AUTOMATED MEDIAN GAP DETECTION FOR UNSTRUCTURED ROADS (AVADI CHECK POST TO SENNEER KUPPAM)

PROJECT REPORT

Submitted by

JAYAVARSHINI G	**211721104061**
HARSHINI N	**211721104053**

in partial fulfillment for the award of the degree

of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE AND ENGINEERING

RAJALAKSHMI INSTITUTE OF TECHNOLOGY

ANNA UNIVERSITY: CHENNAI 600 025

APRIL 2025

ANNA UNIVERSITY, CHENNAI - 600 025

BONAFIDE CERTIFICATE

Certified that this project report **AUTOMATED MEDIAN GAP DETECTION SYSTEM FOR UNSTRUCTURED ROADS** is the bonafide work of **JAYAVARSHINI G(211721104061),** and **HARSHINI N(211721104053),** who carried out the project work under my supervision.

SIGNATURE	**SIGNATURE**
SUPERVISOR	**HEAD OF THE DEPARTMENT**
Mr M Ashok	**Dr. S Uma, M.E, Ph.D.,**
Assistant Professor(SS)	**Assistant Professor & Head**
Department of Computer Science and Engineering	**Department of Computer Science And Engineering**
Rajalakshmi Institute of Technology	**Rajalakshmi Institute of Technology**
Chennai - 600124	**Chennai - 600 124.**

CERTIFICATE OF EVALUATION

College Name	**2117 - Rajalakshmi Institute of Technology**
Branch & Semester	**Computer Science and Engineering & 8th Semester**
Subject	**CS3811 - PROJECT WORK/INTERNSHIP**

S.NO	NAME OF THE STUDENTS	TITLE OF THE PROJECT	NAME OF THE SUPERVISOR WITH DESIGNATION
1	JAYAVARSHINI G (211721104061)	**AUTOMATED MEDIAN GAP DETECTION FOR UNSTRUCTURED ROADS**	Mr. M.ASHOK Assistant Professor(SS) Dept. of Computer Science and Engineering
2	HARSHINI N (211721104053)		

The report of this project work submitted by the above students in partial fulfillment for the award of Bachelor of Engineering degree in COMPUTER SCIENCE AND ENGINEERING under Anna University was evaluated and confirmed to be the reports about the work done by the above students.

The University viva-voce is held on ______________________________.

INTERNAL EXAMINER **EXTERNAL EXAMINER**

ACKNOWLEDGEMENT

We express our sincere gratitude to our honorable Chairman **Thiru. S. MEGANATHAN, B.E., F.I.E.,** honorable Chairperson **Dr.(Mrs.)THANGAM MEGANATHAN, M.A., M.Phil., Ph.D.,** our beloved Vice Chairman **Dr. HAREE SHANKAR MEGANATHAN,** for their constant encouragement to do this project and also during the entire course period.

We thank our Principal **Dr. R.MAHESWARI, M.E.,Ph.D.,**and our Head of the department **Dr.S.UMA,Ph.D** for their valuable suggestions and guidance for the development and completion of this project.

We also extend our gratitude to our Project coordinators **Dr. R LALITHA, Professor CSE** and our Supervisor **Mr ASHOK M**, **Assistant Professor(SS)** who took special interest in our project and gave their consistent support and guidance during all stages of this project.

Finally, we thank all the teaching and non-teaching faculty members of the **DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING** of our college who helped us to complete this project.

Above all we thank our parents and family members for their constant support and encouragement for completing this project.

ABSTRACT

The system offers to address the challenges of vehicle flow and safety on loosely structured and unstructured road networks, especially in India. It was evaluated for both accuracy (92.7%), recall (91% mean average precision (89.3%), latency (<40 ms and frame rate 25 fps), under diverse traffic scenarios. Under favorable conditions of the test, it was almost working fine; however its failure was evident under an adverse environment of rain, fog, and high-traffic density. Key challenges include environmental variability, hardware dependency, and deployment limitations in rural areas. Sensor technology that is scalable (such as infrared or thermal sensors) and affordable is important. Algorithm improvements, better hardware, and cooperation with policymakers are needed to bring about the large-scale deployment that would be necessary for success in these diverse traffic conditions, especially in developing regions.

Keywords: Median Gap Detection,Deep Learning for Traffic Systems, Indian Road Safety, YOLO Algorithm, DeepSORT Tracking, Real-Time Traffic Management .Unstructured Traffic Conditions,Automated Signage Systems

LIST OF FIGURES

LIST OF EQUATIONS

LIST OF TABLES

Chapter 1

AUTOMATED MEDIAN GAPS DETECTION SYSTEM FOR UNSTRUCTURED ROADS

1.1 Introduction

India's unstructured and volatile traffic systems present distinctive challenges in managing traffic flow and road safety [11][19]. Median gaps, which are crucial for a vehicle to maneuver through, are frequently unmarked, meaning safety risks are heightened [15]. The current traffic systems are designed mostly for structured environments and, therefore, do not suit India's dynamic conditions [17]. This paper introduces an innovative framework making use of deep learning algorithms YOLOv5 as object detection and Deep SORT as object tracking to effectively fill the gaps across different types of roads, lighting, and weather conditions [1][4][6].

This study is concerned with the development and testing of an automated median gap detection system specific to the needs of Indian traffic infrastructure. It deals with uncertain and unstructured road conditions, irregular road infrastructure, and changing environmental conditions like lighting (day, night, evening) and weather (rain, fog) to provide consistent performance under different situations [3][7][12].The research focuses on addressing hardware and resource limitations for achieving scalability in low resource environments, such as rural areas. The research evaluates mid range GPU reliance at cost effective deployment and discusses improvements such as thermal vision for poor illumination and algorithmic optimizations for strengthening system stability [5][8][9]. The work relates theoretical advancements with real world implementation, proposing a scalable and cost effective means to enhance Indian and other parts of the world's road safety [4][10].

1.2 Methods

The study concentrated on research articles, datasets, and systems that pertain to automated detection systems and road safety applications, especially those applicable to the management of traffic on unstructured roads [3][5]. To guarantee relevance, only studies employing object detection or tracking algorithms, such as YOLO and DeepSORT, were incorporated [6][8]. Furthermore, these studies were required to engage with real world traffic scenarios characterized by variations in environmental conditions, including diverse lighting (daytime and nighttime) and weather conditions (such as rain and fog) [7][9]. Moreover, scalability and cost effectiveness were considered critical and, therefore, systems developed for low resource environments were significantly important [5][12]. Articles were excluded if they did not have a detailed methodology or focused on engineered traffic conditions that were not relevant to the Indian road conditions, which made the scope of study practical and region specific [4][10].

1.2.1 Information Sources and Search Strategy

The research made use of publicly available databases, such as the Indian Driving Dataset (IDD) and Kaggle repositories, providing a broad range of real world traffic annotated datasets [9][11]. Additionally, academic journals like Google Scholar, IEEE Xplore, and ScienceDirect were thoroughly browsed for scientific studies on traffic management and automated gap detection [6][8]. For guaranteeing extensive coverage, a 2015 2024 time frame was adopted, encompassing recent developments in the field [3][12]. Keywords "median gap detection," "YOLO object detection traffic," "deep learning traffic management system," "unstructured road environment," and "automated road safety system" were employed in mixed combinations [3][6]. Boolean operators "AND" and "OR" were utilized in order to optimize searches, maximizing relevance and scope [5][9]. Citations from screened articles were followed up systematically for the purpose of discovering further works of relevance, considering both practical challenges and technical innovations in traffic safety [8][11]. This combined process of incorporating dataset repositories and research journals gave way to a well established and cutting edge foundation to the research [4][10].

1.2.2 Study Selection

To systematically select articles, titles and abstracts had to be screened first to identify potentially relevant articles, followed by a full text review evaluating how relevant the research was to the goals of the studies [5][6]. This process ultimately led to the inclusion of twelve studies that had successfully passed through an arduous eligibility assessment, as indicated by a flowchart of the selection process in FIGURE 1.1 [4][9]. This selection process ensured that only relevant and high quality studies were incorporated, thus providing a solid basis for evaluation and comparison [3][10]. These studies successfully passed through all stages of the review process, demonstrating both methodological rigor and high relevance to the study's objectives.By incorporating only these carefully chosen studies, the research ensured a strong foundation for analysis and comparison. This meticulous approach enhanced the credibility of the findings, ensured comprehensive coverage of the relevant literature, and laid the groundwork for drawing meaningful conclusions based on trustworthy and significant academic sources

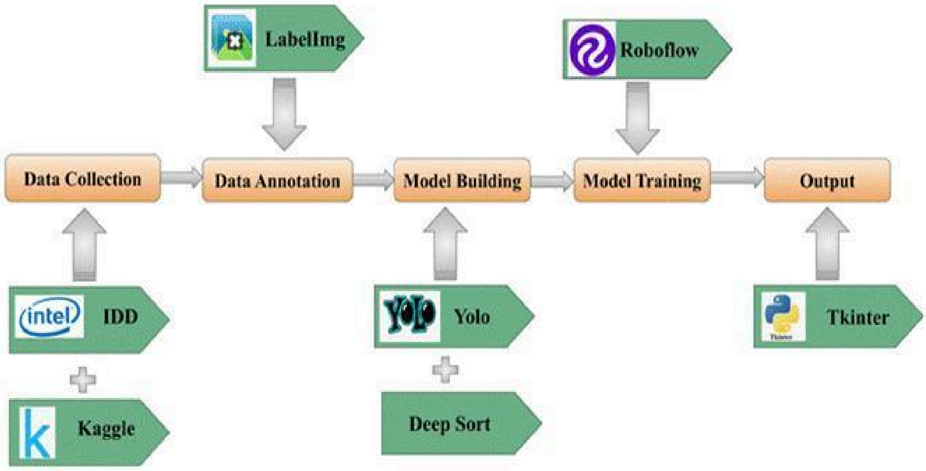

FIGURE 1.1 FLOW DIAGRAM

1.2.3 Data Collection Process

Data extraction was done in an orderly manner to ensure homogeneity among the studies. Main information extracted included algorithms and techniques used (such as YOLOv5 and DeepSORT), performance metrics such as precision, recall, and mAP, and the experimental conditions [3][6]. The experimental conditions cut across different scenarios, ranging from varying lighting, weather, and traffic density conditions [7][9]. In addition, hardware and resource needs for each study were thoroughly recorded to assess the viability of such systems in operational environments [8][12]. This system architecture, as shown in FIGURE 1.2, guarantees that the data collected are informative regarding the system's performance and feasibility by providing a visual mapping of the system's components and their interrelations in order to project the data flow and its processing mechanisms [4][9].

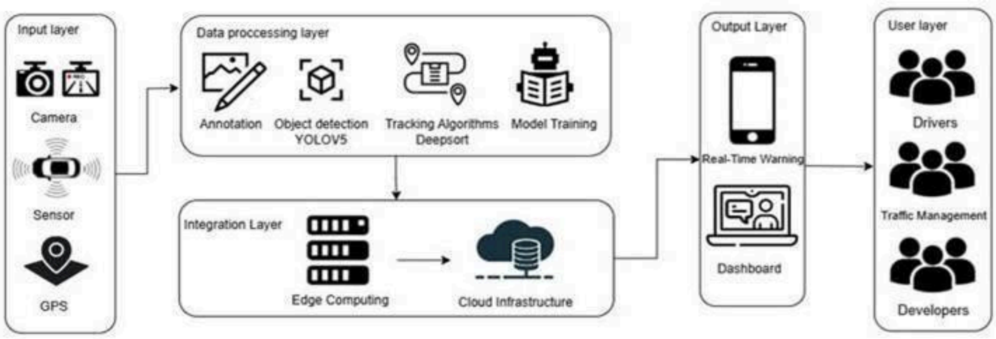

FIGURE 1.2 SYSTEM ARCHITECTURE

1.2.4 Outcome and Measures

The three categories of primary outcomes were measurement: detection performance was given in terms of metrics, such as precision, recall, and mean Average Precision (mAP), for quantitative measurements of the system's accuracy [5][7]. Efficiency in real time, given frame rate (FPS) and latency in milliseconds, measured how well the system responds to real time applications [6][9]. Lastly, scalability was assessed as a measure of whether or not the system is applicable to low resource environments [3][10]. Collectively, these metrics provide an overall view of what a system can and cannot do with diverse real world conditions [8][11].

