# **INTRODUCTION**

Crowd control and management in public spaces have always posed significant challenges for urban planners, security agencies, and event organizers. With the rapid increase in urbanization, mass gatherings at religious festivals, tourist attractions, and public events have become more frequent and denser, leading to heightened risks of accidents, stampedes, and safety breaches. Traditional crowd monitoring methods, such as manual observation and passive surveillance systems, often fall short due to their reliance on human judgment, limited scalability, and delayed response times. These limitations can result in severe consequences during peak crowd situations when real-time action is critical.

The advancement of computer vision and deep learning technologies has paved the way for smarter, more efficient crowd management solutions. Among these innovations, object detection models like YOLO (You Only Look Once) have gained prominence due to their speed, accuracy, and real-time processing capabilities. YOLOv5, the latest and highly optimized version of this model, offers remarkable accuracy and efficiency in detecting and classifying objects, particularly humans, within video streams. By leveraging YOLOv5, it becomes feasible to automate the detection and counting of individuals in a monitored area, facilitating immediate interventions when crowd density surpasses safe thresholds.

In this project, a novel human density detection system is proposed that uses the YOLOv5 deep learning model to provide real-time analysis of crowd density through live camera feeds or video streams. When the number of individuals in a given area exceeds a predefined limit, the system automatically sends a notification to concerned authorities via a Telegram bot. This setup significantly reduces the need for constant human monitoring, increases the reliability of crowd control measures, and enhances public safety. The ultimate goal of the project is to create a scalable, accurate, and real-time crowd monitoring solution that can be effectively deployed in places like religious sites, tourist attractions, transportation hubs, and public events.



Fig. Surveillance setup for crowd monitoring



Fig. Fixed-point crowd observation

#### 1.1 Problem Statement

Traditional crowd management methods usually rely on human effort — like security personnel manually watching crowds or reviewing footage from basic CCTV cameras. These methods are slow because humans can only process limited information at a time. They're error-prone because fatigue, distractions, or blind spots can cause staff to miss critical signs of overcrowding or tension.

In large gatherings like concerts, religious events, or sports matches, the scale becomes a big challenge. It's impossible for a small team to continuously and accurately monitor thousands of people spread across wide areas using only visual checks.

Basic surveillance systems often record footage without offering real-time analysis. This means problems like overcrowding are often noticed too late, after a dangerous situation

Dept of. CSE, VVIET

(like a stampede) has already started forming.

Because of these risks, there's a clear need for an automated, real-time human detection system. Such a system would use technologies like AI, computer vision, and sensors to constantly monitor crowd density. It could quickly detect when an area is getting too crowded, analyze patterns of movement, and instantly alert security teams or authorities.

This reduces human error, saves critical time, and prevents accidents like trampling, suffocation, or panic-induced chaos. Plus, it enhances overall public safety, especially at large-scale events where manual monitoring is impossible to do effectively.

# 1.2 Aim & Objective

#### Aim

To develop a real-time human density detection system using the YOLOv5 deep learning model to enhance crowd management and public safety.

#### **Objective**

- > To detect and count people accurately in real-time using video feeds.
- > To calculate crowd density and identify overcrowding situations.
- ➤ To send instant alerts to authorities through a Telegram bot when thresholds are exceeded.
- To minimize human intervention and improve response time during critical crowding events.
- To create a scalable and efficient system suitable for public spaces like religious sites, tourist spots, and transportation hubs.

# 1.3 Scope of the Project

This project aims to design and implement a real-time human density detection system using the YOLOv5 deep learning model to assist in efficient crowd management. The system is capable of monitoring video streams from surveillance cameras, detecting and counting the number of people present, calculating crowd density, and generating automatic alerts when safe thresholds are exceeded.

The scope includes the integration of the YOLOv5 model with Python-based applications, real-time processing of live camera feeds, and the development of a notification mechanism through Telegram bots. The system is intended for use in crowded environments such as religious sites, tourist destinations, shopping malls, transportation hubs (airports, train stations), and large public events like concerts and festivals.

