-time display of crowd density and movement patterns with 3D grid-based visualization enables the administrators to directly, intuitively grasp the occupancy and flow of space. As mentioned in [6].

f) Automated Security Notifications:

At the point of critical levels - be it of density, motion or object detection - the system, as depicted in Fig. 3, sends out automated alerts via Telegram or other communication devices so that the security personnel can react instantly.

Through introducing this automated and innovative system, the security and situational awareness are improved, operational human personnel burden is reduced, and response times are significantly decreased which eventually leads to a safer public event place [7].

3 Proposed Model

The new approach eliminates the drawbacks of all prior techniques and enables the making of an efficient and precise system.

1) Input Feed Handler:

It has the capability to handle various input types (webcams, IP cameras, video files). It employs Python's multiprocessing technique to process several video feeds simultaneously and in the background. Each feed is assigned an ID that is distinctive and corresponds to a specific region in the visualization [8].

2) Person Detection with YOLOv8:

The Ultralytics YOLOv8 model is used to detect the presence of humans in every frame. The model draws bounding boxes around all the individuals detected. It also computes the total number of individuals present in a frame [9-11].

3) Crowd Density Calculation: The Frame is segmented into spatial grids (e.g., 3×3 or 4×4).For every grid:

Computes local density by finding out how many people are in it. Normalizes based on the area to get density per region. Each individual's bounding box is utilized to estimate position and to build grid density [12].

4) Anomaly Detection

Fight Detection: Estimates optical flow (intensity of movement) and identifies violent movement. As discussed in [13].

Abandoned Object Detection: Tracks objects that remain stationary for extended periods using frame differencing and motion persistence.

Overcrowding Detection: The local grid density crosses a predefined threshold, raising alarms.

5) Real-Time Alert System: If any area exceeds the risk threshold:

A Telegram notification is sent to the security guards or the control room. A black dot representing the UAV/human response is plotted on the 3D graph in that specific region. Shown in Fig. 3 and Fig. 4.

6) Three-Dimensional Visualization of Graphs Using Dash and Plotly:

Live updates every 5-10 seconds. Information from all the video streams is represented in a single integrated three-dimensional plot.

X and Y: Location in space.

Z: Person height or crowd density.

Color-coded Dots: Green (low), Blue (moderate), Red (high), Black (UAV/Human response). As discussed in [6].

Allows the user to inspect areas of congestion by feeding visually. As shown in Fig. 4.

7) Data Synchronization and Communication

The Flask server acts as the go-between that intercepts crowd data from main.py. Dash UI fetches and displays the data.

Feed-specific offsets deliver non-overlapping graphs within the common space.

8) Scalability

Extensible to any number of feeds easily. Can be combined with crowd forecast algorithms or thermal imaging inputs in the future.

4 Prototype Results

This section covers the results that have been achieved.

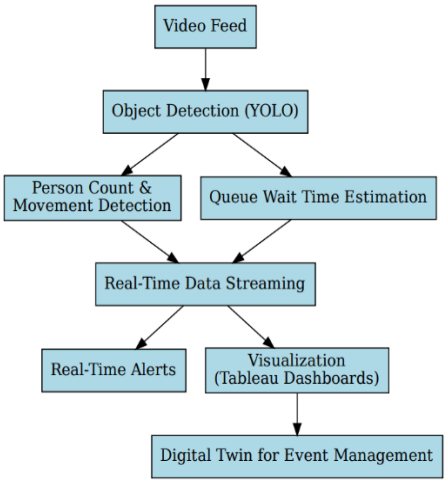


Fig.1. Block Diagram and Flow of Project

The flowchart in Figure 1 is a simple block diagram explaining all the functions this system achieves.

The flowchart illustrates the complete data flow and system architecture of the AI-based Crowd Monitoring and Security System for temple environments (or other public spaces). Each step represents a key real-time detection and alert mechanism stage, forming the Digital Twin framework.