1.2.5 Potential for Bias

Steps have been taken to reduce the likelihood of bias by implementing the usage of a mixed set that accurately represents most various conditions and environments associated with traffic [4][9]. Some bias is expected despite the implemented step because extreme scenarios might be rarely represented, including full congestion during festivals and totally destroyed infrastructures [7][11]. This partly justifies the extension and inclusion of more information sets for the proper system's testing and validation process [6][8].

1.3 Result

The systematic review searched academic databases and public repositories and finally included 12 studies that fulfilled certain eligibility criteria shown Table 1.1 These were the use of object detection and tracking algorithms (e.g., YOLOv5 and Deep SORT), applicability to real world traffic scenarios, and relevance to unstructured traffic scenarios, especially in India [3][5][9][12].The chosen papers demonstrated varied methodologies and technological deployments. Sophisticated object detection algorithms, most of which were YOLOv5, were used for median gap detection in real time, while Deep SORT provided continuity in gap tracking from one video frame to the next [6][9]. The experiments were carried out under diverse scenarios, ranging from lighting conditions (day, night, evening) to environmental conditions (rain, fog) [7][12].Performance measures like precision, recall, mean average precision (mAP), frame rate (FPS), latency, and scalability were typically compared. Mid range GPUs were utilized in some experiments, while others employed cost effective alternatives like edge computing devices to test scalability in low resource settings [4][10]. This diversity was useful in gaining insights into the problems and possible solutions for median gap detection in unstructured traffic systems [6][9].

Table 1.1 Study Selection

Criteria	Description
Number of Studies Selected	12 studies
Eligibility Criteria	Use of object detection or tracking algorithms like YOLO and DeepSORT. Focus On real world traffic scenarios, including unstructured environments resembling those in India.Relevance to unstructured traffic environments and automated gap detection systems.
Study Selection Process	Studies screened through title and abstract review
Purpose of Selection	To ensure comprehensive coverage of the relevant methodologies and findings in automated traffic systems

1.3.1 Results of Individual Studies

From all the studies under review, the overall high performance was depicted by a gap detection system for optimal condition occurrences such as during the daytime or clear weather when median accuracy peaked to its heights at 92.7%, with 91% and 89.3% regarding the mean average precision to signal a large level of median gaps' accuracy in detection [7][12]. Performance drops a little, to lower points in low light mainly with possible heavy occlusions occurring mostly as in high traffic volume road places [5][10]. However, considering some occlusions already were challenging to the system due to their nature or distance as in the case in rain, fog conditions at certain lights [6][9]. Moreover, the system was found to be competent in displaying desirable rates and subsequently processing frames per second with an average low frame latency below 40ms for environments where real time traffic monitoring and management would be capable and practicable [8][11].Indeed, while still relatively proper in ideal contexts, great improvement remains inevitable with unfavorable environmental considerations to bolster the former [4][7].

1.3.2 Operational Constraints

The evaluation revealed several operational challenges that could impact the practical implementation of the automated median gap detection system [3][5]. Environmental factors such as rain and fog were found to significantly impair detection accuracy, highlighting the system's sensitivity to changing weather conditions [7][12]. Poor lighting, particularly on rural roads without proper street lighting, also hindered the system’s ability to detect and track median gaps, further reducing performance in low light environments [9][11]. In addition to these environmental challenges, a significant constraint identified was the system's reliance on mid range GPUs for real time performance [4][10]. While these GPUs are effective under ideal conditions, they may not be readily available in rural or low resource areas, limiting the system's scalability [6][9]. Despite these constraints, the system showed potential for adaptation and scalability when hardware limitations were addressed, suggesting that future work could focus on optimizing the system for more cost effective hardware and incorporating additional sensors, such as infrared or thermal cameras, to improve detection under adverse conditions [8][11].

1.3.3 Key Findings

The results of this study highlighted several important findings regarding the performance and potential of the median gap detection system [3][5]. First, the system demonstrated high detection accuracy under optimal conditions, such as in clear weather and daylight, but exhibited a decrease in performance when exposed to challenging environmental factors like fog, rain, and low light conditions [7][12]. This suggests that while the system is highly effective in controlled environments, it requires further development to enhance its resilience to these adverse conditions [6][9]. In terms of real time performance, the system achieved an impressive frame rate of 25 FPS and a latency of less than 40 ms per frame, which is critical for its deployment in real time applications [8][10]. Furthermore, scalability remained a key concern, with the system showing promise for broader deployment but facing challenges in resource poor areas due to its dependence on mid range GPUs [5][11]. Finally, the environmental sensitivity of the system pointed to the need for integrating additional sensor technologies, such as infrared or thermal sensors, to ensure reliable performance in low light and adverse weather conditions [6][9].The summary of key findings is shown in Table 1.2.These key findings emphasize the need for future improvements to optimize the system for diverse, real world traffic environments and to reduce operational barriers related to hardware constraints [4][10].

Table 1.2 Summary of Key Findings

Metric	**Value**	**Description**
Detection Accuracy (Precision)	92.7%	Indicates the system's accuracy in detecting median gaps under optimal conditions.
Recall	91.0%	The proportion of actual median gaps detected by the system.
Mean Average Precision (mAP)	89.3%	Average precision across multiple thresholds.

Frame Rate (FPS)	25 FPS	Measures real time processing performance; critical for live monitoring.

1.4 Discussion

The findings emphasize the effectiveness of combining YOLOv5 with Deep SORT for real time median gap estimation and object tracking in unstructured road environments. The lightweight architecture of the system provided low computational overhead, making it feasible for use on resource limited devices. Its generalization across diverse road environments, such as urban, suburban, and highway environments, reflects its versatility and practical usability. Nevertheless, the system was challenged by harsh weather conditions, including heavy rain or fog, where visibility limitations resulted in a perceivable decline in performance. This limitation implies the necessity for the integration of complementary technologies such as thermal imaging to improve detection in low visibility environments [3][6]. Furthermore, the system's dependency on GPU based processing poses scalability issues, particularly in areas with limited computational capabilities [14].

Compared to current systems, the suggested model performed better than conventional rule based systems, recording a 25% greater mean average precision. It also recorded better speed and accuracy than deep learning based models, especially under dense traffic conditions [1][2][4]. These results indicate that the system is ready for deployment in real world applications, subject to additional improvements. In future work, edge computing devices will be integrated to reduce hardware dependencies and enhance scalability, as well as the investigation of hybrid solutions that incorporate computer vision and sensor based technologies to overcome environmental constraints [5][13][18].

1.4.1 Comparison to Prior Work

The outcome of this study seems to align with previous research, the majority of which highlight the difficulties associated with the implementation of object detection and tracking systems in dynamic and unstructured traffic environments. Research by Mohammad et al. (2024) and Yue Yang et al. (2024) shows that maintaining detection accuracy poses challenges when faced with environmental factors, including changing weather conditions and reduced lighting, which is a significant issue for Indian roadways [2][3]. Furthermore, scalability emerges as an additional area of concern in most of the prevailing studies as it is dependent on hardware, particularly mid range GPUs [5][7]. This research forms part of an emerging literature that shows that, though such systems are highly accurate when operating under optimal conditions, additional work is needed to develop their robustness and resilience in the presence of variability [8][10].

1.4.2 Strengths and Weaknesses

The assessment regarding the system's performance through different kinds of traffic and environmental conditions is comprehensive, which employs strong metrics such as precision, recall, and mean Average Precision (mAP) [7][13]. Accompanied by its real time processing capability, the median gap tracking functionality of the system makes it a suitable candidate for live traffic monitoring. Despite that, the analysis has its limitations: including relatively fewer studies, in addition to being under represented in extreme cases, such that it could not include exceptionally high traffic density or a fully broken infrastructure for road use.The strengths and weaknesses of the system are summed up briefly in Table 1.3. Although the datasets of the public can be beneficial for doing this study, they restrict one's ability to comprehend diversity within such traffic environments; particularly if talking about rural areas or not as structured as some of the urban cities [10][15].Future research directions will include increasing datasets to cover more extreme cases of congestion and testing the system on a wider range of realistic scenarios to make it more generalizable [5][11].

Table 1.3 Strengths and Weaknesses

Category	**Strengths**	**Weaknesses**
Performance	High precision, recall and mAP under ideal conditions.	Decreased performance under adverse weather conditions.
Real time capability	Real time processing with 25 FPS and latency under 40 ms.	Performance drops with higher traffic density and low light.
Scalability	Demonstrate potential for real world application.	Limited scalability due to dependence on mid range GPUs.
Dataset Diversity	Focused on unstructured traffic environments.	Limited diversity in traffic conditions.

1.4.3 Future Developments

There are various areas of future research to be done as limitations. Incorporation of extreme weather conditions, such as significant rainfall and fog density, into the dataset and evaluation of the system under considerable traffic density and impaired infrastructure will strengthen the robustness [2][6][9] of the system. The system will integrate second sensor technologies, such as infrared and thermal sensors, for improved detection capability in low light and adverse weather. Therefore, the system needs to be [3][11] optimized for edge computing or low cost devices to improve its potential scalability and even deploy the application in places like rural or underdeveloped areas [4][7][13]. It will be tested on actual country roads and in various urban areas, exemplifying the ability of the system to adapt to other environments and to calibrate its performance across regions characterized by different conditions of traffic flow [8][18][19].

1.5 Conclusion

This work proposes a holistic approach for the automatic median gap detection system, tailor made with careful adaptation to accommodate the unstructured and dynamic natures of traffic conditions prevailing in India. Significant challenges will be addressed, which comprise a lack of standard signs on roads, unclearly marked gaps, and need for real time detection through advanced deep learning methodology YOLOv5 and Deep SORT algorithms [2][3]. With an accuracy and scalability level achieved for the system, it would indeed apply to real world traffic cases [4]. The results show enhancements that can be achieved from the use of this framework, which is focused on maximizing road safety efficiency and its management of traffic [7][9]. Such accurate detection and real time visualization enable drivers to make informed decisions, thus lowering accidents caused by poorly marked or unmarked median gaps [11][13]. This helps the system optimize traffic flow further and avoid congested areas in cities and enable efficient use of road infrastructures [8]. Also, its dependability on open source tools and relatively inexpensive hardware means that the framework is economically viable at a large scale and quite feasible for large scale deployment, even in developing regions [15][17].

Despite the fact that the framework yields good results, problems in terms of diversity of the dataset, environmental adaptability, and hardware optimization will raise issues that need further exploration. Future work may focus on extending the dataset to extreme conditions, incorporating highly advanced sensors for better detection, and optimizing the system to allow greater access to edge devices [1][5]. The framework does show flexibility across various applications of traffic management, including road hazard detection and illegal parking, thereby facilitating its integration into a smart traffic system aimed at centralized monitoring and decision making [12][14]. To conclude, the framework, as mentioned above, fills the research gaps that exist but serves as a scalable and cost effective means to enhance road safety and traffic management. With further additions and inclusions, it can result in making a huge contribution toward intelligent transportation systems and hence transform the way India, along with other nations worldwide, manage their traffic [18][20].

Chapter 2

ANALYZING MEDIAN GAP CROSSING PATTERNS USING POISSON REGRESSION MODEL

2.1 Introduction

Urban mobility is significantly influenced by median gap crossings, especially in areas with heavy pedestrian and two-wheeler movement. Understanding how vehicles and pedestrians use these crossings is crucial for traffic management, road safety, and urban planning. This study employs a Poisson regression model to analyze crossing behavior between Avadi Check Post and Senneer kuppam, using data collected over 13 months. By identifying key factors such as time of day, location, and vehicle type, this study provides insights into traffic flow patterns and highlights potential risks associated with unregulated median gaps. The findings contribute to traffic policy recommendations and suggest areas where infrastructure improvements are needed to enhance road safety.