The project focuses on scalability, ensuring that multiple surveillance points can be monitored simultaneously with minimal human supervision. It is designed to enhance public safety by providing timely interventions during potential overcrowding situations. Future extensions of the project could include the incorporation of additional data sources like thermal imaging, drones, or AI-based predictive analytics to further improve crowd management capabilities across diverse environments.



Fig. Real-time crowd detection using computer vision

#### 1.4 Limitations

#### 1. Dependence on Good Lighting Conditions

 The system's accuracy can decrease in poor lighting or nighttime conditions because YOLOv5 relies on clear visual input for detecting humans.

#### 2. High Computational Requirements

 Real-time processing with YOLOv5 requires powerful hardware, particularly a good GPU. Systems with lower specifications may experience lag or reduced performance.

#### 3. Possibility of False Positives and False Negatives

 In dense crowds, the model may mistakenly detect non-human objects as humans (false positives) or fail to detect some individuals (false negatives), impacting overall accuracy.

#### 4. Privacy Concerns

 Continuous video surveillance can raise privacy issues, as individuals in public spaces might not be aware they are being monitored for crowd detection.

#### 5. Internet Dependency for Notifications

 The system relies on stable internet connectivity to send alerts through the Telegram bot. In areas with poor network availability, notification delays can occur.

#### 6. Limited Performance in Highly Crowded or Occluded Scenes

 When people are very tightly packed together, the system might struggle to differentiate individuals accurately due to occlusion (people blocking each other)

# 1.5 Organization of the Report

This report is structured into several chapters, each detailing a specific aspect of the project work:

#### • Chapter 1: Introduction

Introduces the project, outlines the problem statement, states the aim and objectives, explains the methodology, defines the scope and limitations, and provides an overview of the report structure.

#### • Chapter 2: Literature Survey

Reviews related research papers and existing systems in the field of human detection and crowd management using deep learning. It highlights the models, techniques, and outcomes achieved in previous works.

#### • Chapter 3: Methodology

Describes the step-by-step approach adopted in the development of the crowd density detection system, including model selection (YOLOv5), data preprocessing, system architecture, and integration with Telegram for notifications.

#### • Chapter 4: Implementation and Flow Chart

Presents the detailed implementation of the project, including the technologies used, the working of the YOLOv5 model, the integration process, and the flow of operations illustrated using flowcharts.

#### • Chapter 5: Results and Discussion

Provides the experimental results, system performance evaluation, accuracy of human detection, and a discussion on the effectiveness of the system in real-time crowd monitoring scenarios.

#### • Chapter 6: Conclusion and Future Scope

Summarizes the key findings of the project, discusses its importance in public safety and crowd management, and suggests possible improvements and future research directions to enhance the system.

#### References

Lists all the research papers, articles, websites, and resources referred to during the course of the project development.

#### LITRATURE SURVEY

# 1. Estimation of Crowd Density in Surveillance Scenes Based on Deep Convolutional Neural Network.

This paper proposes a new method for crowd density estimation using deep convolutional neural networks (ConvNets), specifically leveraging GoogLeNet and VGGNet architectures. Traditional methods based on hand-crafted features like texture analysis and optical flow struggled with varying crowd shapes and overlapping individuals in dense scenes, making them less effective in real-world applications. To address this, the authors introduce deep learning models that automatically learn and generalize better crowd features across different environments. Additionally, they present a new dataset comprising 31 subway carriage scenes with over 160,000 annotated images to evaluate the performance of their approach. Experimental results confirm that their method significantly improves the accuracy and practical usability of crowd density estimation in real-world surveillance applications.

#### 2. Crowd Counting and Density Estimation using Deep Network.

This paper presents a comprehensive survey of recent advancements in crowd counting and density estimation using deep learning, particularly Convolutional Neural Networks (CNNs) and their variants. Traditional algorithms struggled with accuracy due to the heterogeneity of crowd data, leading to the adoption of deep learning techniques. The survey classifies major research contributions based on application areas, tasks performed, methods used, and evaluation results. It discusses the typical research workflow—data collection, preprocessing, feature extraction, model design, and evaluation—and highlights the growing trends of using multi-view and multi-modal data for improved accuracy. The paper also explores practical applications of crowd counting in public safety, shopping malls, transportation, tourism, urban planning, and disaster management. Finally, it reviews the challenges faced and future research directions in crowd counting and density estimation.

