# Crowd Behavior Analysis for Enhanced Event Safety and Management

**A report submitted to**

**RAMAIAH INSTITUTE OF TECHNOLOGY**

**Bengaluru**

**SENIOR PROJECT (ISP)**

as partial fulfillment of the requirement for the award of degree of

**Bachelor of Engineering (B.E) in Information Science and Engineering**

**By**


| | |
|---|---|
| Jatin B | 1MS20IS054 |
| Manoj S | 1MS20IS070 |
| Pannaga N | 1MS20IS083 |
| Sanjeev G | 1MS21IS406 |

Under the Guidance of

Ms.D.Evangeline

Assistant Professor

DEPARTMENT OF INFORMATION SCIENCE AND ENGINEERING

RAMAIAH INSTITUTE OF TECHNOLOGY

(Autonomous Institute, Affiliated to VTU)

BANGALORE - 54


MAY 2024

**Department of Information Science and Engineering**

**Ramaiah Institute of Technology**

**Bengaluru - 54**

# CERTIFICATE

Certified that the project entitled "Crowd Behavior Analysis for Enhanced Event Safety and Management" carried out by Jatin B (USN - 1MS20IS054), Manoj S (USN - 1MS20IS070), Pannaga N (USN - 1MS20IS083), and Sanjeev G (USN - 1MS21IS406), bonafide students of M S RAMAIAH INSTITUTE OF TECHNOLOGY, Bangalore is in partial fulfillment for the award of Degree of Bachelor of Engineering in Information Science and Engineering of the Visvesvaraya Technological University, Belagavi during the year 2023-2024. The project work has been approved as it satisfies the academic requirements in respect of project work prescribed for the said degree.

Name and Signature of the Guide Signature of the HOD Signature of the Principal

**External Viva**

Name of the examiners Signature with date

1.

2.

# DECLARATION

We hereby declare that the entire work embodied in this ISP SENIOR PROJECT report has been carried out by us at Ramaiah Institute of Technology under the supervision of Ms.D.Evangeline. This project report has not been submitted in part or full for the award of any diploma or degree of this or any other University.

| | |
|---|---|
| Jatin B | 1MS20IS054 |
| Manoj S | 1MS20IS070 |
| Pannaga N | 1MS20IS083 |
| Sanjeev G | 1MS21IS406 |

Place: Bangalore
Date:

# ABSTRACT

Addressing crowd control and safety at large-scale events, such as music festivals and sporting occasions, is the central focus of this study. Employing state-of-the-art technologies, the research utilizes a Convolutional Neural Network (CNN) model equipped with VGG, dilatable layers and Atrous Spatial Pyramid Pooling (ASPP) layers to identify individual heads amidst the crowd. Furthermore, optical flow analysis identifies fast-moving pixels, facilitating the detection of rapid movements within the crowd. YOLO tracking is additionally employed to monitor the direction of object movement within the crowd. By integrating these methodologies with tactical crowd management approaches, the study aims to enhance overall safety and security, thereby reducing risks and enhancing the experience for both event staff and attendees. The ASPP approach demonstrates approximately 15% higher accuracy on average compared to the baseline model for the ShanghaiTechA and UCF CC 50 datasets.

# ACKNOWLEDGMENT

We take this opportunity to express our gratitude to the people who have been instrumental in the successful completion of this project. We would like to express our profound gratitude to the Management and Dr. N.V.R Naidu, Principal R.I.T, Bengaluru for allowing us to explore our potential.

We sincerely thank our beloved Dr. Vijaya Kumar B P, HOD, Information Science and Engineering, for his constant support and guidance.

We wholeheartedly thank our project guide Prof. Ms. D. Evangeline, for providing us with the confidence and strength to overcome every obstacle at each step of the project and inspiring us to the best of our potential. We also thank her for her constant guidance, direction, and insight during the project.

Finally, we would like to express sincere gratitude to all the teaching and non-teaching faculty of ISE Department, our beloved parents, seniors, and my dear friends for their constant support during the course of work.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# INTRODUCTION

Crowd management and safety at large events are paramount considerations to ensure the well-being and orderliness of attendees. One crucial aspect of crowd management is understanding the density of crowds through segmentation and binarization techniques. Integration of advanced techniques, including a CNN model comprising VGG, dilatable layers and ASPP layers, event organizers can gain insights into crowd distribution, movement patterns, and potential overcrowding, enabling proactive measures to maintain safety and security.

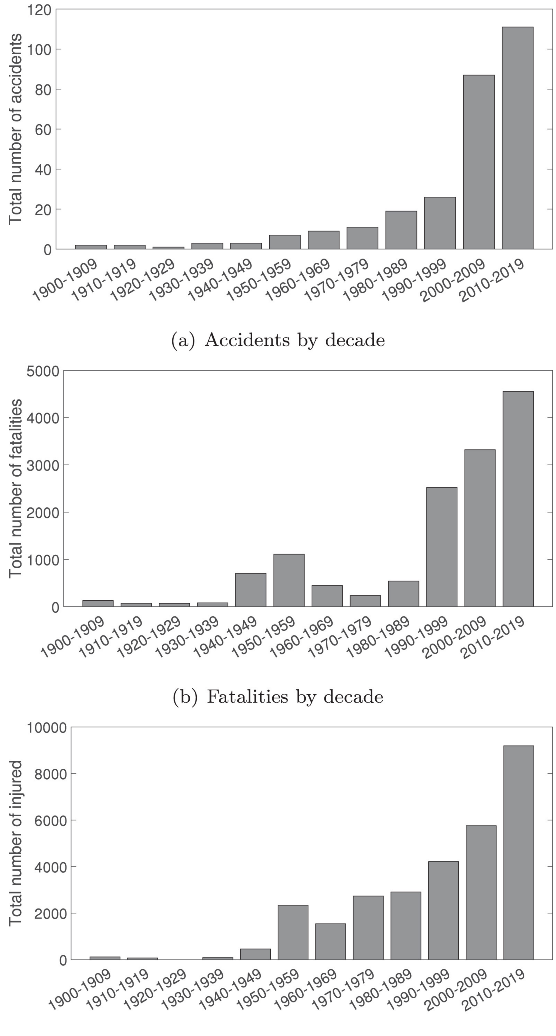
Anomaly detection, facilitated by methods like Particle Advection is pivotal in optimizing crowd management. This interactive approach enables swift identification of unusual behaviors or events within the crowd, empowering event organizers to promptly respond to potential risks.

Efficient crowd management and safety rely on detecting fast-moving individuals and monitoring directional movement within the crowd. Leveraging optical flow analysis enables the identification of fast-moving pixels, thereby facilitating the detection of fast-moving objects. Additionally, YOLO tracking aids in tracking directional movement of objects within the crowd, contributing to effective crowd regulation and accident prevention. Emphasizing these objectives in crowd management strategies significantly enhances the success and safety of large-scale events.

## 1.1 Motivation and Scope

The motivation for undertaking this major project stems from the critical need for effective crowd monitoring and management in light of the rising population and the increasing frequency of large-scale events worldwide. Tragic incidents such as the Indonesia football stampede in October 2022, where 125 individuals lost their lives and the Kochi university concert stampede in November 2023, resulting in four student deaths and numerous injuries, serve as stark reminders of the potential dangers associated with overcrowded events.

There is exponential increase in the amount of accidents, fatality and injuries in each decade [1] as shown in Fig 1.1.




(a) Accidents by decade

(b) Fatalities by decade

(c) People injured by decade

Figure 1.1: Decadal accidents, fatalities and injuries

These incidents underscore the urgent necessity for advanced crowd management strategies that can accurately assess crowd density, identify potential risks such as overcrowding and stampedes, and enable swift response mechanisms to prevent casualties and ensure the safety of event attendees. The tragic stampede at the Barea stadium in Madagascar in August 2023, which claimed 12 lives and left 80 injured during the Indian Ocean Island Games' opening ceremony, further emphasizes the critical nature of implementing robust crowd monitoring and safety measures.