Fig. 2. Live Crowd Detection

This image shows a real-time crowd monitoring system using YOLO-based object detection. Each person in the frame is detected and marked with a bounding box and a confidence score. The top-left corner displays live metrics:

• People Count: 15

• Crowd Density: 7.23

• Average Movement: 2.71

• Flow Density: 5.39

These values help assess crowd size, movement, and congestion in real-time, aiding in efficient crowd management and safety alerts.



Fig. 3. Live Telegram Alerts

This image shown in Figure 3 is an example of automated security alerts sent via Telegram. The system has detected a possible fight in Feed 0 and has triggered multiple urgent notifications:

"Fight detected in Feed 0! Please investigate immediately!"

This kind of alert helps security teams react instantly to violent or suspicious activity, allowing quick intervention and improving overall safety in real-time.

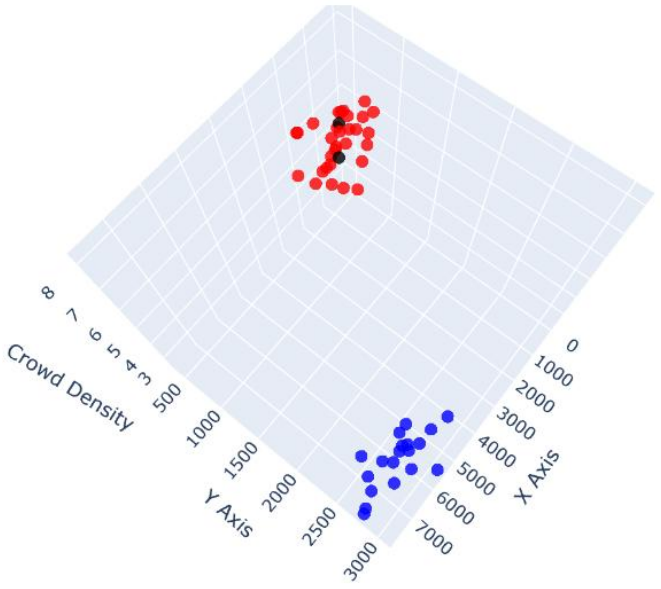


Fig. 4. 3D Graph of Video Feed

This image shows automated overcrowding alerts sent via Telegram. The system detected high crowd density for example, 19 people in Feed 0 and 16 people in Feed 1 and immediately warned the user to take action. The video feed of 3D graph is shown in Figure 4.

These alerts help prevent congestion, manage crowd flow, and avoid real-time safety risks.



Fig. 5 Abandoned Object Detection

Figure 5 shows live detection of abandoned objects in the frame. This also triggers an alert sent via telegram to security personnel.

5 Conclusion

The proposed system used for AI-based security and crowd management is a scalable, cost-effective, and innovative solution to surveillance, especially for public places with a lot of people like, temples, religious events, event centers, and so on. It meets the demand for a modern surveillance system by up to a great extent replacing or complementing the conventional manual monitoring methods. The main component of this system is the YOLO (You Only Look Once) - which is an advanced object detection algorithm. The system is capable of identifying and tracking individuals in real-time and keeping the situational awareness ongoing. This is the basis for the primary characteristics such as the counting of persons, detection of movement patterns, and analysis of crowd behavior [14]. Among the features to be highlighted is the detection of abandoned object, which is very important in environments where there are temples. The system by automatically detecting objects that have been left unattended for a long time assist to prevent security threats such as misplaced belongings or hazardous materials without the need for human observation constantly.In this paper the support of multiple feeds paired with the automated alert system allows the operators to be notified instantly about any irregularities or crowd surges from multiple feeds at the same time. This not only enhances reaction times but also reduces human errors which are for the most part the biggest drawback of surveillance systems that are based solely on human input. To sum it up, it changes traditional monitoring into a digital twin for event and place management that is interactive and supporting the smart future public infrastructure development.