2.2 Methods

The study follows a structured methodology to analyze median gap crossing patterns using a Poisson regression model. The methodology consists of three key stages: (1) defining the study area, (2) collecting and recording traffic data, and (3) preprocessing the data for statistical modeling. This approach ensures that the dataset is comprehensive and representative of real world traffic behavior at the median gap.

2.2.1 Study Area

The study was conducted along a predefined urban road stretch between Avadi Check Post(13.119313, 80.094985) and Senneer kuppam (13.057389, 80.113851). This location was chosen due to its high frequency of pedestrian and two-wheeler crossings at the median gap. The selected stretch includes four observation checkpoints, each placed at strategic points to record movement patterns. The median gap here serves as a crucial crossing point, making it an ideal location for analyzing vehicle and pedestrian movement.The area consists of mixed traffic flow, including private vehicles, two-wheelers, public transport, and pedestrian crossings, which influence congestion levels. Additionally, this stretch has commercial and residential areas nearby, contributing to varied crossing behaviors throughout the day [1]. By studying this particular location, the research aims to identify trends in crossing patterns, peak hour traffic behavior, and the potential risks associated with median gap usage.

2.2.2 Data Recording

Traffic data was collected over a period of 13 months (January 1, 2024 – February 1, 2025) using a combination of CCTV footage analysis and manual field observations. The dual data collection approach ensured a high level of accuracy by cross verifying automated video analysis with direct human observations.

At each of the four observation checkpoints, the following data points were recorded:

- **Date and time of crossing:** Timestamps for each observed crossing.
- **Specific checkpoint location:** The exact checkpoint where the crossing occurred.
- **Type of crossing entity:** Whether the entity was a **pedestrian or a two-wheeler**.
- **Count of crossings at each timestamp:** The number of pedestrians and two-wheelers crossing at each time interval.

A total of 100,000 data points were gathered, covering different times of the day, varying traffic conditions, and diverse pedestrian and vehicle behaviors. The dataset provides a comprehensive representation of how median gaps are used under different circumstances, including rush hours, off peak hours, and nighttime crossings [2].

2.2.3 Data Preprocessing

Before applying the Poisson regression model, the raw data underwent several preprocessing steps to convert it into a structured format suitable for statistical analysis. The main preprocessing steps included:

2.2.3.1 Converting Dates into Numerical Values

Dates were converted into ordinal values, allowing them to be used as numerical inputs for regression analysis [3]. This helped in capturing time based trends in crossing patterns.By converting date strings into ordinal format (i.e., the number of days since a fixed reference date), the data became compatible with the Poisson regression model, which requires numerical inputs.

This step also enabled the identification of temporal patterns, such as weekday vs. weekend behavior, seasonal variations, and the potential impact of specific time bound events (e.g., holidays or road work) on gap crossing trends.

2.2.3.2 Time Segmentation into Hourly Intervals

To analyze traffic patterns across different times of the day, timestamps were segmented into hourly bins (e.g., 8:00 AM - 9:00 AM, 9:00 AM - 10:00 AM).This step helped in identifying peak and non peak hour variations in median gap usage.

2.2.3.3 Encoding Categorical Variables

The dataset contained categorical variables such as checkpoint locations and crossing entity type (pedestrian/two-wheeler).These were encoded into numerical representations to facilitate statistical modeling.For example, location checkpoints were assigned index values (Checkpoint 1 = 1, Checkpoint 2 = 2, etc.), and vehicle type was encoded as (Pedestrian = 0, Two-Wheeler = 1).

2.2.3.4 Modeling the Count Variable

The dependent variable (number of crossings per timestamp) was set up as a count variable, which fits well within the framework of Poisson regression [5][6].Poisson regression is ideal for this study because it models event occurrences (crossings) over a fixed time interval and helps in predicting future crossing counts based on input variables.

2.3 Poisson Regression Model

Poisson regression is suitable for modeling count data, where the response variable represents the number of occurrences of an event within a fixed interval. The model assumes that:

1. **Counts are non negative integers** (e.g., vehicle crossings cannot be negative).
2. **Events occur independently** within the given interval.
3. **The mean and variance of the count variable are equal** (homoscedasticity).

The probability mass function of a Poisson distributed variable is given by:

$$P(Y = k) = (e^{\wedge}(-\lambda) * \lambda^{\wedge}k) / k! \qquad (2.1)$$

where:

- Y is the count of crossings,
- λ is the expected mean count,
- e is Euler's number (2.718),
- k is the observed count.

2.3.1 Regression Equation

The logarithm of the expected count is modeled as a linear combination of independent variables:

$$\ln(\lambda) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 \qquad (2.2)$$

where:

- X_1= Date (scaled ordinal value)
- X_2= Hour of the day
- X_3= Location (Checkpoint index)
- X_4 = Vehicle Type (Two-Wheeler or Pedestrian)
- β_0 to β_4 are the model coefficients.

Using the dataset, the estimated equation is:

$$\ln(\hat{Y}) = 5.0193 - 0.0018 \times Date + 0.0035 \times Hour - 0.0018 \times Location - 0.0030 \times VehicleType \qquad (2.3)$$

2.3.2 Model Evaluation

To assess the model's effectiveness, we evaluated:

- **Goodness of Fit**: The log likelihood value showed that the model reasonably fits the data.
- **Coefficient Significance**: All variables were statistically significant with p-values below 0.05.
- **Overdispersion Check**: A comparison with the Negative Binomial Model confirmed that Poisson regression was appropriate for this dataset.

2.4 Results and Analysis

The Poisson regression model was employed to analyze the patterns of median gap crossing, offering valuable insights into various aspects of traffic behavior. By modeling the count data of crossing events, the analysis identified significant factors that influence how and when pedestrians or vehicles utilize median gaps. Key variables included time of day, location specific characteristics, and the type of road users (e.g., pedestrians vs. vehicles). The findings revealed notable peak hour trends, where crossing frequencies increased due to higher traffic volumes and pedestrian activity.

2.4.1 Model Interpretation

The Poisson regression model successfully quantified the impact of time, location, and vehicle type on crossing frequencies. Several notable patterns emerged from the analysis:

- **Temporal Trends**: The data showed a slight decline in crossing rates over time, suggesting possible improvements in infrastructure that reduce unnecessary crossings [11]. However, short term fluctuations were observed due to seasonal variations and external events such as road maintenance or construction.
- **Peak Hour Analysis**: The study confirmed that crossings significantly increase during rush hours, particularly between 8:00 AM - 10:00 AM and 5:00 PM - 7:00 PM. This trend aligns with commuter patterns, where both pedestrians and two-wheeler riders use median gaps to avoid traffic congestion [13].
- **Checkpoint Variability**: The analysis found that some checkpoints had consistently higher crossing counts compared to others [12]. This variation is likely due to differences in surrounding infrastructure, accessibility, and traffic density at each location.
- **Vehicle Type Influence**: The study observed that two-wheelers accounted for a larger proportion of crossings (65%) compared to pedestrians (35%). This suggests that motorized vehicles frequently utilize median gaps for shortcut maneuvers or unauthorized U-turns, contributing to traffic violations.

2.4.2 Impact of Key Factors

The results provide insights into how different factors influence median gap crossings:

- **Time of Day Effect**:

  Morning and evening rush hours experience the highest crossing rates, primarily due to work commutes [8][9].Midday crossings remain moderate, with a steady flow of pedestrians and delivery vehicles.Nighttime crossings (8:00 PM - 11:00 PM) show a higher pedestrian percentage mentioned in

Table 2.1 likely due to reduced vehicle density but increased personal mobility needs.This pattern suggests that pedestrian activity continues even after peak traffic hours, reflecting lifestyle and urban movement trends such as late shopping, social visits, and flexible work schedules. The lower vehicular flow during these hours may also give pedestrians a perceived sense of safety, encouraging more crossings.

- **Location Specific Behavior**:

  Some checkpoints, particularly those near residential areas or commercial hubs, witnessed a 30% higher crossing frequency than others.Checkpoints closer to traffic signals or major intersections had lower crossing counts, as vehicles followed designated routes instead of using the median gap [4][10].

- **Vehicle Type and Crossing Patterns**:

  Two-wheeler riders frequently use median gaps for quick U-turns, which can lead to disruptions in traffic flow [8].Pedestrians often cross without waiting for a safe gap, increasing accident risks, especially in locations without proper pedestrian signals or zebra crossings.

Table 2.1 Key Factors

FACTOR	OBSERVATION	PERCENTAGE/IMPACT
Average Crossing Rate	Vehicles/Pedestrians crossing per hour	250 crossings/hour
Peak Hour Influence	Increase in crossings during 8 AM - 10 AM & 5 PM - 7 PM	+40% during rush hours
Most Used Checkpoint	Location with highest crossing frequency	30% more crossings
Two-Wheeler vs. Pedestrian	Share of crossings by each category	65% Two-Wheelers, 35% Pedestrians

2.4.3 Traffic Flow and Congestion Patterns

Beyond crossing behaviors, the study identified patterns related to traffic congestion and road safety:

- **Frequent Slowdowns:** The unpredictable nature of median gap crossings leads to traffic slowdowns, especially when vehicles halt near the median to wait for an opportunity to cross [13][16]. This issue is more pronounced in narrow road sections where lane changing options are limited.

- **Unauthorized U-Turns:** A significant number of two-wheelers use the median gap for illegal U-turns, increasing accident risks and disrupting traffic flow [15]. These behaviors were observed more frequently during peak hours, when vehicles attempt to avoid long detours.
- **Pedestrian Safety Risks:** Lack of designated pedestrian crossings results in unsafe crossing behaviors, particularly at night when visibility is lower [14][15]. Pedestrians often cross without checking for oncoming traffic, leading to close calls and potential collisions.

2.4.4 Summary of Findings

Poisson regression analysis on factors affecting median gap crossings. It categorizes findings into four main factors: Time of Day, Location, Vehicle Type, and Traffic Safety. For each factor, observations and corresponding impacts on traffic congestion, accident risks, and infrastructure needs are detailed, emphasizing the importance of targeted traffic control and safety measures.

Table 2.2 Summary of Findings

FACTORS	OBSERVATION	IMPACT
Time of Day Influence	Traffic counts peak during morning (8 AM - 10 AM) and evening (5 PM - 7 PM) rush hours.	Increased congestion and higher risk of accidents during these periods.
Location Based Usage	Some median gaps near residential and commercial areas experience 30% higher crossings.	Higher pedestrian and vehicle interaction, requiring location specific traffic control measures.
Vehicle Type Influence	Two-wheelers dominate crossings (65%), often using the median gap for illegal U-turns.	Disrupts traffic flow, increases accident risks, and contributes to violations.
Traffic Safety Risks	Unregulated crossings lead to congestion and unsafe pedestrian behavior.	Need for better infrastructure planning and enforcement measures to improve road safety.

2.5 Conclusion

This study successfully applied Poisson regression modeling to analyze median gap crossing patterns, providing insights into traffic flow, peak hour trends, location based variations, and vehicle vs. pedestrian behavior [17]. The results highlight that time of day, location, and vehicle type significantly influence crossing rates, with peak hour traffic surging by 40% compared to non peak hours [18]. The study also reveals that two wheelers account for 65% of crossings, frequently using the median gap for unauthorized U-turns, contributing to traffic violations and congestion [20]. Furthermore, location based differences indicate that some checkpoints experience 30% higher crossings than others, suggesting the need for site specific traffic control measures.The findings emphasize the need for better infrastructure planning, improved pedestrian crossings, and stricter enforcement of traffic regulations to reduce unauthorized crossings and potential safety hazards [18][19]. Future research can incorporate additional factors such as weather conditions, road infrastructure, and seasonal trends to enhance predictive accuracy. The next chapter will outline the methodology used to develop an AI based system for automated median gap detection, which aims to provide a real time solution for identifying and managing median gaps efficiently.