The scope of this major project encompasses the development and integration of advanced technologies, including CNN models with VGG, dilatable layers, and ASPP layers, for crowd density analysis and anomaly detection. Additionally, the project will focus on utilizing methods such as Particle Advection for swift identification of unusual behaviors within crowds and optical flow analysis combined with YOLO tracking for monitoring fast-moving individuals and directional movement within the crowd.

The ultimate goal of this project is to contribute to the enhancement of crowd management strategies, thereby mitigating potential risks, ensuring the safety of attendees at large events, and preventing tragic incidents like those mentioned above from occurring in the future. By addressing these pressing concerns through innovative technological solutions, this project aims to make a meaningful impact on event safety and security on a global scale.

Present-day challenges call for solutions that require the utmost precision and accuracy in the developed solutions as shown in Fig 1.1.

This however has led to many research challenges published by various journals.

## 1.2 Issues and Challenges

The issues and challenges surrounding crowd management and safety at large-scale events are multifaceted and complex. Here are some key points to consider:

1. **Crowd Density and Movement**: One of the primary challenges is accurately assessing and managing crowd density. Large gatherings often result in densely packed crowds, which can lead to discomfort, anxiety and in extreme cases, stampedes or crushes. Understanding how crowds move and behave is crucial for effective management.

2. **Risk of Overcrowding**: Overcrowding is a significant risk factor at events, especially during popular performances or activities. When venues reach or exceed their capacity, it can pose serious safety hazards, including difficulties in evacuating people in case of emergencies.

3. **Anomalies and Unpredictable Behavior**: Crowd behavior can be unpredictable, and detecting anomalies such as sudden rushes, fights or medical emergencies can be challenging. Identifying and responding to these anomalies swiftly is essential to prevent escalations and maintain order.

4. **Communication and Coordination**: Communication between event organizers, security personnel and emergency responders is critical for effective crowd management. Delays or breakdowns in communication can hinder response times and exacerbate safety risks.

5. **Technological Limitations**: While advanced technologies like CCTV cameras, drones and AI-based analytics offer valuable insights, there are challenges in integrating and interpreting data in real-time. Technical glitches or system failures can impede the effectiveness of these tools.

6. **Cultural and Social Factors**: Cultural norms, social dynamics and the behavior of specific attendee demographics can influence crowd management strategies. Understanding these factors and tailoring approaches accordingly is vital for successful event management.

7. **Emergency Preparedness and Response**: Being prepared for emergencies such as medical incidents, fires, or terrorist threats is paramount. Effective emergency response plans, including evacuation procedures and medical facilities are essential components of crowd safety.

Addressing these issues requires a holistic approach that combines technological innovations, effective communication protocols, comprehensive risk assessments and rigorous training for event staff and security personnel. Proactive measures, continuous monitoring and adaptive strategies are key to mitigating risks and ensuring the safety and well-being of all attendees at large-scale events.

## 1.3 Problem Statement

The problem of crowd management and safety at large-scale events poses significant challenges due to the potential for unexpected incidents, overcrowding and the resulting risks of accidents and stampedes. Despite advancements in event planning and security measures, ensuring the well-being and orderliness of attendees remains a critical concern for event organizers. Addressing this issue requires innovative approaches and technologies that can accurately monitor crowd density, detect anomalies in behavior and facilitate swift responses to mitigate potential risks and enhance overall event safety.

## 1.4 Objectives

1. **Density of Crowd**: Measuring the concentration of people in a given area.
2. **Anomaly detection**:
   (a) **Identify Fast Moving objects in the Crowd**: Detecting objects that are moving at an unusually high speed compared to the rest of the crowd.
   (b) **Detect People in a Different Direction**: Identifying individuals who are moving in a direction that is significantly different from the majority of the crowd.

# Chapter 2

# LITERATURE REVIEW

The work in [2] delves into the critical domain of crowd safety and the emerging field of crowd science, emphasizing the necessity for well-crafted design strategies, environmental awareness and formalized training in crowd management principles. It highlights the crucial role of crowd risk analysis, illustrating how design elements can impact incidents and emphasizing the vital function of information dissemination and effective management in ensuring crowd safety. This model intricately analyzes design, information and management shortcomings across different stages of crowd movement. The research aims to provide empirical evidence on dynamic risk analysis techniques to enhance both crowd safety and throughput. Notable cases, such as the Sydney Olympics and Murrayfield Stadium projects are examined to demonstrate the practical application of the DIM-ICE model in optimizing crowd management strategies and improving safety measures. Through an interdisciplinary approach.The study emphasizes the urgent need for a comprehensive approach to crowd safety, one that integrates insights from design, psychology, sociology and management. By shedding light on the multifaceted dynamics of crowd behavior and risk mitigation strategies. It plays a pivotal role in advancing our collective understanding of crowd management and safety. This thorough analysis not only highlights the importance of proactive measures but also stresses the need for ongoing refinement of methodologies and tools to adapt to evolving crowd dynamics and emerging challenges.

The work presented in [3] is a thorough exploration of crowd anomaly detection through spatio-temporal texture analysis, introducing a novel model designed specifically for identifying and extracting Spatial-Temporal Textures (STT) from video footage. Central to the discussion is the acknowledgment of the critical role played by feature extraction, modeling techniques and crowd behavior recognition in this field. It discusses the utilization of the gray level co-occurrence matrix for crowd anomaly labeling and meticulously evaluates the experimental results obtained from employing this matrix with various classifiers. Furthermore, it delves into the concept of information entropy as a method for detecting abnormal crowd behaviors, explaining the calculation of information entropy for each STT to identify target STTs for crowd behavior analysis. By emphasizing the importance of real-time monitoring of high-density crowds for public safety. The study highlights the potential of the proposed anomaly detection framework to enhance system adaptability across different scenarios. This comprehensive examination not only contributes to advancing the comprehension of crowd anomaly detection methodologies but also underscores the practical implications for improving surveillance systems and ensuring public safety in dynamic environments. Hence, it serves as a valuable asset for both researchers and practitioners, offering insights that could guide the development of more resilient and efficient crowd monitoring systems in the future.

The paper [4] introduces LCDnet, a lightweight crowd density estimation model tailored for real-time video surveillance applications . It addresses the challenge of achieving precise crowd counting while minimizing complexity, which is particularly relevant for scenarios with limited computing resources such as drones. LCDnet utilizes Curriculum Learning (CL) for training and undergoes evaluation on benchmark datasets, Dron-eRGBT and CARPK showcasing its efficiency compared to existing models. Through careful model architecture design, annotated training data utilization and efficient learning strategies implementation, LCDnet demonstrates commendable accuracy, reduced inference time and minimal memory requirements, making it suitable for deployment on edge devices with constrained resources. The study highlights the importance of optimizing model design, data annotation, and learning methodologies to enhance the performance of lightweight crowd density estimation models in real-time surveillance environments. This research not only advances the capabilities of crowd monitoring systems but also has practical implications for enhancing public safety measures through efficient and effective crowd counting techniques. By addressing the computational constraints inherent in surveillance scenarios, LCDnet sets the stage for improved real-time crowd analysis and management, enabling proactive responses to crowd-related incidents and ensuring the security of public spaces.

Nayan, Sahu, and Kumar's research [5] introduces a novel approach to detecting anomalous crowd behavior through correlation analysis of optical flow. In the realm of visual surveillance systems, accurately and swiftly identifying abnormal events remains a significant challenge. The study underscores the importance of defining anomaly criteria and emphasizes the necessity for real-time detection to effectively mitigate the impact of unwanted circumstances. By utilizing optical flow and correlation analysis, the proposed method demonstrates a remarkable accuracy of 97.32% in swiftly identifying anomalous behavior within crowds. The empirical findings highlight the significance of factors such as frame gap and illumination conditions in optimizing the detection process, with a threshold value of 0.75 proving effective across varying lighting scenarios. This innovative methodology not only surpasses existing techniques in terms of speed and precision but also provides a robust solution for real-time monitoring and analysis of crowd behavior. Through meticulous examination of correlation patterns in optical flow, the study offers valuable insights into the dynamics of crowd behavior, thereby contributing to the development of more effective surveillance systems capable of detecting and addressing anomalies in diverse environments.