References

[1] A. T. Islam, J. Pudota and R. Tomar, "Accurate Localization Based on Non-Visual Sensor Data for Digital Twin Model," 2025 IEEE International Conference on Artificial Intelligence and eXtended and Virtual Reality (AIxVR), Lisbon, Portugal, 2025, pp. 125-129

[2] C. Zhang, Y. Gao, M. Ye, M. Zhang, B. Wang and Z. Li, "A Risk evaluation and control method of highway based on digital twin," 2023 IEEE 3rd International Conference on Digital Twins and Parallel Intelligence (DTPI), Orlando, FL, USA, 2023, pp. 1-5

[3] J. Qiao et al., "Research on Digital Twin Modeling Technology for Distribution Networks Integrating Multiple Digital Resources," 2023 International Conference on Intelligent Computing, Communication & Convergence (ICI3C), Bhubaneswar, India, 2023, pp. 312-317

[4] Q. Chen et al., "Video Tracking and Auxiliary Obstacle Avoidance of Power Plant Inspection Robot based on Deep Learning and 3D Digital Twin System," 2024 6th International Conference on Robotics, Intelligent Control and Artificial Intelligence (RICAI), Nanjing, China, 2024, pp. 100-104

[5] M.-S. Baek, E. Jung, Y. S. Park and Y. -T. Lee, "Federated Digital Twin Implementation Methodology to Build a Large-Scale Digital Twin System," 2024 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), Toronto, ON, Canada, 2024, pp. 1-2

[6] Y. Wang, C. Liu, Y. Wang, S. Wu, C. Wang and L. Chen, "Crowd OS Powered Active 3D Mapping Leveraging Heterogeneous Crowd Robots," 2023 International Conference on Artificial Intelligence of Things and Systems (AIoTSys), Xi'an, China, 2023, pp. 51-54

[7] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," arXiv preprint arXiv:1804.02767, 2018. [Online]. Available: https://arxiv.org/abs/1804.02767

[8] M. Shah, S. Goyal, and S. Salvi, "Crowd Density Estimation and Real-time Monitoring using Digital Twins in Transportation Hubs," 2024 IEEE Pune Section International Conference (PuneCon), Pune, India, 2024, pp. 1-6
[9] R. Hao, W. Lu, W. Fan, S. Zhang, X. Li and K. Wang, "A Preliminary Study on Railway Digital Twin Model and System," 2023 IEEE 14th International Symposium on Parallel Architectures, Algorithms and Programming (PAAP), Beijing, China, 2023, pp. 1-5
[10] R. Casadei, G. Delnevo, R. Girau and S. Mirri, "Modelling Groups of Humans: Towards Crowd Digital Twins," 2024 IEEE Symposium on Computers and Communications (ISCC), Paris, France, 2024, pp. 1-6
[11] W. Sultani, C. Chen, and M. Shah, "Real-World Anomaly Detection in Surveillance Videos," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 2018, pp. 6479-6488.
[12] M. Rezaei and M. Azarmi, "DeepSOCIAL: Social Distancing Monitoring and Infection Risk Assessment in COVID-19 Pandemic," Applied Sciences, 68vol. 10, no. 21, p. 7514, 2020.
[13] B. Zhan, D. N. Monekosso, P. Remagnino, S. A. Velastin, and L. Q. Xu, "Crowd analysis: a survey," Machine Vision and Applications, vol. 19, no. 5-6, pp. 345–357, 2008.
[14] Reddy, C. V., Reddy, N. U., Sharma, S., Kumar, J. P., Lakhanpal, S., and Albawi, A. "Advanced Deep Learning Architectures for Real-Time Human Emotion Recognition and Behavioral Prediction". In 2024 IEEE International Conference on Communication, Computer Sciences and Engineering (IC3SE), pp. 1306-1311, May 2024. https://doi.org/10.1109/IC3SE62002.2024.10592940