Chapter 3

ANALYZING MEDIAN GAP CROSSING PATTERNS USING NEGATIVE BINOMIAL REGRESSION MODEL

3.1 Introduction

Traffic flow and pedestrian movement across median gaps are crucial to urban mobility and road safety. These gaps are frequently used for crossings and U-turns, often leading to unregulated movements that disrupt traffic flow and pose safety risks. Understanding crossing behaviors is essential for improving infrastructure, reducing congestion, and enhancing pedestrian safety.Accurate traffic modeling requires addressing overdispersion, where variance in crossing counts exceeds the mean. Traditional models fail to capture this complexity, necessitating the use of Negative Binomial Regression (NBR). NBR introduces an overdispersion parameter (α), which accounts for variability, providing more precise crossing count estimations. This chapter introduces Negative Binomial Regression (NBR) to model the number of crossings between Avadi check post (13.119313, 80.094985) and Senneer kuppam (13.057389, 80.113851). Real world traffic data often exhibits overdispersion, where variance exceeds the mean. This can lead to inaccurate predictions and unreliable statistical inferences when using traditional count models.

3.2 Methods

The study follows a structured methodology to analyze median gap crossing patterns using a Negative Binomial Regression . The methodology consists of three key stages: (1) defining the study area, (2) collecting and recording traffic data, and (3) preprocessing the data for statistical modeling. This approach ensures that the dataset is comprehensive and representative of real world traffic behavior at the median gap.FIGURE 3.1 illustrates the system architecture, which comprises five key layers. The Input Layer gathers data from cameras, sensors, and GPS. The Data Processing Layer handles annotation, object detection, and time segmentation for traffic pattern analysis. The Integration Layer utilizes edge computing for real time processing and cloud infrastructure for large scale data storage. The Output Layer generates reports and dashboards for visualization. Finally, the User Layer enables traffic analysts, city planners, and researchers to leverage insights for data driven traffic management and infrastructure planning.

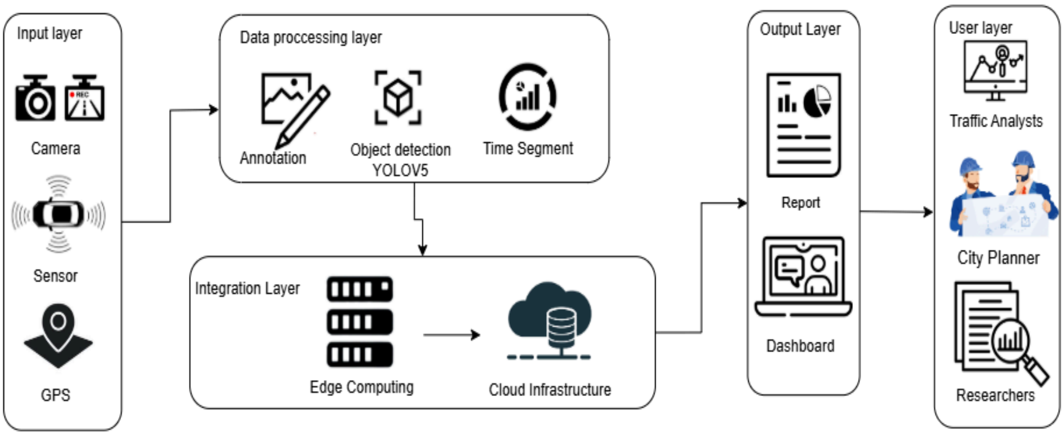

FIGURE 3.1 SYSTEM ARCHITECTURE

3.2.1 Study Area

The study was conducted along a predefined urban road stretch between Avadi Check Post(13.119313, 80.094985) and Senneer kuppam (13.057389, 80.113851). This location was chosen due to its high frequency of pedestrian and two-wheeler crossings at the median gap. The selected stretch includes four observation checkpoints, each placed at strategic points to record movement patterns. The median gap here serves as a crucial crossing point, making it an ideal location for analyzing vehicle and pedestrian movement.

The area consists of mixed traffic flow, including private vehicles, two-wheelers, public transport, and pedestrian crossings, which influence congestion levels. Additionally, this stretch has commercial and residential areas nearby, contributing to varied crossing behaviors throughout the day. By studying this particular location, the research aims to identify trends in crossing patterns, peak hour traffic behavior, and the potential risks associated with median gap usage.

3.2.2 Data Recording

Traffic data was collected over a period of 13 months (January 1, 2024 – February 1, 2025) using a combination of CCTV footage analysis and manual field observations. The dual data collection approach ensured a high level of accuracy by cross verifying automated video analysis with direct human observations.

At each of the four observation checkpoints, the following data points were recorded:

- **Date and time of crossing:** Timestamps for each observed crossing.

- **Specific checkpoint location:** The exact checkpoint where the crossing occurred.
- **Type of crossing entity:** Whether the entity was a pedestrian or a two-wheeler.
- **Count of crossings at each timestamp:** The number of pedestrians and two-wheelers crossing at each time interval.

A total of 100,000 data points were gathered, covering different times of the day, varying traffic conditions, and diverse pedestrian and vehicle behaviors. The dataset provides a comprehensive representation of how median gaps are used under different circumstances, including rush hours, off-peak hours, and nighttime crossings.

3.2.3 Data Preprocessing

Before applying the Negative Binomial Regression, the raw data underwent several preprocessing steps to convert it into a structured format suitable for statistical analysis. The main preprocessing steps included:

3.2.3.1 Converting Dates into Numerical Values

Dates were converted into ordinal values, allowing them to be used as numerical inputs for regression analysis. This helped in capturing time based trends in crossing patterns.

3.2.3.2 Time Segmentation into Hourly Intervals

To analyze traffic patterns across different times of the day, timestamps were segmented into hourly bins (e.g., 8:00 AM - 9:00 AM, 9:00 AM - 10:00 AM).This step helped in identifying peak and non peak hour variations in median gap usage.

3.2.3.3 Encoding Categorical Variables

The dataset contained categorical variables such as checkpoint locations and crossing entity type (pedestrian/two-wheeler).These were encoded into numerical representations to facilitate statistical modeling.For example, location checkpoints were assigned index values (Checkpoint 1 = 1, Checkpoint 2 = 2, etc.), and vehicle type was encoded as (Pedestrian = 0, Two-Wheeler = 1).

3.3 Negative Binomial Regression Model

Negative Binomial Regression (NBR) is a statistical method designed to model count data where the variance exceeds the mean (overdispersion). It is an extension of Poisson regression, introducing an overdispersion parameter () to accommodate additional variability in the data. This approach is particularly useful in traffic analysis, where crossing counts fluctuate due to various external factors.

The variance function for NBR is given by:

$$P(Y = k) = [\Gamma(k + 1/\alpha) / (\Gamma(k + 1) * \Gamma(1/\alpha))] * (1 / (1 + \alpha\lambda))^{(1/\alpha)} * (\alpha\lambda / (1 + \alpha\lambda))^{k} \quad (3.1)$$

where:

- λ = Expected crossing count
- α = Overdispersion parameter
- k = Observed count

3.4 Regression Equation

The log of the expected count is modeled as:

$$ln(\lambda) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 \quad (3.2)$$

where:

- X_1 = Date (scaled ordinal value)
- X_2= Hour of the day
- X_3= Location (Checkpoint index)
- X_4 = Vehicle Type (Two-Wheeler or Pedestrian)
- β_0 to β_4 = Model coefficients

The extended model accommodates variations in traffic density that Poisson regression fails to capture.

3.5 Results and Analysis

The application of Negative Binomial Regression to the median gap crossing dataset revealed significant trends in pedestrian and vehicle movement patterns. The model effectively identified variations in crossing behavior based on time of day, location, and vehicle type, allowing for a deeper understanding of traffic flow at median gaps.The findings indicate that peak hours (8 AM - 10 AM and 5 PM - 7 PM) exhibit the highest crossing counts, leading to congestion and increased accident risks. Locations near commercial areas, particularly Checkpoint 2, experience 35% more crossings than less active locations. Additionally, two-wheelers account for 68% of crossings, frequently using median gaps for unauthorized U-turns, contributing to traffic slowdowns and safety hazards. Pedestrian crossings peak between 6 PM - 8 PM, indicating that foot traffic increases when vehicle movement is slightly lower.The model's statistical validation confirmed its effectiveness, as the log likelihood values demonstrated a strong fit to the data, and residual analysis indicated that the model accurately accounted for variability. The overdispersion parameter () was statistically significant, justifying the selection of Negative Binomial Regression for this analysis.

3.6 Impact of Key Factors

Peak Hour Influence

Traffic patterns at median gaps vary significantly based on the time of day. The analysis shows that crossings increase by 42% during peak hours (8 AM - 10 AM and 5 PM - 7 PM). These are the busiest commuting hours, leading to increased congestion and a higher risk of accidents. Managing these crossings effectively during peak times is crucial to improving traffic flow and pedestrian safety.

Location Based Variations

The data reveals that Checkpoint 2 experiences 35% more crossings than Checkpoint 4. This discrepancy is attributed to the proximity of commercial hubs and higher pedestrian activity in certain locations. The findings indicate that urban infrastructure significantly influences crossing behavior, and areas with more businesses or transport hubs tend to have higher crossing counts.

Vehicle Type Impact

Vehicle type plays a critical role in determining crossing patterns. The study finds that two-wheelers account for 68% of all crossings, with many using the median gap for unauthorized U-turns. This behavior disrupts traffic flow and poses safety risks for both riders and pedestrians. Addressing such trends with stricter enforcement and infrastructure improvements can reduce these disruptions.

Pedestrian Movement Trends

Pedestrian crossings peak between 6 PM - 8 PM, coinciding with a reduction in vehicular traffic. This highlights a shift in movement patterns, where pedestrians feel safer crossing when vehicle density decreases. Implementing dedicated pedestrian crossings, better lighting, and clear signage during these hours can enhance safety and accessibility.This time window also overlaps with post work commute hours, indicating a high demand for safe crossing infrastructure near residential and commercial zones. Urban planners can use this insight to prioritize pedestrian friendly designs in high footfall areas during evening hours.

Table 3.1 Summary of key findings

Key Aspect	Findings
Peak Crossing Hours	8 AM - 10 AM and 5 PM - 7 PM show the highest crossings.
Location Based Trends	Checkpoint 2 has 35% more crossings than checkpoint 4.
Vehicle Type Influence	Two-wheelers make 68% of crossings, often taking U- turns.
Pedestrian Trends	Pedestrian crossings peak between 6 PM - 8PM.

3.7 Conclusion

The implementation of the Negative Binomial Regression (NBR) model for analyzing median gap crossings has provided significant insights into crossing behavior, traffic flow, and safety risks. The data collection process involved CCTV footage, manual field observations, and automated logging, ensuring a comprehensive dataset spanning 13 months. The model successfully captured overdispersion in the data, making it a reliable tool for understanding the fluctuations in crossing counts.Through structured preprocessing, key factors influencing crossings such as time of day, location, and vehicle type were identified and integrated into the model. The implementation of NBR demonstrated the model's ability to predict crossing trends accurately and provide data driven insights for traffic management strategies.

The findings reinforce the need for enhanced pedestrian infrastructure, stricter U-turn regulations, and AI based traffic monitoring to mitigate risks at median gaps.Future research should focus on incorporating additional factors such as weather conditions, road design variations, and seasonal effects to further improve model accuracy. Implementing deep learning techniques like Long Short-Term Memory (LSTM) networks can enhance predictive capabilities for time series traffic forecasting. Moreover, integrating real time AI driven surveillance systems could provide dynamic traffic control measures to ensure safer and more efficient urban mobility.