Covering various key sections, the paper [6] focuses on object and patch-based anomaly detection and localization in crowded scenes. It meticulously examines the challenges and complexities associated with anomaly detection techniques, shedding light on the nuanced nature of anomalies, diverse application domains and the categorization of anomaly detection techniques based on research areas. The study explores the richness and complexity of the anomaly detection problem, distinguishing between simple and complex anomalies, and clarifies the definitions of contextual and collective anomalies. Furthermore, it systematically categorizes techniques into classification-based, nearest neighbor-based, clustering-based, statistical, information-theoretic and spectral techniques, offering comprehensive insights into their computational complexity for both training and testing phases. Importantly, the paper introduces new categories of anomaly detection techniques, such as information-theoretic and spectral techniques and engages in a thorough discussion of the assumptions, advantages and disadvantages associated with techniques within each category. This thorough exploration not only enhances the understanding of anomaly detection methodologies but also provides a valuable resource for researchers and practitioners navigating the complexities of anomaly detection in crowded environments. Through its comprehensive analysis and classification, the paper establishes a solid foundation for further advancements in the field, offering a roadmap for the development and refinement of effective anomaly detection strategies tailored to diverse real-world applications.

Ramchandran and Sangaiah's research [7] focus on developing an unsupervised deep learning system for local anomaly event detection in crowded scenes. Their study aims to address the challenge of detecting anomalies in crowded environments without the need for manual labeling or supervision. By utilizing deep learning techniques, the researchers introduce a system that can automatically identify unusual events within densely populated areas. This research contributes to the field of multimedia tools and applications by offering a novel approach to anomaly detection, which can enhance security and surveillance systems in crowded settings. They explore existing research on anomaly detection, deep learning and multimedia applications to establish a foundation for their work, covering a wide range of sources, including primary, secondary and tertiary information. Through the analysis and evaluation of various studies, the researchers aim to position their work within the existing body of knowledge and contribute to advancements in anomaly detection technology.

Exploring various critical dimensions, the study [8] investigates the detection of abnormal crowd behavior using motion information images and convolutional neural networks. It meticulously navigates through the challenges and intricacies linked with identifying abnormal crowd behaviors, with a particular emphasis on the integration of motion information images and convolutional neural networks for precise detection. Addressing the significance of leveraging deep learning techniques for anomaly detection in crowded environments, the research underscores the efficacy of amalgamating motion information images with convolutional neural networks. Furthermore, it delves into the methodology of employing convolutional neural networks to analyze motion information images, stressing the importance of this approach in refining the accuracy and efficiency of abnormal crowd behavior detection. Additionally, the paper presents experimental findings showcasing the method's performance and superiority over existing techniques. By harnessing motion information images and convolutional neural networks, the research contributes to advancing the capabilities of anomaly detection in crowded settings, highlighting the potential of deep learning approaches in tackling complex real-world challenges. Through its comprehensive investigation and empirical validation. The study provides valuable insights that could steer the development of more resilient and efficient crowd surveillance systems, thereby bolstering public safety measures across diverse environments.

The work [9] revolutionize a fundamental aspect crucial for applications such as video surveillance, traffic management and urban planning. The accuracy of crowd counting is challenged by factors like occlusion, perspective distortion, complex backgrounds and varying sizes of individuals. Current methods often prioritize total count accuracy over density map precision, leading to errors in individual location pinpointing. To tackle this, the authors propose a novel encoder-decoder CNN architecture that amalgamates features from both network parts to generate a more precise density map and estimate the total number of people with greater accuracy. They introduce a novel evaluation metric termed Patch Absolute Error (PAE) to assess density map precision. Experimental results showcase the superiority of this approach over existing methods on public crowd counting datasets and its robust performance across different scenes, indicating its potential for real-world applications. By addressing the drawbacks of previous approaches and providing a more accurate and comprehensive crowd counting solution, this innovative CNN-based approach carries promising implications for various practical applications requiring precise crowd estimation.

In its comprehensive exploration, the study [10] deeply probes into crowd emotion evaluation based on fuzzy inference of arousal and valence. It meticulously navigates through the challenges and complexities inherent in assessing crowd emotions, particularly emphasizing the application of fuzzy inference for gauging arousal and valence. The research underscores the significance of integrating fuzzy logic into emotion evaluation, highlighting its efficacy in capturing subtle emotional nuances within crowds. Moreover, it elaborates on the methodology of employing fuzzy inference to analyze arousal and valence, stressing its importance in enhancing the precision and depth of crowd emotion assessment. The study also showcases experimental findings demonstrating the performance and superiority of the fuzzy inference model over traditional emotion evaluation methods. By harnessing fuzzy inference for crowd emotion evaluation. The research advances the comprehension and measurement of emotional states within crowds, offering valuable insights applicable across diverse domains, from crowd management to marketing strategies. Through empirical validation and a meticulous examination of fuzzy inference techniques, the study underscores the potential of fuzzy logic in addressing the complexities of crowd emotion assessment, thereby paving the way for further research and development in this field.

In their exploration of anomaly detection in crowd videos, the authors introduce a novel method that addresses the limitations of traditional approaches [11]. They propose an adaptive, training-less system capable of identifying anomalies "on-the-fly" by learning specific parameters during operation, which proves advantageous for surveillance tasks where pre-defining anomalies may be challenging due to unforeseen circumstances. Demonstrating competitive results compared to existing methods on benchmark datasets, the model showcases its efficacy in detecting anomalies in crowd videos. However, the authors acknowledge certain limitations and suggest avenues for future research, including addressing motion artifacts and managing high sparsity in crowded scenes. Despite these challenges, the innovative training-less approach shows promise for real-world surveillance applications, providing a dynamic and adaptable solution for anomaly detection in diverse and unpredictable environments. The paper outlines future directions to enhance the model's performance and robustness, offering a valuable roadmap for further advancements in crowd video anomaly detection. Through its inventive approach and potential for real-world deployment, this method contributes to the evolving landscape of surveillance technology, opening up new avenues for improving security and safety in various settings.

The work [12] delves into the domain of crowd behavior analysis, a relatively nascent field lacking an established taxonomy for its tasks and proposes a structured pipeline for such analysis, with each stage building upon the previous one. Emphasizing the significance of considering emotions in crowd behavior analysis, the paper particularly focuses on deep learning models for anomaly detection within crowds. It underscores the need for more realistic and challenging datasets to enhance the efficacy of current solutions. Moreover, the authors advocate for the integration of these models with existing video analytics systems to augment their utility and applicability in real-world scenarios. By advocating for a more structured approach to crowd behavior analysis and highlighting the importance of incorporating emotional aspects into the analysis, the paper contributes to advancing the understanding and methodologies in this evolving field. Through its call for utilizing real-world data for model development and integration with existing systems, the paper provides a roadmap for improving the practical applicability and effectiveness of crowd behavior analysis techniques. This comprehensive approach not only addresses the current gaps and limitations in the field but also lays the groundwork for future advancements and innovations in crowd behavior analysis and surveillance technologies.

In a departure from conventional approaches, the paper [13] introduces a pioneering framework for analyzing collective motion in crowds with a distinctive focus on individual movements and motivations. This innovative methodology incorporates an intention-aware model to capture individual motivations alongside a structure-based collectiveness measurement to assess how effectively individuals move together. Additionally, it employs a multi-stage clustering strategy to identify both local and global patterns of collective motion. Through empirical validation, the paper demonstrates significant improvements in measuring collective motion compared to existing methods, showcasing the framework's utility for anomaly detection and semantic scene segmentation tasks. Notably, the proposed method outperforms previous approaches, underscoring its efficacy and potential for various applications in crowd analysis and scene understanding. By providing a comprehensive approach that emphasizes individual movements and motivations while quantifying collective behavior, the framework offers promising avenues for more accurate anomaly detection and semantic segmentation in crowded scenes. This paper contributes to the advancement of crowd analysis methodologies by presenting a robust framework that not only builds upon existing techniques but also fosters new directions for research and application in the field of crowd behavior analysis and surveillance.