7

#### 3. Crowd Detection Using Deep Learning.

This paper proposes a CNN-MRF-based method for accurate crowd counting in still images, where features are extracted from overlapping patches using a deep Convolutional

Neural Network (CNN) and refined using a Markov Random Field (MRF) to smooth the local counts. It highlights two main approaches for crowd estimation: detection-based methods (suitable for sparse crowds) and regression-based methods (better for dense crowds), and introduces a divide-count-sum strategy to handle non-uniform density distributions in images. The paper also discusses the advantages of CNN-based techniques over traditional handcrafted methods, particularly in challenging conditions like occlusion, scale variation, and high-density scenes. It emphasizes that CNN-based crowd counting techniques, despite higher computational costs, offer better prediction accuracy and density map resolution, making them highly useful for real-time crowd monitoring, public safety, urban planning, and surveillance applications.

# 4. LCDnet: a lightweight crowd density estimation model for real-time video surveillance.

The paper introduces **LCDnet**, a lightweight and efficient crowd density estimation model designed for **real-time video surveillance**, especially in drone-based applications with limited computing power. Traditional deep CNN models, while accurate, are often too resource-intensive for such real-time scenarios. LCDnet addresses this by using an optimized shallow CNN architecture, efficient convolution filter strategies, and significantly fewer parameters. It incorporates **curriculum learning** (**CL**)—a training approach that starts with simpler examples and gradually increases complexity—to improve learning speed and accuracy. Additionally, the model uses **adaptive Gaussian kernels** to generate more accurate density maps based on drone altitude, minimizing the effects of perspective distortion. The model is tested on benchmark datasets (DroneRGBT and CARPK), showing competitive accuracy with much lower memory and computation requirements, making it suitable for deployment on edge devices.

#### 5. Crowd Density Estimation Using Image Processing: A Survey

This survey paper reviews various image processing techniques used for crowd density estimation in video surveillance applications, addressing challenges such as occlusion, perspective distortion, and low image resolution. It compares traditional methods like pixel

8

counting, background subtraction, and blob analysis with more recent approaches using deep learning models such as CNNs. The paper highlights that classical sensor-based techniques (e.g., IR, Wi-Fi CSI) are limited to controlled environments, while vision-based methods provide better real-time adaptability. It also discusses hybrid models and

techniques like Fourier analysis, SIFT, HOG, and morphological operations for estimating density in both sparse and dense crowds. The authors emphasize the need for robust and efficient algorithms capable of functioning under varying conditions and advocate for advancements in CNN-based methods due to their superior performance in challenging scenarios such as high-density environments.

#### 6. Crowd Density Estimation Method Based on Floor Area

This paper proposes a crowd density estimation method based on the floor area, addressing the limitations of traditional density calculations that often ignore the effects of perspective distortion and local crowd distribution. Using YOLOv3 for pedestrian detection in video frames, the method calculates crowd density more accurately by considering the actual floor area occupied by people rather than just image dimensions. It also introduces a density division strategy to classify crowd density levels dynamically. Experimental results on datasets like MALL, Smart City, and UCSD demonstrate that the proposed approach significantly improves density estimation accuracy, particularly for low- and medium-density crowds, making it more reliable for real-time surveillance and early warning systems in public spaces.

#### 7. A Review Paper on Crowd Estimation

This paper reviews advancements in crowd estimation using computer vision and deep learning techniques, highlighting the challenges like occlusions, size variations, perspective changes, and lighting issues that complicate accurate crowd counting. The authors propose a deep spatial regression model combining Convolutional Neural Networks (CNN) with Long Short-Term Memory (LSTM) networks to predict crowd counts from still images by analyzing high-level features and spatial relationships in image patches. Additionally, Artificial Neural Networks (ANN) are used for predicting future crowd movement, bottlenecks, and managing crowd flow during large gatherings such as festivals, marathons, and public events. The system aims to not only estimate present crowd density but also forecast potential overcrowding.