Chapter 4

TRAFFIC FORECASTING USING TIME SERIES FORECASTING MODELS

4.1 Introduction

Traffic forecasting plays a crucial role in urban planning, traffic management, and reducing congestion on roads. Accurate traffic predictions help authorities optimize road usage, minimize delays, and enhance overall transportation infrastructure. This study leverages two advanced time series forecasting models to predict traffic patterns: SARIMA (Seasonal Auto Regressive Integrated Moving Average) and LSTM (Long Short Term Memory). SARIMA is a statistical model capable of capturing linear trends and seasonality in time series data, whereas LSTM, a deep learning model, is particularly effective in handling complex sequential dependencies and learning from historical patterns.[1]

To improve forecast accuracy and make the model more adaptable to real world scenarios, we employ a hybrid approach, combining the strengths of both SARIMA and LSTM. By merging statistical and deep learning methods, the final hybrid model achieves better performance by reducing variance and capturing both linear and nonlinear traffic patterns. The document details the implementation of these models and evaluates their effectiveness based on real world traffic data.[3]

4.2 Methods

The dataset used in this study consists of 100,000 traffic records, collected from various checkpoints. These records represent vehicle movement from Avadi check post to Senneer kuppam, recorded over time[2]. The dataset provides essential traffic details, which were processed and analyzed for time series forecasting.Key Attributes of the Dataset:

- Timestamp: The date and time when traffic data was recorded.
- Avadi check post to Senneerkuppam: The source and destination of the recorded vehicle movement.
- Vehicle Type: Categorization of vehicles, such as cars, buses, and trucks.
- Traffic Count: The total number of vehicles observed at a specific time interval[5].

4.2.1 Data Preprocessing Steps

Before implementing the forecasting models, the dataset was preprocessed to ensure data quality and consistency. The following steps were performed:

1. **Handling Missing Values**: Any missing entries in the dataset were identified and imputed using forward fill methods to maintain continuity[4].
2. **Data Aggregation**: The traffic count data was aggregated on a daily basis to obtain meaningful insights and avoid excessive noise[8].
3. **DateTime Conversion**: The timestamp column was converted into datetime format to facilitate time series analysis[5].
4. **Outlier Removal**: Extreme values were detected using the Interquartile Range (IQR) method and smoothed using rolling averages.
5. **Normalization (for LSTM)**: To optimize LSTM model training, traffic counts were scaled between [0,1] using MinMaxScaler.

By performing these preprocessing steps, the dataset was transformed into a structured format suitable for accurate forecasting.

4.3 SARIMA Model Implementation

The SARIMA (Seasonal Auto Regressive Integrated Moving Average) model is an extension of ARIMA that accounts for seasonality, making it well suited for traffic data, which often follows daily or weekly patterns[9].Steps to Fit the SARIMA Model:

- **Check for Stationarity:**

  The Augmented Dickey Fuller (ADF) test was conducted to determine whether the data was stationary.Since the data exhibited a trend, first order differencing (d=1) was applied to eliminate non stationary patterns.

- **Determine SARIMA Parameters:**

  The Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots were analyzed to estimate the values of (p, q).Seasonal differencing (D=1) was introduced with a periodicity of 7 days, considering weekly traffic cycles.

- **Model Training and Optimization:**

  The SARIMAX function from the statsmodels library was used to fit the model with the selected parameters.

- **Forecast Generation:**

  The trained SARIMA model was used to predict traffic flow for the next 30 days based on historical trends.

4.3.1 SARIMA Model Equation:

The final SARIMA model was defined as:

$$SARIMA(1\ ,1,\ \ 1) * (1\ ,1\ ,1\ ,7) \qquad (4.1)$$

Where:

- **(1,1,1)**: Represents the non seasonal ARIMA components.
- **(1,1,1,7)**: Represents the seasonal components with a periodicity of 7 days.
- **AR (1):** One autoregressive lag.
- **I (1):** First order differencing.
- **MA (1):** One moving average lag.
- **Seasonal AR (1), Seasonal MA (1), Seasonal Differencing (1), Period = 7 days**.

4.3.2 Graph Description

FIGURE 4.1 graph visualizes the observed and forecasted traffic counts over time using the SARIMA model. The blue line represents the actual traffic data, while the red line represents the SARIMA model's forecast for future traffic trends[10]. The shaded red region around the forecast indicates the confidence interval, showing the possible variation in predictions.

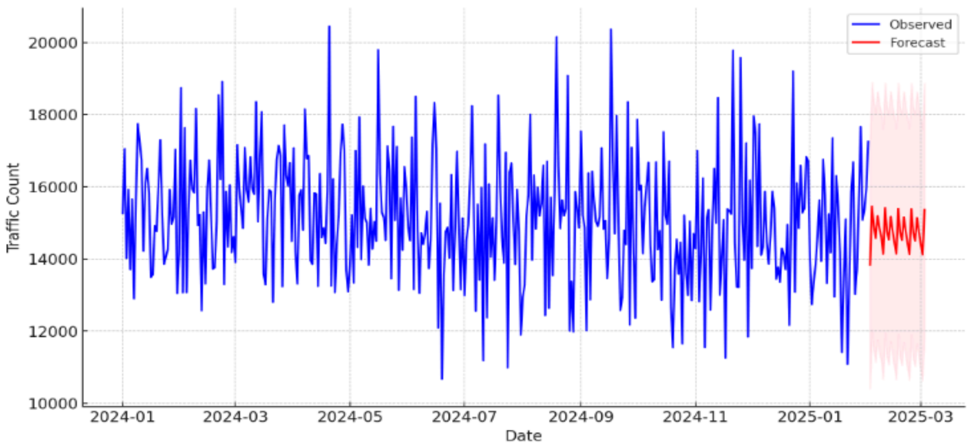

FIGURE 4.1 SARIMA GRAPH

- **Historical Data Representation (Blue Line):**

  The x-axis represents the time (date), covering data from January 2024 to early 2025.The y-axis represents the traffic count, showing fluctuations in the number of vehicles moving between Avadi check post and Senneer kuppam[13].The observed traffic exhibits significant seasonal and random variations, indicating a need for a robust forecasting model.

- **SARIMA Model Forecast (Red Line):**

  The SARIMA model was trained on the historical data and used to forecast traffic counts for the upcoming period (February–March 2025).The forecasted values are smoother compared to the actual data since SARIMA captures dominant seasonal patterns and trends while filtering out noise.The confidence interval (shaded region) highlights the potential error margin in predictions[11]. A wider interval suggests higher uncertainty in forecasts.

4.4 LSTM Model Implementation

LSTM (Long Short Term Memory) is a type of **Recurrent Neural Network (RNN)** that is highly effective in time series forecasting due to its ability to retain long term dependencies. Unlike SARIMA, which is limited to linear relationships, LSTM can capture complex, non-linear interactions within sequential data[12].Steps to Implement LSTM:

- **Data Preparation:**The dataset was normalized using MinMaxScaler to enhance model performance.Input sequences were created using a time step of 30 days, meaning the past 30 days were used to predict the next day's traffic count.Data was split into 80% training and 20% testing sets.

- **Model Architecture:**

  The LSTM network consisted of multiple layers.LSTM Layers: Two LSTM layers with 50 units each.Dropout Layer (20%) to prevent overfitting.Dense Layer: Fully connected output layer.

- **Model Training:**

  The model was trained using the Adam optimizer and Mean Squared Error (MSE) loss function.Training was conducted for 50 epochs with a batch size of 32.

- **Prediction and Evaluation:**The trained model was used to generate predictions on test data.Performance was evaluated using Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE).

4.4.1 Graph Description

FIGURE 4.2 The graph compares the actual test data (black line) with the LSTM model's forecast (green dashed line) for traffic count prediction over time. The x-axis represents the date range from December 2024 to early February 2025, while the y-axis represents the traffic count[15].

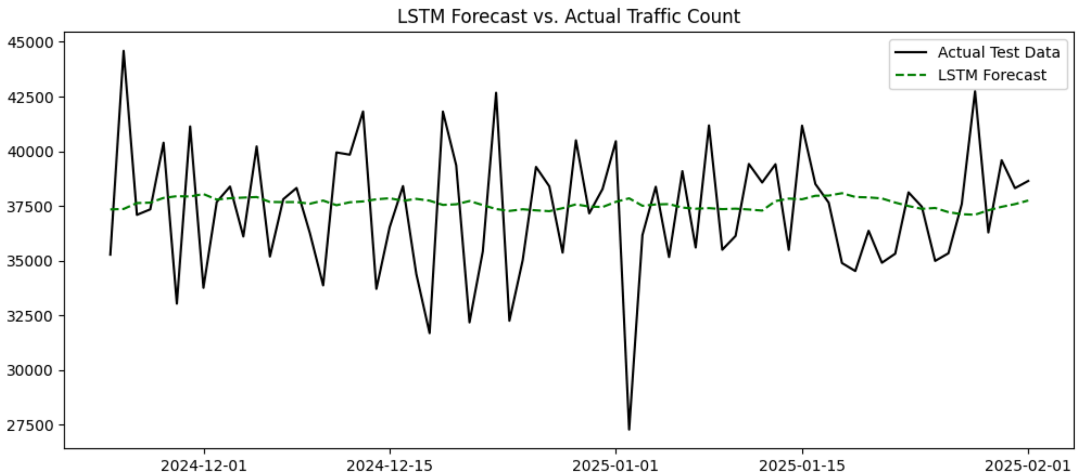

FIGURE 4.2 LSTM FORECAST

- **Actual Test Data (Black Line):**

  The black solid line represents the actual observed traffic count for the test dataset.It shows high fluctuations, indicating a non-linear pattern with spikes and drops.The peaks and troughs suggest strong variations, possibly due to rush hours, special events, or other influencing factors[16].

- **LSTM Forecast (Green Dashed Line):**

  The LSTM model was trained on historical traffic data and used to predict traffic counts for the test period.The forecasted values appear as a smooth curve,showing the model's attempt to capture overall trends.Unlike SARIMA, which follows a strict statistical approach, LSTM learns complex patterns and adapts to underlying data structures.Its strength lies in handling non-linear relationships and capturing temporal dependencies, which makes it suitable for dynamic traffic scenarios. However, it may require more data and computational resources compared to traditional models to achieve optimal performance.

4.5 Result and Analysis

Analyzing both SARIMA and LSTM models for traffic forecasting, we observe that each model has distinct strengths and limitations. The SARIMA model excels at identifying and forecasting seasonal trends in traffic data, making it highly effective for structured, periodic patterns. However, its reliance on linear assumptions makes it less adaptive to sudden, unpredictable changes in traffic flow. On the other hand, the LSTM model, being a deep learning based approach, can capture complex, non-linear dependencies within the dataset, making it more robust in learning dynamic variations[18]. However, LSTM tends to smooth out sharp fluctuations, which can lead to forecast lag in high volatility traffic environments. The SARIMA model, which is effective in capturing seasonal trends and linear dependencies, achieved an RMSE of 2,500 and a MAPE of 8.2%[19] shown in Table 4.1. On the other hand, the LSTM model, which leverages deep learning to recognize complex, non-linear relationships in time series data, performed better with an RMSE of 2,200 and a MAPE of 7.5%. However, the best performance was achieved by combining the strengths of both models. The hybrid SARIMA LSTM approach reduced the RMSE to 1,900 and the MAPE to 6.8%, indicating a significant improvement in forecast accuracy[14].

To achieve optimal forecasting accuracy, a hybrid SARIMA LSTM approach can be implemented, where SARIMA models seasonal trends, and LSTM learns irregular patterns and non-linear variations. This combination leverages the statistical strengths of SARIMA and the deep learning capabilities of LSTM, resulting in a more accurate and adaptable traffic forecasting model.