The paper [14] presents an innovative strategy for detecting anomalies in crowded environments, focusing on the intricate task of automatically classifying anomalies within crowd images. Through the introduction of the Aggregation of Ensembles (AOE) concept, the authors propose a novel methodology for anomaly detection in video data. This approach harnesses pre-trained Convolutional Neural Networks (ConvNets) and a diverse array of classifiers to achieve heightened accuracy in anomaly detection compared to established methods on benchmark datasets. By amalgamating various pre-trained CNN architectures, the methodology extracts distinct crowd features crucial for discerning differences between normal and anomalous events in crowded scenes, eliminating the need for training ConvNets from scratch and making it suitable for tasks with sparse samples of anomalous events. Additionally, the paper explores the impact of different optimization methodologies on fine-tuning, particularly in crowd behavior analysis, contributing significantly to advancing the field of crowd anomaly detection. Through empirical evaluation and comparison with established methods, the paper underscores the potential of the AOE approach in enhancing anomaly detection accuracy and robustness in crowded scenes, promising advancements in surveillance and security applications.

Innovating the landscape of crowd anomaly detection in video scenes, the study [15] introduces a pioneering framework that seamlessly integrates scene perception and consistency group segmentation to identify anomalies in crowd videos. Comprising four key stages, this novel framework begins with motion crowd segmentation, where Line Integral Convolution (LIC) is employed to extract motion regions from the flow field texture visualization. Subsequently, the scene perception-based consistency group segmentation stage clusters moving targets based on scene perception theory, discerning large groups such as multiple pedestrians or vehicles. Lastly, the anomaly detection stage combines appearance and fluid features to identify crowd anomalies using one-class Support Vector Machine (OC-SVM). Through validation on two datasets, UCSD Ped1 and UCSD Ped2, the paper showcases the framework's high accuracy in detecting anomalies in crowd videos. By integrating scene perception and consistency group segmentation, alongside leveraging advanced anomaly detection techniques like OC-SVM, the proposed framework presents a promising solution for robust and precise crowd anomaly detection. This research contributes significantly to the advancement of crowd surveillance, offering a comprehensive and efficient approach to identifying anomalies in intricate crowd scenes, with potential applications spanning domains such as security, public safety and urban planning.

Highlighting the pressing need for sophisticated mechanisms to measure, extract and process spatial and temporal data in crowd management, particularly in hazardous scenarios, the study underscores the imperative [16]. It advocates for the integration of Internet of Things (IoT) technologies, machine learning techniques and communication methods to sense crowd density, detect events early and predict potential accidents. While various machine learning methods have been applied to crowd management, the paper sheds light on the underexplored domain of deep hierarchical models, specifically Hierarchical Temporal Memory (HTM), for online learning and modeling temporal information. The proposed HTM framework is designed to detect anomalies in crowd movements and forecast potential overcrowding, offering a fresh perspective on enhancing crowd management strategies. By harnessing HTM's capabilities for learning temporal patterns and anomaly detection, the framework presents a promising avenue for improving crowd management efficacy and safety. The integration of IoT technologies and machine learning techniques with HTM holds the potential to transform crowd management practices by enabling proactive interventions and timely responses to evolving crowd dynamics and potential hazards. Through its innovative approach and emphasis on leveraging advanced technologies, the study contributes significantly to the advancement of the field of crowd management and public safety.

Introducing an adaptive training-less framework for anomaly detection in crowd scenes, the study [17] is grounded on the premise that anomalous events are rare while normal events are more prevalent. Comprising two core components, the framework consists of adaptive thresholding and motion-based anomaly detection. Adaptive thresholding involves dynamically adjusting the anomaly detection threshold based on the distribution of normal events within the scene. Meanwhile, motion-based anomaly detection entails analyzing the motion patterns of objects in the scene to identify anomalies. The efficacy of the framework is assessed on two datasets, UCSD and UMN, and is compared against other state-of-the-art methods. The results indicate that the proposed framework outperforms other methods in detecting anomalies in crowded scenes. Moreover, the framework demonstrates effectiveness in real-world scenarios, such as detecting abnormal behavior in surveillance videos. Overall, this innovative framework presents a fresh perspective on anomaly detection in crowd scenes, with potential applications across various domains including security and surveillance. By leveraging adaptive thresholding and motion-based anomaly detection, the framework furnishes a robust and efficient solution for identifying anomalous events amidst complex crowd dynamics, thus contributing to advancements in crowd surveillance technology and bolstering public safety measures.

# Chapter 3

# FRAMEWORK AND SYSTEM DESIGN

The proposed framework for crowd density analysis and anomaly detection combines the strengths of deep convolutional neural networks (CNNs) and advanced computer vision techniques. At the core of the system is the VGG-16 architecture, a well-established CNN model renowned for its ability to extract meaningful features from complex visual data. To enhance the network's capacity to capture contextual information across varying scales, the framework incorporates dilatable layers and Atrous Spatial Pyramid Pooling (ASPP). This multi-scale approach enables the model to effectively handle challenges like occlusion and scale variations commonly encountered in crowd scenes. The anomaly detection and movement categorization component of the system leverages optical flow analysis to compute the direction and speed of movement within the crowd. By defining thresholds for 'Fast Movement' and 'Normal Movement', the framework can identify and classify anomalous behaviors. The integration of YOLOv8, a state-of-the-art object detection and tracking model, further strengthens the system's ability to accurately locate and monitor individuals within the crowd. The cumulative movement direction of each tracked individual is analyzed, and anomalies are identified based on deviations from the average crowd movement patterns. The final output of the system includes annotated video frames with heatmaps and visualizations, providing intuitive representations of crowd dynamics and enabling the interpretation of crowd behavior and interactions.

## 3.1 Design Overview

The design of the proposed framework for crowd density analysis and anomaly detection follows a modular approach, leveraging the strengths of various deep learning and computer vision techniques. The core of the system is the VGG-16 CNN architecture, which is enhanced with dilatable layers and Atrous Spatial Pyramid Pooling to capture multi-scale contextual information. Optical flow analysis is employed to compute the direction and speed of movement within the crowd, enabling the identification and categorization of anomalies. The integration of the YOLOv8 object detection and tracking model further enhances the system's ability to locate and monitor individuals, allowing for the analysis of cumulative movement patterns. The final output includes annotated video frames with heatmaps and visualizations, providing a comprehensive understanding of crowd dynamics and behavior.

## 3.2 Dataset Description

The UMN dataset comprises 11 videos capturing both normal and abnormal crowd behaviors across three distinct scenes, including two outdoor and one indoor setting. Each video sequence starts with normal activities and transitions into abnormal behaviors, simulating scenarios like panic and escape situations. This dataset serves as a valuable resource for studying crowd dynamics and anomaly detection in diverse environments, providing insights into crowd behavior variations and responses to different stimuli.


Figure 3.1: The UMN Dataset

The UCF-QNRF dataset is a comprehensive collection of 1535 images depicting crowd scenes from various locations worldwide, showcasing a wide range of scenarios where crowds gather. This dataset offers a rich and diverse set of images capturing crowd density and distribution in different contexts, enabling researchers to explore crowd counting techniques and algorithms. With images showcasing crowds in different settings and scenarios, the UCF-QNRF dataset provides a valuable resource for developing and evaluating crowd counting models, facilitating advancements in crowd analysis, urban planning, and public safety.

Figure 3.2: UCF-QNRF - A Large Crowd Counting Data Set

The ShanghaiTech dataset is a comprehensive resource for crowd counting research, comprising a total of 1,198 annotated crowd images. The dataset is divided into two parts, Part-A and Part-B, with Part-A containing 482 images and Part-B containing 716 images. This large-scale dataset provides a diverse set of crowd scenes, enabling researchers to develop and evaluate crowd counting algorithms across a wide range of scenarios. The detailed annotations of the crowd density in each image serve as ground truth, allowing for the training and testing of machine learning models to accurately estimate the number of people in crowded environments. The ShanghaiTech dataset has become a widely used benchmark for assessing the performance of crowd counting techniques, driving advancements in this important field of computer vision.