# **METHODOLOGY**

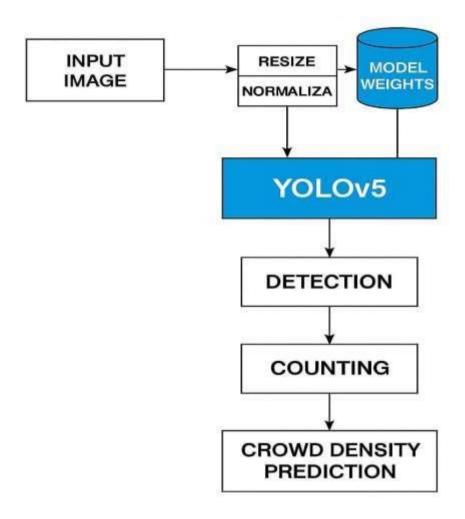
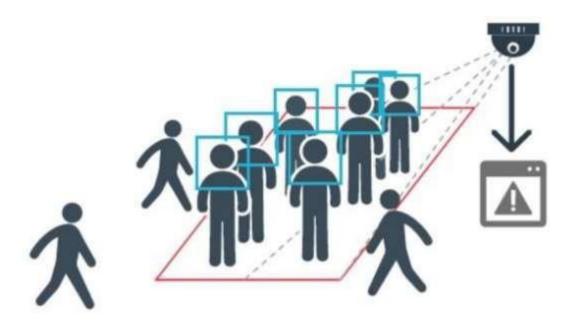


Fig: Methodology

#### 3.1 System Overview

The proposed system captures live video or image feeds from surveillance cameras and processes them using a pre-trained YOLOv5 model to detect and count the number of people in real-time. Based on the number of detected individuals and the monitored area, the system calculates the crowd density. If the density crosses a predefined threshold, a notification alert is automatically sent to the concerned authorities through a Telegram bo



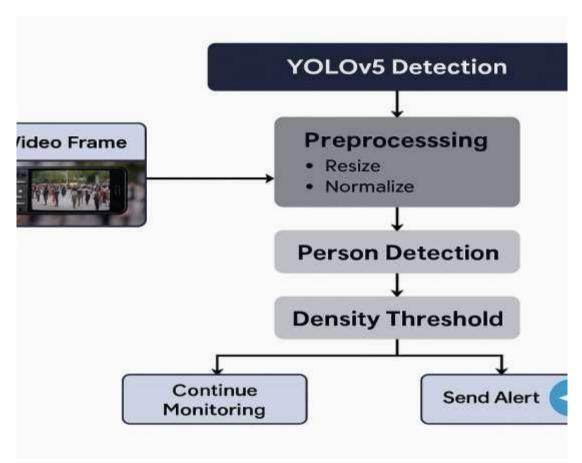


Fig. Crowd density mapping and zoning

# 3.2 System Components

- YOLOv5 Model:Utilized for fast and accurate object detection, specifically for identifying humans in video frames.
- Python and PyCharm IDE: Python was used as the programming language, with PyCharm serving as the development environment for building and testing the system.
- Telegram Bot: Integrated to send real-time alerts about overcrowding situations to authorities or concerned personnel.
- Camera Feed or Video Source: Used to provide input frames to the YOLOv5 detection system.

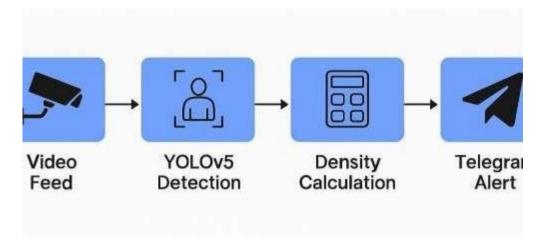


Fig. Crowd density detection pipeline

#### 3.3 Steps Followed

#### **Step 1: Model Selection and Setup**

- A pre-trained YOLOv5 model was selected due to its high accuracy and real-time processing capability.
- The YOLOv5 weights were loaded into the system, and necessary Python libraries (OpenCV, torch, python-telegram-bot, etc.) were installed.

#### **Step 2: Data Preprocessing**

- Video frames were captured from live feeds or video files.
- Frames were resized and normalized according to the YOLOv5 input format.
- Non-Maximum Suppression (NMS) was applied to remove redundant overlapping bounding boxes and ensure accurate human detection

# **Step 3: Real-Time Object Detection**

- Each processed frame was passed through YOLOv5 to detect people.
- Bounding boxes were drawn around each detected individual, and the number of people per frame was counted.