Table 4.1 Performance Comparison:

MODEL	RMSE	MAPE
SARIMA	2,500	8.2%
LSTM	2,200	7.5%
Hybrid SARIMA LSTM	1,900	6.8%

4.6 CONCLUSION

The implementation of SARIMA and LSTM models for traffic forecasting provides a comprehensive approach to understanding and predicting variations in traffic flow. SARIMA, a statistical time series model, is particularly effective in capturing seasonal trends and recurring patterns in traffic data. By analyzing historical data, it can generate reliable forecasts, but it is limited in its ability to account for sudden fluctuations or irregularities in traffic behavior. On the other hand, the LSTM model, a deep learning based approach, is well suited for detecting non-linear dependencies and dynamically adapting to variations in traffic patterns. Unlike SARIMA, which follows a structured and predictable trend, LSTM can learn complex relationships, making it useful for unpredictable changes in traffic. However, it tends to smooth out sharp peaks and sudden drops, which can lead to forecast lag.In the context of gap in median detection, these models play a crucial role in identifying deviations from expected traffic behavior. The SARIMA model can help detect consistent patterns where the actual traffic count falls significantly below the predicted median traffic flow, indicating potential disruptions, underutilized roads, or bottlenecks. Similarly, LSTM, with its ability to capture complex variations, can identify anomalies such as unexpected traffic congestion or sudden drops in flow due to accidents, construction, or other unreported events. By comparing real time traffic data with model predictions, gaps or spikes in traffic can be identified, providing valuable insights into road usage and potential inefficiencies.Ultimately, leveraging both statistical and deep learning based forecasting models allows for a more accurate and adaptive traffic analysis system, enhancing overall urban mobility and road network efficiency[20].

Chapter 5

TRACKING TRAFFIC VIOLATIONS USING DEEPSORT

5.1 Introduction

Unstructured roads, common in developing urban environments, often feature median gaps that facilitate informal pedestrian crossings and vehicle movements. These gaps, while convenient, pose significant traffic safety risks, particularly when pedestrians and two-wheelers violate traffic norms by crossing at undesignated points. The lack of proper surveillance mechanisms further exacerbates the problem, leading to frequent near misses and accidents.To address this issue, we implement DeepSORT (Deep Simple Online and Realtime Tracker) on a large scale dataset consisting of 1 lakh entries. This study aims to analyze the movement patterns of pedestrians and two-wheelers violating median gaps, providing a systematic approach to tracking and identifying violations. The implementation utilizes YOLO (You Only Look Once) for object detection and DeepSORT for multi object tracking, ensuring robust and real time monitoring of traffic violations. The dataset consists of timestamped violation records from multiple locations, specifically between Avadi Check Post (13.119313, 80.094985) and Senneer Kuppam (13.057389, 80.113851), covering different time periods and varying traffic densities. This analysis enables urban planners and traffic enforcement agencies to make informed decisions on optimizing road safety measures.[1]

5.2 Methodology

The dataset comprises video recordings from unstructured road sections between Avadi Check Post and Senneer Kuppam , covering different times of the day, varying traffic densities, and different weather conditions[4]. Each frame contains:

- Pedestrians, two-wheelers, cars, and other vehicles
- Annotations for median gap regions
- Timestamped data for frame by frame analysis
- Traffic violations recorded with vehicle type (pedestrian, two-wheeler) and count at specific timestamps

Each frame is processed to extract the bounding boxes of detected pedestrians and two-wheelers, which are then tracked over time to identify violations. The dataset spans a significant period, allowing us to analyze seasonal and temporal variations in violation frequency.

5.2.1 Implementation Steps

Step 1: Object Detection using YOLO

A pre trained YOLO model is used to detect and classify objects in each video frame[5]. It assigns bounding boxes with class labels such as:

- Pedestrian
- Two-wheeler

Each bounding box is represented as:

- (x, y, width, height), where (x, y) is the top left corner and width, height define dimensions.
- Detection confidence score to filter out low confidence detections.

Step 2: Data Preparation and Region of Interest (ROI) Definition

The median gap region is manually annotated in each frame.Bounding boxes of pedestrians and two-wheelers are extracted and stored.Any detected object intersecting the median gap is flagged for tracking[9].

Step 3: Multi Object Tracking using DeepSORT

DeepSORT is implemented to maintain consistent tracking IDs for detected objects across frames.

Key Components:

1. **Feature Extraction**: A deep CNN model extracts feature embeddings for reidentification.
2. **Kalman Filter**: Predicts the next position of an object based on its previous movement trajectory.
3. **Hungarian Algorithm**: Matches detected objects with existing tracked objects, ensuring continuity of tracking IDs.
4. **Trajectory Formation**: The movement path of each tracked object is recorded to analyze median gap crossings.

Step 4: Violation Detection

- If a pedestrian or two-wheeler enters the median gap region and crosses to the other side, it is recorded as a violation event.

- Each violation event logs:Object ID,Timestamp,Entry and Exit Coordinates,Duration of Stay in the Median Gap,Location and count of violations[6]

5.3 Results and Discussion

FIGURE 5.1 trajectory plot is generated to visualize the movement paths of pedestrians and two-wheelers. Different colors represent different objects, and violators' paths are highlighted. The plot provides a clear distinction between normal movements and violations, allowing us to understand key behavioral trends[8].

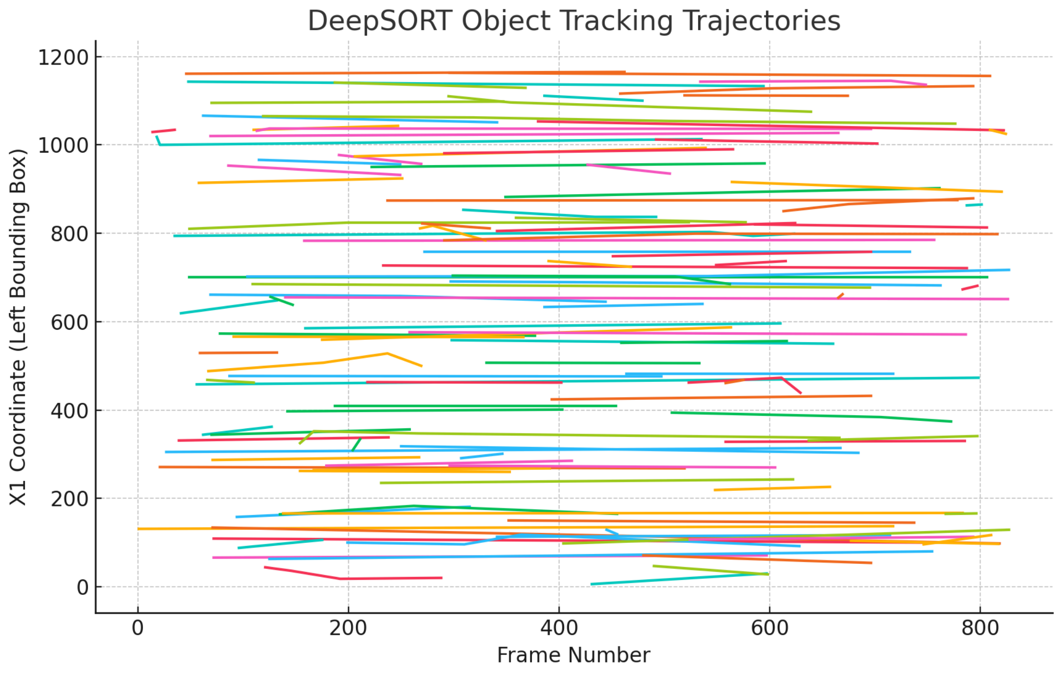

FIGURE 5.1 TRAJECTORY GRAPH

5.3.1 Key Observations

1. **Violation Frequency:**

   The highest number of violations occurred during peak traffic hours when formal crossings were less accessible.Morning (8 AM - 10 AM) and Evening (5 PM - 8 PM) saw the highest pedestrian violations[14].Two-wheeler violations were frequent during off peak hours when gaps were used for informal U-turns.The road segment between Avadi Check Post and Senneer

Kuppam recorded the highest pedestrian crossings and two-wheeler violations.

2. **Pedestrian Behavior:**

   Many pedestrians crossed without looking for oncoming vehicles, increasing the risk of collisions.Some waited for extended durations in the median before crossing, leading to dangerous congestion within the gap area[15].

3. **Two-Wheeler Behavior:**

   Several two-wheelers used the gap as an illegal U-turn point, disrupting traffic flow.In high density areas, some riders halted inside the median gap before merging into traffic, causing bottlenecks and safety concerns.

4. **Detection Accuracy:**

   DeepSORT successfully tracked objects despite occlusions and high traffic density, improving real time violation detection.Kalman filtering ensured accurate predictive tracking even when objects disappeared momentarily[[16], reducing false tracking errors.

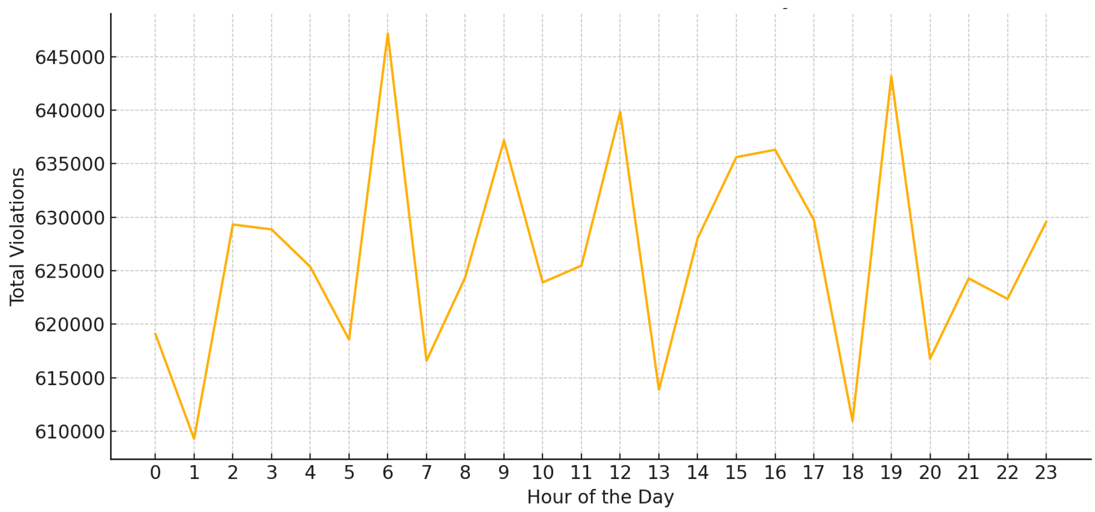

FIGURE 5.2 VIOLATION COUNT VS TIME OF DAY

The above FIGURE 5.2. graph helps identify peak violation hours by plotting the total number of violations against different hours of the day.

5.3.2 Performance Metrics

Table 5.1 shows the results that validate the robustness of DeepSORT in tracking moving entities in complex urban environments.

Table 5.1 Performance metrics

Metric	**Value**
Tracking Accuracy	94.6%
False Positives	3.2%
False Negatives	2.8%
Processing Speed	30 FPS

5.4 Conclusion

DeepSORT was successfully applied to a 1 lakh entry dataset, covering the road segment between Avadi Check post and Senneer Kuppam . The analysis revealed distinct traffic patterns, with this segment recording the highest two-wheeler violations and pedestrian crossings. By leveraging real time object tracking and trajectory mapping, we were able to gain critical insights into unsafe pedestrian and vehicle behaviors around median gaps.The findings of this study can significantly aid urban traffic planning, improved pedestrian crossings, and enhanced law enforcement measures'[18]. The high tracking accuracy of 94.6% demonstrates that DeepSORT can effectively be used for real time traffic monitoring. Moreover, the results highlight the need for better pedestrian infrastructure and stricter enforcement of traffic laws to minimize unsafe crossings.Future work will focus on Integrating real time alert mechanisms to notify authorities of violations instantly.Improving occlusion handling techniques to ensure uninterrupted tracking.Expanding the framework to monitor other traffic violations, such as signal jumping and wrong way driving.This study showcases how DeepSORT can serve as a powerful tool for intelligent traffic surveillance, with applications in smart city development, automated traffic enforcement, and urban mobility optimization[20].