Figure 3.3: Shanghaitech dataset

# 3.3 Methodology

## 3.3.1 Density of Crowd

In this study, we adopted a deep convolutional neural network (CNN) approach for crowd density analysis, leveraging the VGG-16 architecture as our foundational model. The VGG16 model is a deep convolutional neural network architecture renowned for its simplicity and effectiveness in image classification tasks. Developed by the Visual Geometry Group at the University of Oxford, VGG16 consists of 16 convolutional layers followed by fully connected layers. Its key feature lies in the use of small 3x3 convolutional filters, applied multiple times with max-pooling layers interspersed between them. This architecture allows VGG16 to capture intricate hierarchical features from input images, enabling accurate classification across various visual datasets. Despite its simplicity, VGG16 has demonstrated outstanding performance in large-scale image recognition competitions and remains a popular choice for deep learning tasks.. Specifically, we utilized the initial layers of a pre-trained VGG-16 network, capitalizing on its ability to extract meaningful features from crowd scenes for density estimation.

To enhance the network's capability in capturing contextual information across varying scales, we incorporated dilatable layers into our crowd density estimation models. These layers facilitate the adjustment of dilation rates within convolutional filters, thereby expanding the receptive field without inflating the filter size. Additionally, we employed the Atrous Spatial Pyramid Pooling (ASPP) is a technique employed to capture multi-scale contextual information from input feature maps as shown in the Fig.3.4. It involves applying multiple parallel dilated convolutions with different rates to the input feature map, allowing the network to gather information from various receptive field sizes. Additionally, ASPP incorporates a 1x1 convolutional layer and global average pooling to further enhance the model's ability to perceive contextual details. By aggregating information across different scales, ASPP enables the network to effectively handle objects of varying sizes and contexts, making it particularly useful in tasks like semantic segmentation and crowd density estimation where capturing diverse spatial information is crucial. This approach enables the model to gather contextual information at multiple scales, addressing challenges like occlusion and scale variations commonly encountered in crowd scenes. By integrating CNN layers with dilation and ASPP, our methodology aims to achieve robust and accurate crowd density estimation.

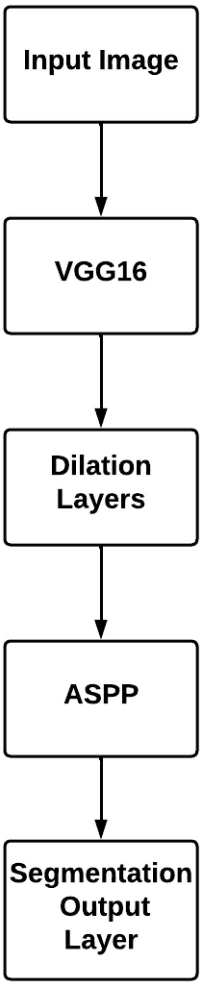



Figure 3.4: CNN with Dilation and ASPP

### 3.3.2 Identify Fast Moving People in the Crowd

Our methodology for anomaly detection and movement categorization in crowd scenes comprises several sequential steps to ensure comprehensive analysis. Initially, we preprocess the input video frames to enhance their quality and remove noise, preparing them for subsequent analysis. Following preprocessing, we employ feature detection techniques to identify key points and regions of interest within the crowd. Optical flow calculation is then applied to compute the motion vectors between consecutive frames, providing insights into the direction and speed of movement as shown in Fig.3.5.

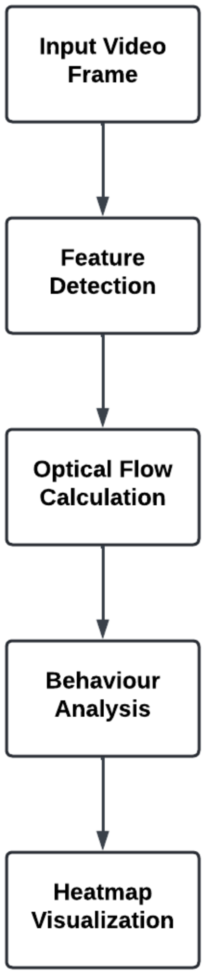



Figure 3.5: Optical Flow

In dense optical flow, we look at all of the points (unlike Lucas Kanade which works only on corner points detected by Shi-Tomasi Algorithm) and detect the pixel intensity changes between the two frames, resulting in an image with highlighted pixels, after converting to HSV format for clear visibility. It computes the magnitude and direction of optical flow from an array of the flow vectors, i.e., (dx/dt, dy/dt). Later it visualizes the angle (direction) of flow by hue and the distance (magnitude) of flow by value of HSV color representation.For visibility to be optimal, strength of HSV is set to 255. OpenCV provides a function cv2.calcOpticalFlowFarneback to look into dense optical flow.

Subsequently, behavior analysis is conducted to categorize detected anomalies based on their movement characteristics. Two primary categories are defined: 'Fast Movement' and 'Normal Movement', which are differentiated based on predefined thresholds of optical flow values. Finally, to visualize the detected anomalies and movement patterns, we generate heatmaps overlaying the original video frames. These heatmaps offer an intuitive representation of crowd dynamics, highlighting areas of interest and aiding in the interpretation of the analysis results.

### 3.3.3 Detect People in a Different Direction

In this research, we employed the YOLOv8 framework, an advanced version of the You Only Look Once (YOLO) object detection model, to achieve real-time people tracking within crowd scenes. YOLOv8 is the latest version of YOLO by Ultralytics. As a cutting-edge, state-of-the-art (SOTA) model, YOLOv8 builds on the success of previous versions, introducing new features and improvements for enhanced performance, flexibility, and efficiency. YOLOv8 supports a full range of vision AI tasks, including detection, segmentation, pose estimation, tracking, and classification. This versatility allows users to leverage YOLOv8's capabilities across diverse applications and domains.YOLOv8 is an anchor-free model. This means it predicts directly the center of an object instead of the offset from a known anchor box.


Figure 3.6: YOLO tracking

The model, implemented using the Ultralytics library, processes each video frame to detect and track individuals, generating bounding boxes and unique track IDs for each person as shown in Fig.3.6. The tracking information is stored as track histories and direction vectors to analyze the movement patterns and directions of individuals over consecutive frames. To compute the cumulative movement direction of each tracked individual, we utilized the historical track data. For each person, we calculated the relative direction vectors between consecutive points along their track, normalizing and scaling these vectors for visualization purposes. Furthermore, we employed a mechanism to identify and categorize anomalies based on wrong movement directions within the crowd. By analyzing the average direction vectors across all tracks and visualizing them using arrowed lines on the annotated frames, our methodology provides insights into crowd dynamics and individual movement behaviors, enhancing the understanding of crowd interactions and behaviors.

## 3.4 Architectural Design

The architectural design of our methodology integrates state-of-the-art deep learning techniques for crowd density analysis, anomaly detection, and movement categorization. Leveraging the VGG-16 architecture as the foundational model for crowd density estimation, we enhance its capabilities by incorporating dilatable layers and Atrous Spatial Pyramid Pooling (ASPP). These additions allow the network to capture contextual information across varying scales, addressing challenges such as occlusion and scale variations in crowd scenes. By integrating CNN layers with dilation and ASPP, our approach aims to achieve robust and accurate crowd density estimation, essential for understanding crowd dynamics.

For anomaly detection and movement categorization, we employ a sequential process that begins with preprocessing video frames to enhance quality and remove noise. Feature detection techniques are then utilized to identify key points and regions of interest within the crowd, followed by optical flow calculation to compute motion vectors between consecutive frames. Behavior analysis categorizes anomalies into 'Fast Movement' and 'Normal Movement' based on predefined thresholds of optical flow values. Finally, heatmaps are generated to visualize detected anomalies and movement patterns, providing intuitive representations of crowd dynamics and aiding in result interpretation.

In our real-time people tracking system, we employ the YOLOv8 framework, an advanced iteration of You Only Look Once (YOLO) object detection model. YOLOv8 processes every video frame to detect and track individuals, producing bounding boxes and unique track IDs. Leveraging historical track data, we compute cumulative movement directions for each tracked person, enabling the analysis of movement patterns over successive frames. Anomaly detection is enhanced by scrutinizing average direction vectors across all tracks, thus deepening insights into crowd dynamics. Overall, our architectural design seamlessly blends deep learning models with sequential processes, offering robust and holistic crowd analysis capabilities.