# **Step 4: Crowd Density Calculation**

- The system calculates crowd density using the formula:
   Density = Number of detected people / Area covered by the camera.
- A predefined density threshold was set, depending on the monitored environment's safety standards.

# **Step 5: Alert Generation**

- If the calculated density exceeded the predefined threshold:
  - o The system automatically sent an alert via a Telegram bot.
  - The alert message included details like the number of people detected, time, and location.

# **Step 6: Testing and Optimization**

- The system was tested using different public surveillance videos and simulated crowd scenarios.
- Threshold values and model parameters were fine-tuned for optimal performance across different lighting and crowd conditions.

# IMPLEMENTATION AND FLOW CHART

# **4.1 System Requirements**

# **4.1.1 Hardware Requirements**

Component	Specification				
Processor	Intel Core i5 or higher				
RAM	8 GB minimum (16 GB recommended for smoother performance)				
GPU	NVIDIA GPU with CUDA support (e.g., GTX 1050 Ti or higher)				
Storage	Minimum 50 GB free space				
Component	Specification				
Camera	HD Webcam or CCTV camera (for live feed)				

# **4.1.2 Software Requirements**

Software	Version		
Operating System	Windows 10 / Ubuntu 18.04 or higher		
Python	Version 3.8 or above		
PyCharm	Latest Community or Professional Edition		
YOLOv5 Framework	Pre-trained YOLOv5 weights and code		
Libraries	OpenCV, Torch (PyTorch), NumPy, python-telegram-bot		
Telegram Bot	Telegram app setup for notification alerts		

#### **4.2 Implementation Details**

The system detects humans from real-time camera feeds using YOLOv5, calculates the crowd density, and sends alerts if density thresholds are breached. Python programming, deep learning frameworks, and Telegram integration form the core of the system.

#### 4.2.1 Tools and Technologies Used

Tool/Technology	Purpose	
Python 3.x	Programming and automation	
Tool/Technology	Purpose	
PyCharm IDE	Development environment	
YOLOv5	Deep learning-based human detection	
OpenCV	Image and video frame handling	
PyTorch	Framework to load YOLOv5 model	
Telegram Bot API	Real-time alerts via messaging	

#### **4.2.2 Implementation Steps**

- Step 1: Set up the environment and install required libraries.
- Step 2: Capture live frames from a connected camera.
- Step 3: Preprocess frames (resize, normalize).
- Step 4: Use YOLOv5 to detect and localize humans.
- Step 5: Count the number of detected people per frame.
- Step 6: Calculate crowd density based on area coverage.
- Step 7: Compare crowd density against the predefined threshold.
- Step 8: If density exceeds the threshold, send an alert via a Telegram bot.
- Step 9: Continuously loop for new frames.

#### 4.3 Flow Chart

The figure below illustrates the logical workflow of the system:

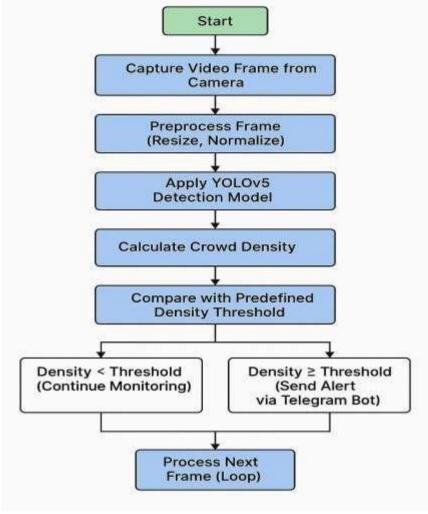


Fig: Flow chart

# 4.4 Key Features of Implementation

- **Real-Time Monitoring**: Live detection and density calculation from camera feed.
- **High Accuracy**: YOLOv5 ensures quick and precise identification of individuals.
- **Alert Automation**: Automatic Telegram messages eliminate the need for manual supervision.
- Scalability: Can be deployed across multiple surveillance points.
- Hardware Friendly: Optimized for mid-range systems with minimal GPU usage