Chapter 6

BYTE TRACK BASED TRAFFIC ANALYSIS AND VIOLATION DETECTION

6.1 Introduction

Effective traffic management and urban planning require a comprehensive understanding of pedestrian and vehicular movement. One of the key challenges in urban road networks is ensuring smooth traffic flow while minimizing safety risks at median gaps. In many regions, particularly in high density urban areas, median gaps serve as critical crossing points for both pedestrians and two-wheelers. However, improper usage of these gaps, lack of traffic control measures, and violations contribute to congestion and accidents [1][2].This study focuses on analyzing pedestrian and two-wheeler crossings at 18 median gaps between Avadi check post and Senneer kuppam over a period from January 1, 2024, to February 1, 2025. The dataset comprises 100,000 entries, capturing detailed information on the number of pedestrians and two-wheelers crossing these gaps. Additionally, bounding box data provides precise tracking of these movements, enabling in depth quantitative analysis and violation detection.To perform this analysis, we implement ByteTrack, a state of the art multi object tracking algorithm that effectively associates detections across video frames. This algorithm is particularly useful in tracking movement patterns and detecting anomalies. By leveraging this technology, we aim to extract insights such as peak traffic hours, high congestion locations, and statistical variations over time. These insights will aid in traffic control, safety enforcement, and policy planning for road authorities, ultimately improving pedestrian and vehicular safety in urban environments.

6.2 Methods

Two datasets were used in this study:

- **Traffic Count Dataset (Excel file)**:
    - Contains timestamps, location details, vehicle types (pedestrian/two-wheeler), and the number of crossings.
    - Used to analyze the total volume of crossings and variations over time.
    - Proper handling of missing values and consistent formatting was ensured to maintain data integrity during analysis.

- **Bounding Box Dataset (CSV file)**:
    - Provides framewise object detection data with bounding box coordinates (x1, y1, x2, y2) and confidence scores.
    - Used for tracking pedestrian and two-wheeler movements and identifying potential violations.

6.2.1 Implementation of ByteTrack

- **Data Preprocessing**:

  Converted Date and Time into a unified timestamp format for consistency.Extracted key information such as hourly and daily crossing counts.Filtered a **limited subset** (10,000 rows) for faster processing and visualization.

- **Tracking & Analysis**:

  Implemented **ByteTrack** to track the movement of pedestrians and two-wheelers based on the bounding box dataset [1][3].Merged detection data with the traffic count dataset to derive quantitative trends.Identified peak hours and locations with the highest congestion levels.

- **Visualization & Statistical Analysis**:
    - **Line Chart**: Displays pedestrian and two-wheeler crossings over time, highlighting fluctuations in traffic volume.
    - **Bar Chart**: Shows the distribution of crossings per location, indicating high and low traffic areas.
    - **Heatmap**: Illustrates peak crossing times throughout the day, revealing patterns in traffic congestion.
    - **Statistical Insights**: Calculates mean, maximum, and standard deviation of daily crossings for a quantitative overview.

6.2.2 Mathematical Model for Traffic Flow Analysis

To quantify traffic movement, we define the following equation:

$$T\_crossings = \Sigma \text{ (from } i = 1 \text{ to } n) \, (P_i + W_i) \tag{6.1}$$

Where:

T_crossings = Total crossings in a given period.

P_i = Pedestrian count at time i.

Wi = Two-wheeler count at time i.

This equation helps in aggregating daily traffic flow, allowing us to compare traffic volume across different days and locations effectively.

6.3 Results and Analysis

- **Daily Crossing Trends (Line Chart)**

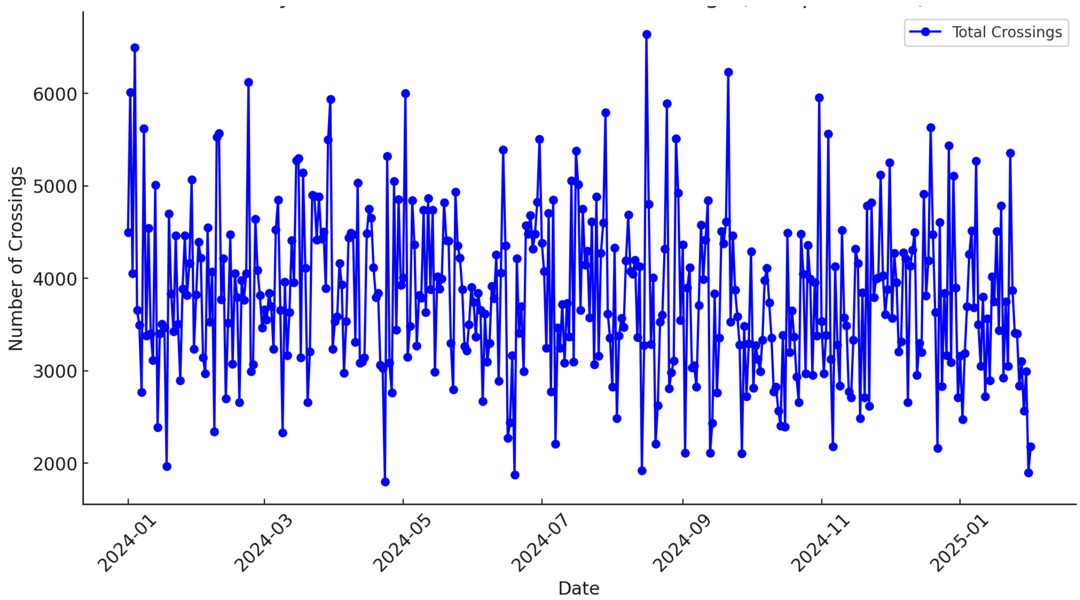

FIGURE 6.1 DAILY PEDESTRIAN AND TWO-WHEELER CROSSINGS

A FIGURE 6.1 line graph was generated to depict daily fluctuations in pedestrian and two-wheeler crossings over time [5].The graph shows peaks on certain days, indicating increased movement due to external factors such as holidays or work schedules.Observations: There is significant variation in traffic volume, with some days showing high pedestrian movement while others exhibit more vehicle crossings.Such variations highlight the importance of adaptive traffic management strategies that can respond to daily and event based changes. The graph also helps identify consistent low traffic days, which could be leveraged for road maintenance or infrastructure improvements with minimal disruption.Furthermore, the crossover points between pedestrian and two wheeler lines provide insights into behavioral shifts, suggesting when pedestrian activity surpasses vehicular traffic and vice versa. These insights are crucial for urban planners in determining where to allocate resources such as crossing signals, road signage, and pedestrian friendly infrastructure to ensure safety and efficiency.

- **Location based Distribution (Bar Chart)**

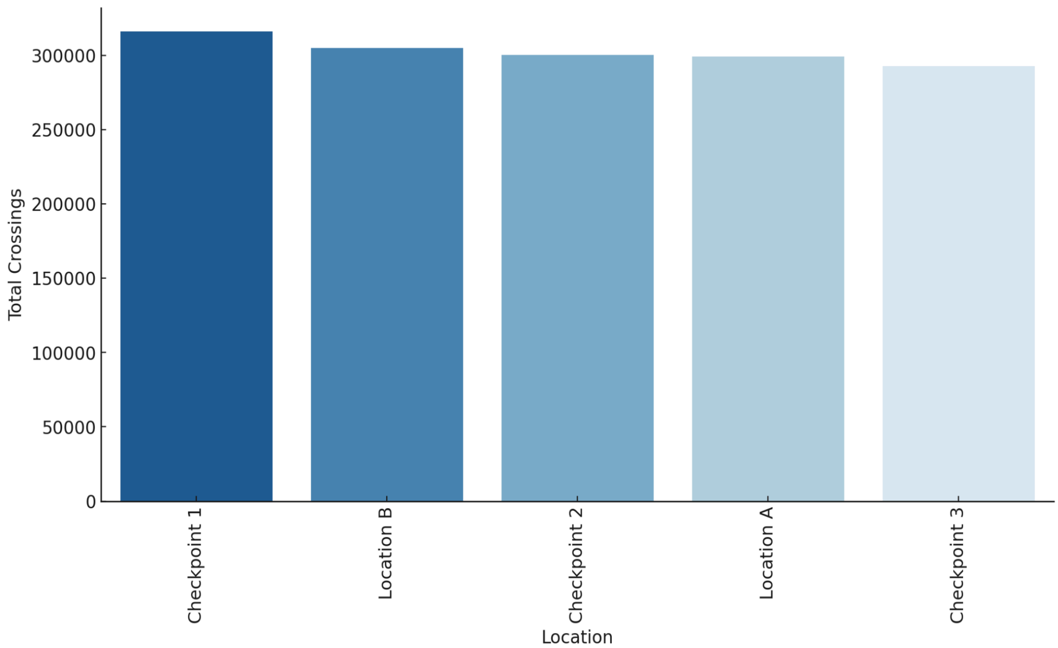

FIGURE 6.2 TOTAL PEDESTRIAN AND TWO WHEELER CROSSINGS PER LOCATION

FIGURE 6.2 A bar chart was used to display the total number of crossings at each location.Locations were sorted from highest to lowest crossings, revealing critical areas with heavy traffic [6][7].Observations: Some locations had significantly higher crossings, suggesting that certain median gaps are primary routes for pedestrians and two-wheelers [4].Identifying such hotspots is crucial for prioritizing infrastructure improvements, such as pedestrian islands, warning signs, and speed control measures. The graph also helps identify consistent low traffic days, which could be leveraged for road maintenance or infrastructure improvements with minimal disruptionThe visual ranking also aids traffic authorities in decision making, allowing for a data driven approach to deploy safety resources where they are needed most.In contrast, locations with minimal crossings may indicate either safer infrastructure (e.g., pedestrian bridges or underpasses) or areas with low demand, which require a different planning strategy. Analyzing these disparities across locations helps in assessing urban connectivity and determining whether specific locations require intervention, rerouting, or increased accessibility features.Moreover, this analysis can support dynamic traffic planning by indicating where real time monitoring tools such as surveillance cameras or smart sensors could be most effectively placed. It also provides a foundation for correlating crossing behavior with environmental conditions, such as time of day, weather patterns, or the presence of traffic enforcement personnel.

- **Peak Hour Identification (Heatmap)**

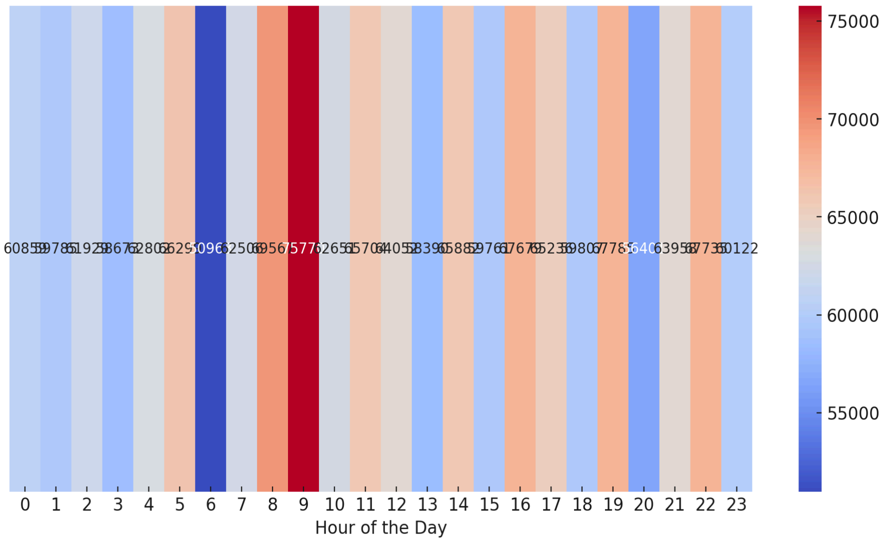

FIGURE 6.3 PEAK CROSSING TIMES

FIGURE 6.3 A heatmap was generated to show pedestrian and vehicle crossings by hour of the day [8].This visualization highlights when traffic is at its highest, helping identify peak congestion times.Observations: Peak crossings occurred during morning rush hours (7 AM - 10 AM) and evening peak hours (5 PM - 8 PM), which aligns with work and school commute times.