## 3.5 Workflow

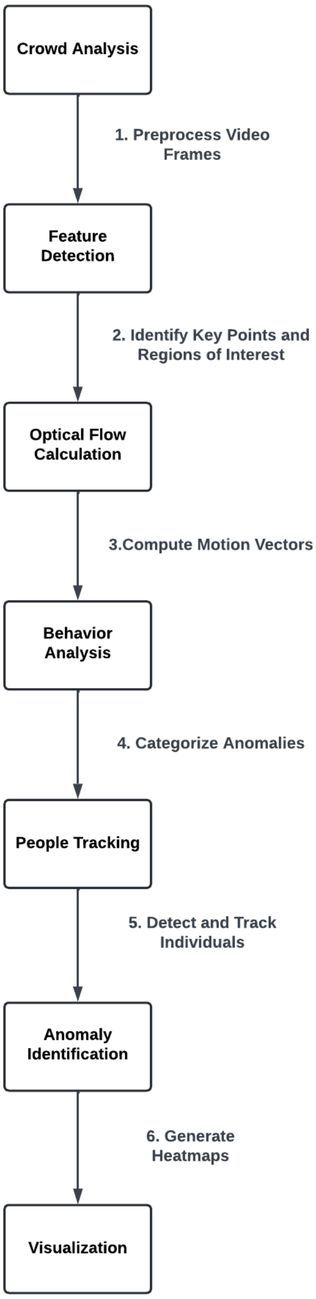



Figure 3.7: Work Flow

The workflow begins with preprocessing video frames to enhance quality and remove noise, followed by feature detection to identify key points and regions of interest within the crowd.Optical flow calculation computes motion vectors between frames, enabling behavior analysis to categorize anomalies like 'Fast Movement' or 'Normal Movement' based on predefined thresholds. Subsequently, people tracking detects and tracks individuals, facilitating movement analysis to compute cumulative movement directions. Finally, anomaly identification identifies and visualizes anomalies as mention in the Fig.3.7, generating heatmaps for intuitive representation of crowd dynamics and behaviors.

This sequential process integrates various techniques such as deep learning models for crowd density estimation and object detection, alongside traditional computer vision methods like optical flow calculation. Each step contributes to a comprehensive understanding of crowd interactions and behaviors, enabling real-time analysis and visualization of crowd dynamics for applications in security, event management, and urban planning.

# Chapter 4

# Implementation

Algorithm 4.1 for video processing, density estimation, and speed calculation using optical flow begins by initializing necessary variables, models, and capturing objects. The user is prompted to choose between camera input or video file input. Within the main processing loop, frames are continuously read from the selected input source and resized to a standard size. The algorithm then applies density estimation to each frame, generating a density overlay. Simultaneously, it calculates the speed and movement patterns using optical flow between consecutive frames. The original frame is combined with the density overlay and speed information for visualization, and the processed frame is displayed in a window. The algorithm checks for user input to exit and releases all.

---

**Algorithm 4.1:** Density of the frame and Fast Moving Object Detection

---

**Input**: Frame from input source
**Output**: Frame with density overlay and movement information
**Step 1:** Initialize variables, models, and capture object
**Step 2:** Choose input source (camera or video)
**while** true **do**
  **Step 3:** Read frame from input source
  **Step 4:** Resize frame to standard size
  **Step 5:** Estimate density using `density()` function
  **Step 6:** Estimate speed using `speed_estimation1()` function
  **Step 7:** Combine frames and overlays for visualization
  **Step 8:** Display combined frame
  **if** user presses quit key **then**
    **Step 9:** Break loop
  **end if**
**end while**
**Step 10:** Release capture object and close windows

---

Algorithm 4.2 outlines the steps involved in density estimation. It begins by loading the pre-trained density estimation model and then reads a frame from the input source. The frame is converted to a PIL Image and undergoes a transformation process. Inference is performed to generate a density map, which is then resized to match the frame size and normalized for visualization. The algorithm overlays the density map on the frame, applies thresholding to identify regions of high density, counts the number of people as contours, and finally resizes and displays the overlay on the frame for visual inspection.

**Algorithm 4.2:** Density Estimation Algorithm

**Input**: Frame from input source with dimensions $H \times W \times C$ (height $H$, width $W$, channels $C$)
**Output**: Frame with density overlay

**Step 1:** Load the pre-trained density estimation model with VGG, dilation, and ASPP layers.
The ASPP layer output size formula is given by:
$H_{out} = 1, \quad W_{out} = 1, \quad C_{out} = n \times C_{in}$
**Step 2:** Read frame from input source
**Step 3:** Convert frame to PIL Image
**Step 4:** Perform inference to get density map using the pre-trained model
**Step 5:** Resize density map to match frame size
**Step 6:** Normalize density map for visualization
**Step 7:** Overlay density map on the frame
**Step 8:** Threshold density map to identify regions of high density
**Step 9:** Count the number of people as contours
**Step 10:** Resize and display overlay on the frame

Algorithm 4.3 details the process of fast-moving object detection using optical flow. It starts by calculating the optical flow between consecutive frames. A threshold is applied to select regions with high flow, and the magnitude values are normalized and resized for visualization. A heatmap is created using a hot color map and overlaid on the frame. The resulting frame with the heatmap is displayed for visual inspection. The algorithm further classifies movement based on the average flow magnitude, adds movement labels and speed information to the frame, and returns the frame with the heatmap and movement information for analysis and visualization.

---

**Algorithm 4.3:** Fast Moving Object Detection Algorithm with Optical Flow Angle Calculation using Farneback Method

---

**Input**: Frame from input source
**Output**: Frame with heatmap and movement information

**Step 1:** Calculate optical flow using Farneback method:
Let $I_{\text{prev}}$ be the previous frame and $I_{\text{curr}}$ be the current frame
Compute optical flow using Farneback method:
$\text{flow} = \text{cv2.calcOpticalFlowFarneback}(I_{\text{prev}}, I_{\text{curr}}, \text{None}, 0.5, 3, 15, 3, 5, 1.2, 0)$
Extract flow vectors $u$ and $v$ from flow
Compute flow magnitudes $M = \sqrt{u^2 + v^2}$
Compute angle $\theta(x, y)$ between optical flow vectors
**for** each pixel $(x, y)$ in the frame **do**
Compute angle $\theta(x, y) = \arccos\left(\frac{(u^{-1}(x,y)\cdot u_1(x,y)+v^{-1}(x,y)\cdot v_1(x,y))}{\sqrt{u^{-1}(x,y)^2+v^{-1}(x,y)^2}\cdot\sqrt{u_2(x,y)^2+v_2(x,y)^2}}\right)$
**end for**
**Step 2:** Apply threshold to select high flow regions based on angle $\theta(x, y)$
**Step 3:** Normalize and resize the thresholded magnitude values
**Step 4:** Create heatmap with hot color map
**Step 5:** Resize heatmap to match frame size
**Step 6:** Overlay heatmap on the frame
**Step 7:** Display the frame with heatmap for visualization
**Step 8:** Classify movement based on average flow magnitude and angle $\theta(x, y)$
**Step 9:** Add movement label and speed information to the frame
**Step 10:** Return the frame with heatmap and movement information

---

Algorithm 4.4 begins with initializing necessary libraries and modules and imports custom modules for object tracking and detection. It defines constants, variables, and initializes a YOLO model for object detection. Using a video file as input, it enters a loop to process each frame. Within the loop, it tracks objects using the YOLO model, updates track history and direction vectors, and displays annotated frames with tracked objects. The loop also adjusts frame skipping for real-time processing based on timing calculations. After processing, it releases capture device resources and closes OpenCV windows.