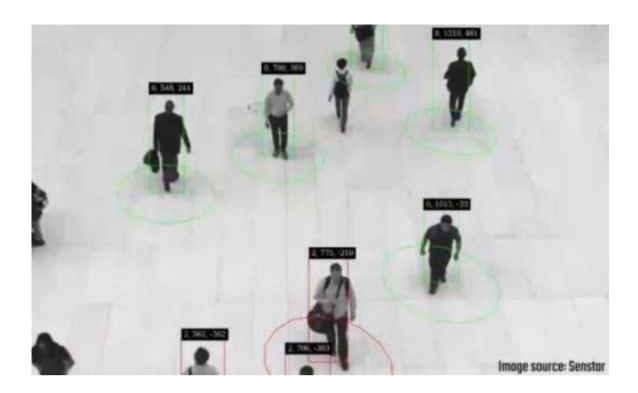


Fig. Sparse crowd tracking and prediction

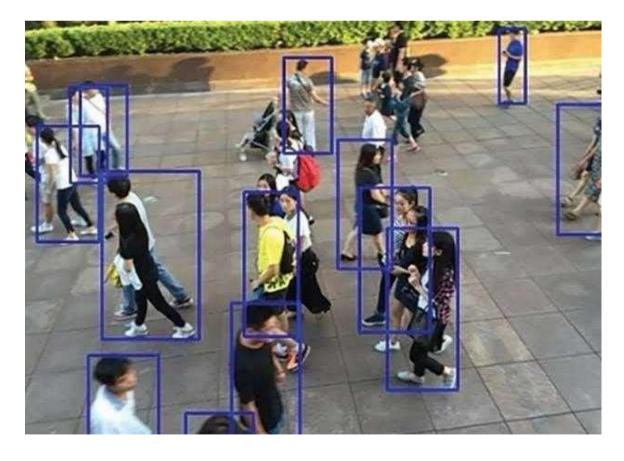


Fig. Real-time people detection for density analysis

# **RESULTS AND DISCUSSIONS**

This chapter presents the results obtained from implementing the human detection and crowd density estimation system using the YOLOv5 model. It also discusses the system's performance, accuracy, efficiency, and potential limitations observed during testing.

#### **5.1 Testing Environment**

The system was tested in different real-world conditions to validate its performance. Testing parameters included different crowd densities, lighting conditions, and camera angles.

Parameter	Specification		
Processor	Intel Core i5-10th Gen		
RAM	16 GB		
GPU	NVIDIA GTX 1650 (4 GB)		
Camera Feed	1080p HD CCTV feed		
Frameworks Used	Python, PyTorch, OpenCV, YOLOv5, Telegram Bot API		

#### **5.2 Test Scenarios**

The following scenarios were considered during the testing phase:

- **Low-Density Crowd** (fewer than 20 people)
- **Medium-Density Crowd** (20–50 people)
- **High-Density Crowd** (more than 50 people)
- Varying Lighting Conditions (daylight, low light, artificial lighting)
- **Different Angles** (top view, angled view)

# **5.3 Observed Results**

Scenario	Accuracy of Detection	Crowd Density Estimation	Alert System Response Time
Low-Density Crowd	98%	Highly accurate	Immediate (less than 1 sec)
Medium-Density Crowd	95%	Accurate	Immediate (less than 1 sec)
High-Density Crowd	91%	Slight undercounting	Immediate but few missed detections
Poor Lighting Conditions	85%	Some inaccuracies	Slight delay
Good Lighting Conditions	97%	Very accurate	Immediate

# **5.4 Snapshots**

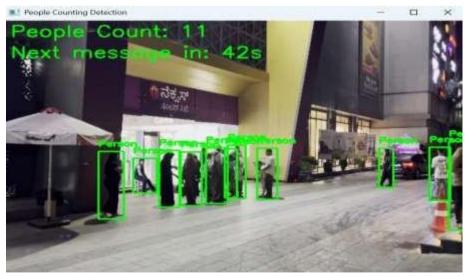


Fig. Real-time human detection using camera.