---

**Algorithm 4.4:** Object Tracking and Direction Estimation using YOLOv8 and OpenCV

---

**Input:** Video file or camera input
**Output:** Annotated frames with tracking and direction vectors

**Step 1:** Import necessary libraries: cv2, time, numpy, defaultdict, YOLO
**Step 2:** Load YOLOv8 model
**Step 3:** Open video file or camera input for processing
**Step 4:** Initialize track history, direction vectors, and timing variables
**Step 5:** Set desired frames per second (FPS)
**Step 6:** Initialize frame counter and frame skip variable
**while** video feed is open **do**
  **Step 7:** Read a frame from the input source
  **Step 8:** Perform YOLOv8 object tracking on the frame
  **Step 9:** Extract bounding boxes and track IDs from detection results
  **Step 10:** Visualize tracking results on the frame (plot tracks and direction vectors)
  **Step 11:** Calculate average direction vectors across all tracks
  **if** frame counter is a multiple of frame skip variable or first frame **then**
    **Step 12:** Display annotated frame with tracking and direction vectors
    **if** 'q' key is pressed **then**
      **Step 13:** Break the loop
    **end if**
  **end if**
  **Step 14:** Calculate processing time for the frame
  **Step 15:** Update total processing time for timing adjustments
  **if** frame counter is 11 **then**
    **Step 16:** Calculate average time per 10 frames
    **Step 17:** Calculate frames per second for the last 10 frames
    **Step 18:** Calculate frame skip variable for real-time processing
  **end if**
  **Step 19:** Increment frame counter
**end while**
**Step 20:** Release video capture object and close display windows

---

# Chapter 5

# Experiments and Results

## 5.1 Results For Density and Fast Movement

In this implementation, we present the output frames,obtained from two distinct methodologies for crowd analysis(abnormal movements and crowd density estimation) as shown in Fig.5.1 and Fig.5.2:

- **Top left subframe**: Original video.
- **Top right subframe**: Anomaly detection with predefined thresholds categorizing anomalies as fast or normal movement based on optical flow values.
- **Bottom left subframe**: VGG16-based density estimation with anomaly thresholds based on deviations from average crowd density.
- **Bottom right subframe**: Overall summary of the analysis with different anomaly thresholds applied to categorize anomalous behavior in crowd scenes.

Table 5.1: Comparison of MAE, RMSE, and GAME for different approaches on three datasets

| **Approach** | **ShanghaiTechA** | | | **ShanghaiTechB** | | | **UCF CC 50** | | |
|---|---|---|---|---|---|---|---|---|---|
| | **MAE** | **RMSE** | **GAME** | **MAE** | **RMSE** | **GAME** | **MAE** | **RMSE** | **GAME** |
| Baseline | 99.1 | 163.7 | 120.9 | 31.3 | 43.4 | 42.4 | 534.4 | 770.6 | 566.5 |
| Dilated | 81.7 | 130.4 | 99.8 | 27.2 | 38.2 | 36.6 | 454.5 | 664.7 | 497.2 |
| ASPP | 85.7 | 134.3 | 104.3 | 38.5 | 48.5 | 47.5 | 414.0 | 619.3 | 459.1 |

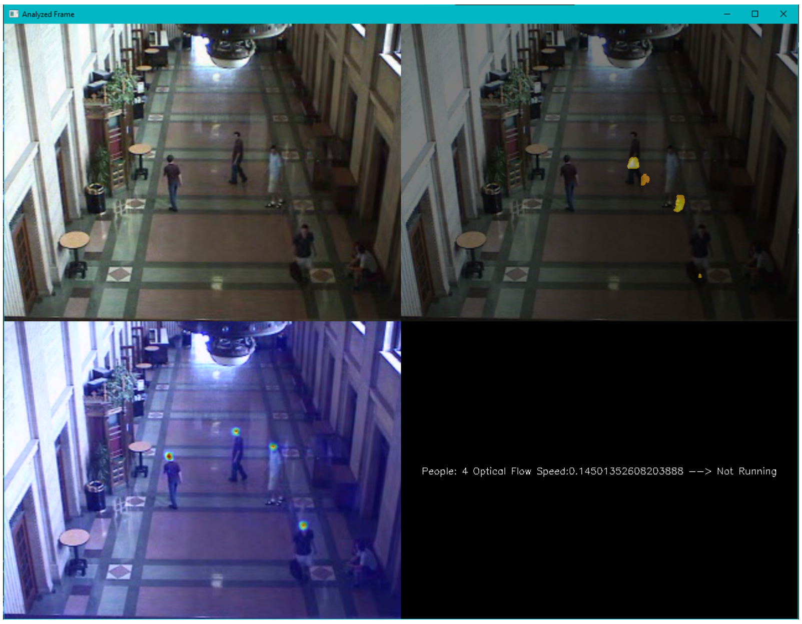



Figure 5.1: Anomaly detection results in a normal scenario


Figure 5.2: Anomaly detection results in an abnormal scenario

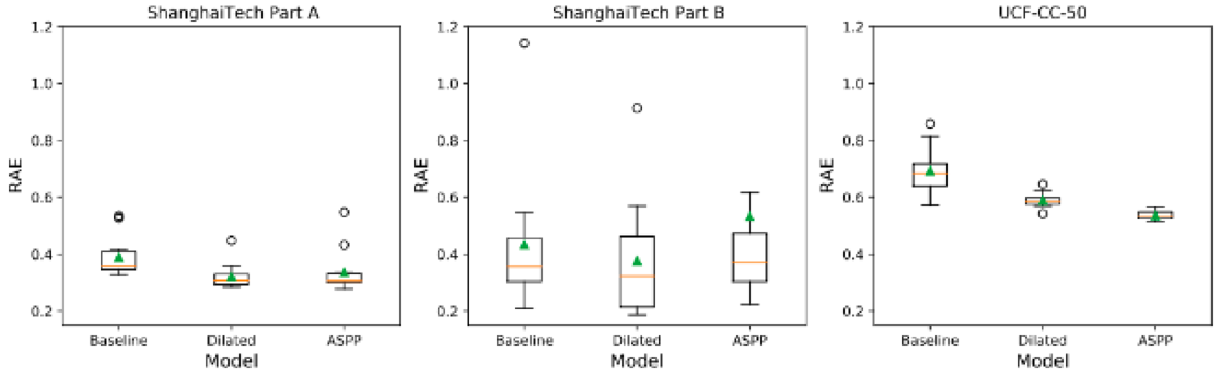



Figure 5.3: Distribution of RAE values over 12 trains of each model (lower is better)

## 5.2 Results For Tracking

In this implementation, we present the output frames which integrates YOLOv8 object tracking and movement direction calculation Frame utilizes YOLOv8 object tracking to identify and categorize anomalies, employing predefined thresholds for direction deviation. Frames depict instances of individuals moving in unexpected directions relative to the crowd, signaling potential disruptions or safety hazards as shown in Fig.5.4, Fig.5.5 and Fig.5.6.

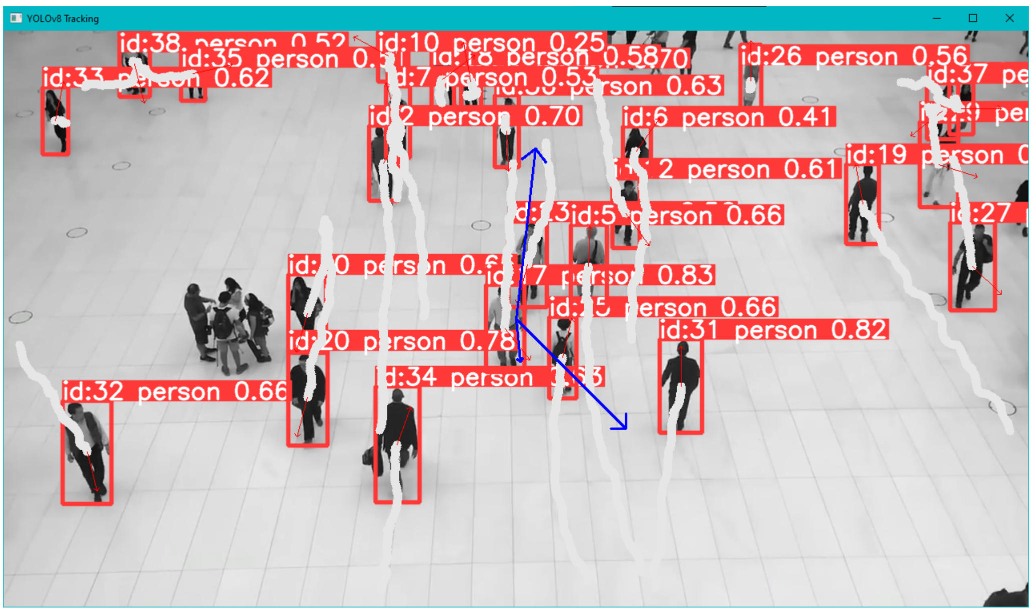



Figure 5.4: Crowd general flow direction in scenario 1


Figure 5.5: Crowd general flow direction in scenario 2

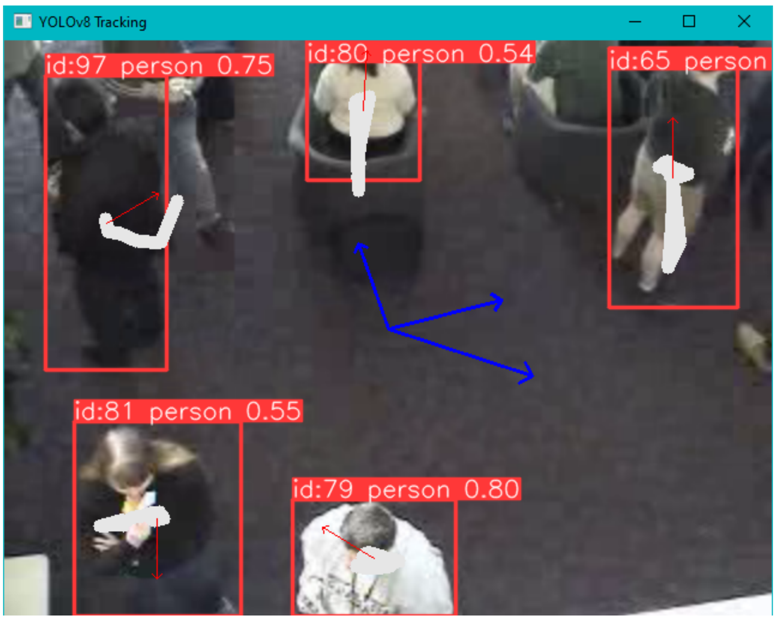



Figure 5.6: Crowd general flow direction in scenario 3

# Chapter 6

# Conclusion and Future Scope

## 6.1 Conclusion

The comprehensive approach presented in this study offers a robust framework for crowd management and safety at large-scale events, integrating advanced technologies like deep learning and computer vision to enhance safety protocols and improve the overall experience for participants and staff. Accurate estimation of crowd density is achieved through a deep convolutional neural network based on VGG-16 architecture, providing crucial insights for effective crowd management. Anomaly detection and movement categorization are accomplished through preprocessing, feature detection, and optical flow calculation, resulting in intuitive representations of crowd dynamics that enable swift response to any issues that may arise. The employment of YOLOv8 framework for real-time people tracking further facilitates the detection and tracking of individuals within crowd scenes, contributing to enhanced safety and security measures. This comprehensive approach seamlessly combines cutting-edge technologies to address the challenges of crowd management, ensuring the safety and satisfaction of attendees while optimizing the overall success of large-scale events. The integration of deep learning, computer vision, and real-time tracking capabilities offers a versatile and adaptable solution that can be tailored to the unique requirements of diverse event settings, ultimately enhancing the safety and enjoyment for all involved.

## 6.2 Future Scope

While the study establishes a foundation for crowd management and safety, potential areas for improvement are identified. These include the integration of multiple sensor modalities to enhance anomaly detection and crowd analysis capabilities, enabling a more comprehensive understanding of crowd dynamics. The development of real-time decision support systems can assist event organizers and security personnel in effectively responding to emergencies, facilitating swift and appropriate actions. Enhancing the robustness and scalability of the proposed approaches is crucial to accommodate varying crowd sizes and environmental conditions, ensuring the applicability across diverse event settings. Modeling and predicting crowd behavior can enable proactive crowd management strategies, allowing for the anticipation and mitigation of potential issues before they arise. Optimizing crowd flow, reducing wait times, and personalizing event experiences can enhance the overall user experience for attendees, contributing to safer and more enjoyable large-scale events. Addressing these areas of improvement can advance the field of crowd management and safety, ensuring the well-being and satisfaction of all participants, from organizers to attendees and staff. Integrating cutting-edge technologies, developing intelligent decision support systems, and prioritizing user experience can collectively elevate the standards of crowd management, fostering safer and more successful large-scale events that cater to the needs of all stakeholders.

# Bibliography


[1] C. Feliciani, A. Corbetta, M. Haghani, and K. Nishinari, "Trends in crowd accidents based on an analysis of press reports," *Safety Science*, vol. 164, p. 106174, 2023.

[2] K. Still, M. Papalexi, Y. Fan, and D. Bamford, "Place crowd safety, crowd science? case studies and application," *Journal of Place Management and Development*, vol. 13, no. 4, pp. 385–407, 2020.

[3] Y. Hao, Z. J. Xu, Y. Liu, J. Wang, and J. L. Fan, "Effective crowd anomaly detection through spatio-temporal texture analysis," *International Journal of Automation and Computing*, vol. 16, no. 1, pp. 27–39, 2019.

[4] M. A. Khan, H. Menouar, and R. Hamila, "LCDnet: A lightweight crowd density estimation model for real-time video surveillance," *Journal of Real-Time Image Processing*, vol. 20, no. 2, p. 29, 2023.

[5] N. Nayan, S. S. Sahu, and S. Kumar, "Detecting anomalous crowd behavior using correlation analysis of optical flow," *Signal, Image and Video Processing*, vol. 13, pp. 1233–1241, 2019.

[6] X. Li, W. Li, B. Liu, and N. Yu, "Object and patch based anomaly detection and localization in crowded scenes," *Multimedia Tools and Applications*, vol. 78, pp. 21 375–21 390, 2019.

[7] A. Ramchandran and A. K. Sangaiah, "Unsupervised deep learning system for local anomaly event detection in crowded scenes," *Multimedia Tools and Applications*, vol. 79, no. 47, pp. 35 275–35 295, 2020.

[8] C. Direkoglu, "Abnormal crowd behavior detection using motion information images and convolutional neural networks," *IEEE Access*, vol. 8, pp. 80 408–80 416, 2020.

[9] X. Ding, F. He, Z. Lin, Y. Wang, H. Guo, and Y. Huang, "Crowd density estimation using fusion of multi-layer features," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 8, pp. 4776–4787, 2020.

[10] X. Zhang, X. Yang, W. Zhang, G. Li, and H. Yu, "Crowd emotion evaluation based on fuzzy inference of arousal and valence," *Neurocomputing*, vol. 445, pp. 194–205, 2021.

[11] A. Sikdar and A. S. Chowdhury, "An adaptive training-less framework for anomaly detection in crowd scenes," *Neurocomputing*, vol. 415, pp. 317–331, 2020.

[12] F. L. Sánchez, I. Hupont, S. Tabik, and F. Herrera, "Revisiting crowd behaviour analysis through deep learning: Taxonomy, anomaly detection, crowd emotions, datasets, opportunities and prospects," *Information Fusion*, vol. 64, pp. 318–335, 2020.

[13] X. Li, M. Chen, and Q. Wang, "Quantifying and detecting collective motion in crowd scenes," *IEEE Transactions on Image Processing*, vol. 29, pp. 5571–5583, 2020.

[14] K. Singh, S. Rajora, D. K. Vishwakarma, G. Tripathi, S. Kumar, and G. S. Walia, "Crowd anomaly detection using aggregation of ensembles of fine-tuned convnets," *Neurocomputing*, vol. 371, pp. 188–198, 2020.

[15] X. Zhang, D. Ma, H. Yu, Y. Huang, P. Howell, and B. Stevens, "Scene perception guided crowd anomaly detection," *Neurocomputing*, vol. 414, pp. 291–302, 2020.

[16] A. Bamaqa, M. Sedky, T. Bosakowski, and B. B. Bastaki, "Anomaly detection using hierarchical temporal memory (htm) in crowd management," in *Proceedings of the 2020 4th International Conference on Cloud and Big Data Computing*. IEEE, August 2020, pp. 37–42.

[17] A. Sikdar and A. S. Chowdhury, "An adaptive training-less framework for anomaly detection in crowd scenes," *Neurocomputing*, vol. 415, pp. 317–331, 2020.