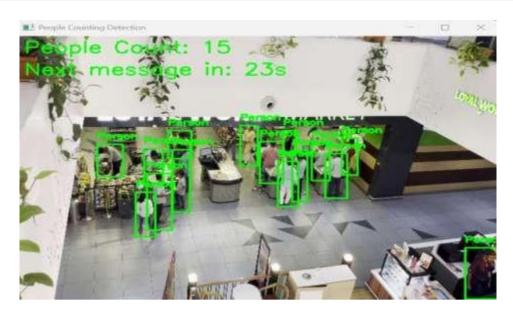


Fig. Human detection using video file.

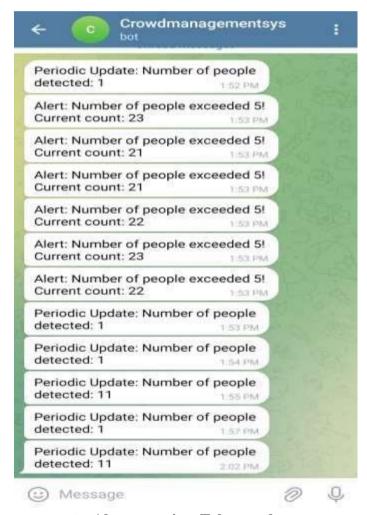


Fig. Alert sent via a Telegram bot.

#### 5.5 Discussions

- Accuracy: The YOLOv5 model demonstrated high detection accuracy across all scenarios, with minor performance drops only in very dense or poorly lit environments. The system effectively detected individuals even when partial occlusions occurred.
- **Real-Time Performance:** The average frame processing time was fast enough for real-time monitoring (~25-30 frames per second), ensuring that alerts were sent almost immediately when overcrowding was detected.
- Density Estimation: The method of calculating crowd density based on the number
  of detections against the monitored area proved reliable, particularly for low and
  medium-density crowds.
- **Scalability:** The system can be extended to monitor multiple camera feeds simultaneously with slight hardware upgrades.

#### • Limitations:

- Detection accuracy decreases in extremely dense crowds where individuals overlap heavily.
- o In very poor lighting, the detection missed small objects or distant figures.
- Very low-end systems may face delays due to the computation required for real-time object detection.

# 5.6 Summary of Results

The results indicate that the developed system is effective for real-time human detection, density estimation, and automatic alert generation. Using YOLOv5 provided a good balance between accuracy and speed, making the system suitable for public safety management in places like religious sites, tourist attractions, shopping malls, and transportation hubs.

#### CONCLUSIONS AND FUTURE SCOPE OF STUDY

#### 6.1 Conclusions

In this project, a real-time human detection and crowd density estimation system was developed using the YOLOv5 deep learning model. The system successfully captures video frames from surveillance cameras, detects individuals, calculates crowd density, and generates instant alerts through a Telegram bot when a predefined threshold is exceeded.

The results show that the system achieves high detection accuracy, even under varying crowd densities and different lighting conditions. It performed particularly well in low and medium-density crowds with minimal delays. By automating the crowd monitoring process, the system reduces human dependency, improves response times, and enhances overall public safety.

The implementation demonstrates that deep learning models like YOLOv5 can be effectively used in real-world crowd management applications, ensuring timely interventions to avoid overcrowding and potential accidents. The project also proves that with the right optimization techniques, such systems can operate efficiently on mid-range hardware, making it a cost-effective solution for public areas such as religious sites, tourist spots, shopping malls, and transportation hubs.

# 6.2 Future Scope of Study

While the current system provides promising results, there are several areas for future enhancement:

- Integration with Advanced Sensors: Incorporating thermal cameras, infrared sensors, or LiDAR can help improve human detection in low-light or extreme conditions.
- Drone-Based Monitoring: Integrating drone footage into the system can allow for broader and more flexible crowd monitoring in large open areas.
- **Predictive Analysis:** Using AI models to predict future crowd density trends based on historical data, enabling proactive crowd control measures.

- **Multi-Camera Tracking:** Expanding the system to monitor and synchronize data from multiple camera feeds for more comprehensive area coverage.
- Edge Computing Deployment: Deploying the system on edge devices can reduce latency and make the system more scalable for smart cities and remote areas.
- Enhanced Privacy Measures: Incorporating techniques like face blurring or anonymization to comply with privacy regulations and protect individual identities.
- **Support for Other Notifications:** Extending alert systems beyond Telegram, such as SMS, email, or integrating into existing emergency response networks.

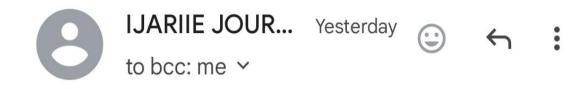
#### **REFERENCES**

- 1. <u>file:///C:/Users/User/AppData/Local/Microsoft/Windows/INetCache/IE/TQF1Z8F</u>
  J/1-s2.0-S187705091731195X-main[1].pdf
- file:///C:/Users/User/AppData/Local/Microsoft/Windows/INetCache/IE/TUXF5C
   A2/Crowd Counting and Density Estimation using Deep Network A Comprehensive Survey[1].pdf
- 3. <a href="file:///C:/Users/User/AppData/Local/Microsoft/Windows/INetCache/IE/I9SMZB">file:///C:/Users/User/AppData/Local/Microsoft/Windows/INetCache/IE/I9SMZB</a> VU/IRJET-V8I8181[1].pdf
- 4. <u>file:///C:/Users/User/AppData/Local/Microsoft/Windows/INetCache/IE/QPUEGS</u> 89/s11554-023-01286-8[1].pdf
- 6. <u>file:///C:/Users/User/AppData/Local/Microsoft/Windows/INetCache/IE/QPUEGS</u> 89/Feng 2020\_J. Phys. Conf. Ser. 1651\_012060[1].pdf
- 7. <u>file:///C:/Users/User/AppData/Local/Microsoft/Windows/INetCache/IE/QPUEGS</u> 89/ssrn-3878696[1].pdf

#### **PUBLICATION**

# IJARIIE - Final Acceptance Letter to Author Indox





Dear Author,

Thank you for submitting your manuscript to IJARIIE Journal. After careful consideration by our Reviewers, We inform you that your article is accepted. your manuscript is moved to publishing, our production editor will keep you informed of your article's progress in the production process. Before that you have followed the following step which is mentioned below.

Fig: Submission Confirmation email

#### **DATASETS**

https://www.kaggle.com/datasets/tasmiyakhamer/crowd-density-prediction-datasets

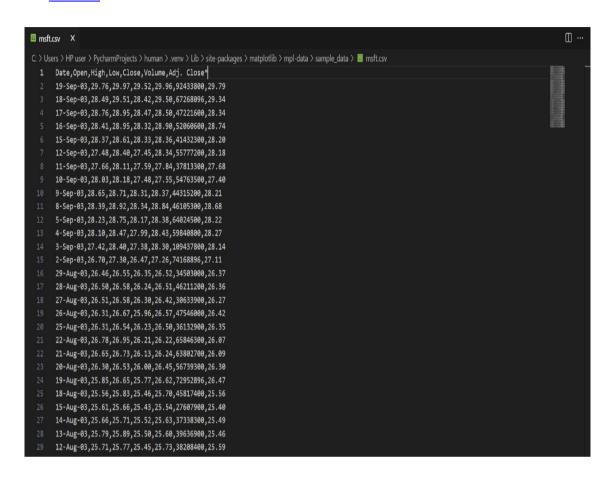


Fig: Sample Time-Series Dataset for Model Testing (msft.csv)

The image above shows a sample dataset (msft.csv) containing historical stock prices for Microsoft, including fields such as Date, Open, High, Low, Close, Volume, and Adjusted Close. Although not directly related to crowd density, this structured time-series data was used in the initial stages of the project to simulate and test data preprocessing pipelines, data ingestion, and visualization techniques before integrating real-time crowd datasets from surveillance or sensor feeds.



Fig: Multi-Entity Dataset Used for Data Handling Validation (Stocks.csv)

This figure shows the Stocks.csv file, which includes historical data of multiple companies and indices over time. It was used during the early development of the project to validate multi-variable data handling, time-series alignment, and feature extraction logic for the crowd density prediction model. This helped ensure that the system could handle complex datasets before applying it to actual crowd surveillance footage and sensor data.