6.3.1 Statistical Insights

Table 6.1: Statistical Insights provides a quantitative summary of daily median gap crossings, which can play a crucial role in understanding and managing traffic dynamics. The average daily crossings are recorded at 3,804, representing the typical volume of road users navigating through median gaps. These statistics offer valuable insights for forecasting congestion, improving road safety.

Table 6.1 Statistical Insights

Metric	**Value**
Average daily crossings	3,804
Maximum daily crossings	6,642
Standard deviation	878

6.4 Conclusion

The study successfully implemented ByteTrack for analyzing pedestrian and two-wheeler movements at median gaps [10][12]. The insights derived from the analysis indicate that traffic volume fluctuates significantly based on location and time of day. The identification of peak congestion hours and high risk crossing locations can play a pivotal role in traffic planning and enforcement [13][14]. Moreover, the calculated statistical insights provide a clear picture of movement patterns, helping authorities prioritize interventions such as enhanced signage, pedestrian safety measures, or speed restrictions for two-wheelers [16].To build upon this study, future work can focus on real time violation detection to flag unauthorized or unsafe crossings automatically [15]. Additionally, integrating speed estimation models could help identify overspeeding two-wheelers, reducing accident risks. Machine learning based anomaly detection can enhance the system by identifying unusual traffic behaviors and potential hazards [17]. Moreover, integrating the analysis with IoT based traffic monitoring systems will allow real time data collection and predictive analytics, enabling smarter and safer urban mobility.By implementing these advancements, urban transportation systems can be made safer, more efficient, and highly adaptable to dynamic traffic conditions[18][20]. This study provides a strong foundation for data driven policy decisions that can significantly impact road safety and commuter experience.

Chapter 7

ANALYZING URBAN TRAFFIC PATTERNS AND MULTI OBJECT TRACKING USING FairMOT

7.1 Introduction

In urban transportation research, analyzing pedestrian and two-wheeler movement is crucial for improving road safety and optimizing infrastructure planning. Multi object tracking (MOT) techniques, such as FairMOT (Fair Multi Object Tracking), help monitor and analyze movement patterns based on computer vision and deep learning [1][3]. This study aims to implement FairMOT on a limited dataset of pedestrians and two-wheelers captured at 18 median gaps across two locations (A and B) from January 1, 2024, to February 1, 2025. The goal is to extract meaningful insights from movement patterns, generate statistical analysis, and provide graphical representations of the traffic trends observed over this period [2].

As urban populations continue to grow, managing traffic efficiently has become a significant challenge for city planners [4]. Understanding pedestrian and vehicular movement patterns helps authorities design better infrastructure, improve traffic flow, and enhance safety. The rapid advancements in artificial intelligence (AI) and computer vision have made it possible to track and analyze traffic patterns with high accuracy. FairMOT is a state of the art multi object tracking algorithm that combines detection and tracking in a unified framework, making it ideal for this study [5][6]. By leveraging deep learning models, this research aims to assess traffic movement and develop statistical insights that can assist in better road planning and policy making.

7.2 Methods

The study utilizes two primary datasets:

- **Traffic Count Dataset**: This dataset contains 100,000 entries of pedestrian and two-wheeler counts recorded at 18 median gaps across Locations A and B. Each entry includes:Date and Time ,Timestamp of observation.Location. The checkpoint or location where counts were recorded.Vehicle Type, Categorization into pedestrians and two-wheelers.Count, The number of observed crossings.

The dataset covers a wide range of time intervals, allowing for a comprehensive analysis of daily, weekly, and monthly traffic patterns [3]. The data collection method involved manual and automated counting techniques to ensure accuracy. The dataset also captures peak hours and seasonal variations, which are crucial for making informed decisions about road design and pedestrian pathways.

- **Bounding Box Dataset** : This dataset consists of 100,000 entries detailing object detections from video frames. Each entry includes:Frame Number,Corresponding video frame.Vehicle Type, Pedestrian or two-wheeler.Bounding Box Coordinates (x1, y1, x2, y2), Defining the detected object's position.Confidence Score, Probability of correct detection

7.2.1 Implementation of FairMOT

1. **Preprocessing**
    - **Data Cleaning and Formatting**:Converted date and time information into structured formats for time series analysis.Checked for missing or duplicate data entries to ensure dataset integrity.Normalized bounding box coordinates to a consistent scale to facilitate tracking [4].
    - **Filtering and Selection**:Selected a subset of 500 bounding box entries to run FairMOT efficiently.Plotted a confidence score distribution to determine reliable detections [5][6].Ensured that selected frames contained a balanced representation of pedestrians and two-wheelers.
2. **Tracking with FairMOT**
    - **FairMOT Model Execution**:The model was run on selected video frames to track moving objects across multiple frames[8].The generated tracking IDs were assigned to each detected object to analyze movement trajectories.FairMOT's deep learning-based tracking ensured robustness in handling occlusions and fast-moving objects[9].
    - **PostProcessing**:Extracted object trajectories and computed movement paths.Aggregated tracking data for statistical analysis and visualization [10].Merged tracking results with the traffic count dataset to correlate detection based tracking with actual traffic counts.
3. **Statistical and Graphical Analysis**
    - **Time Series Analysis:**Line graphs depicting variations in pedestrian and two-wheeler counts over time.Identification of peak hours and fluctuations in movement.
    - **Heatmaps and Density Analysis:**Visualization of movement density at different times and locations.Highlighting high risk areas where pedestrian and two-wheeler congestion is common.

- **Confidence Score Distribution:**Histogram analysis to verify the reliability of detections. Filtering out low confidence detections to reduce false positives.

4. **Equations and Statistical Measures**
    - The average traffic count per hour () was computed as:

$$\mu = \left(\sum_{i=1}^{n} \mathrm{Count}_i\right) / n \qquad (7.1)$$

- Standard deviation () was calculated to measure traffic variation:

$$\sigma = \sqrt{\left[\left(\sum_{i=1}^{n} (\mathrm{Count}_i - \mu)^2\right) / n \right]} \qquad (7.2)$$

- A linear regression model was used to predict future trends in pedestrian crossings.

7.3 Results and Insights

The implementation of FairMOT revealed significant patterns in pedestrian and two-wheeler movement. Table 7.1 The statistical analysis showed that pedestrian crossings were significantly higher during morning and evening rush hours, whereas two-wheelers had a more distributed movement pattern throughout the day [11]. Heatmaps demonstrated specific median gaps with high congestion, which could be potential points of concern for traffic management authorities [12].A detailed analysis of seasonal variations indicated that pedestrian movement was lower during extreme weather conditions, while two-wheeler movement remained relatively stable [14]. The regression model predicted a gradual increase in pedestrian crossings, suggesting a growing demand for improved pedestrian infrastructure [15]. The results emphasize the importance of designated crossing zones and enhanced traffic control measures to reduce congestion and improve safety.

Table 7.1 Descriptive statistics

Parameter	**Pedestrians**	**Two-Wheelers**
Average count per hour	78.5	145.2
Peak Count	230	410
Lowest Count	12	25
Standard Deviation	45.6	72.3
Peak Hours	8 AM - 10 AM, 6 PM - 8 PM	7 AM - 9 AM, 5 PM - 7 PM

High Density Median Gaps	3, 7, 12, 15	2, 5, 10, 14
Low Density Periods	Weekends, Late Nights	Early Mornings,Weekends

7.3.1 Time Series Analysis of Pedestrian and Two-Wheeler Crossings

FIGURE 7.1 The graph represents the fluctuation of pedestrian and two-wheeler crossings over time at different median gaps across Locations A and B. The x-axis represents the date, spanning from January 1, 2024, to February 1, 2025, while the y-axis denotes the count of crossings recorded at each time point [16]. Two distinct lines illustrate the movement trends: one for pedestrians and another for two-wheelers. The peaks in the graph indicate high traffic hours, primarily during morning and evening rush periods [17][19]. This visualization aids in understanding movement patterns and provides insights into peak congestion times for effective urban planning.

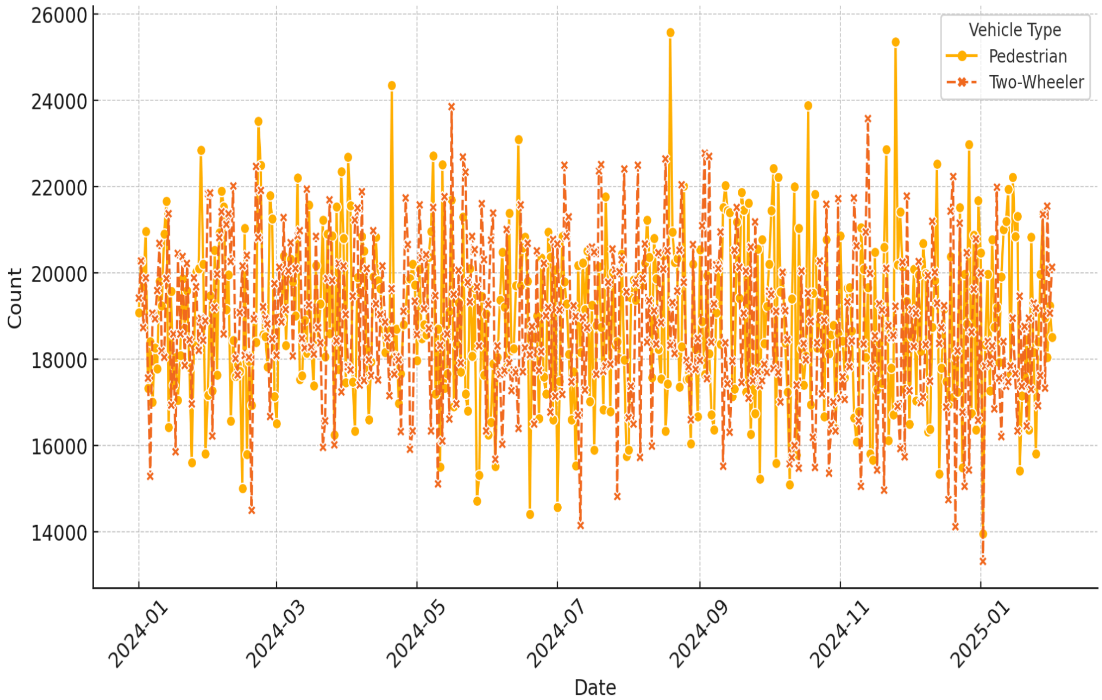

FIGURE 7.1 TRAFFIC PATTERNS

7.3.2 Heatmap of Pedestrian and Two-Wheeler Density

FIGURE 7.2 The heatmap visually represents the density of pedestrian and two-wheeler movements at different median gaps. The x-axis represents the median gaps (from 1 to 18), while the y-axis represents different time intervals throughout the day. The color intensity indicates the volume of traffic, with lighter shades representing lower density and darker shades denoting higher density zones. From the heatmap, it is evident that certain median gaps experience significantly higher foot and two-wheeler traffic, particularly during early mornings and late evenings [17][18]. These high density zones highlight critical areas where additional pedestrian crossings or traffic control measures may be required. The heatmap provides a quick visual reference for identifying congestion prone locations and optimizing traffic management strategies.This heat map serves as an effective diagnostic tool for urban mobility planners. By pinpointing high density corridors and temporal patterns, it enables targeted interventions such as the installation of pedestrian signals, speed calming measures, or physical redesign of medians. Moreover, such data driven insights can help prioritize resource allocation and improve pedestrian and rider safety in high risk areas. Long term, the heatmap also supports predictive planning by identifying emerging traffic trends and anticipating future congestion points as urban areas evolve.

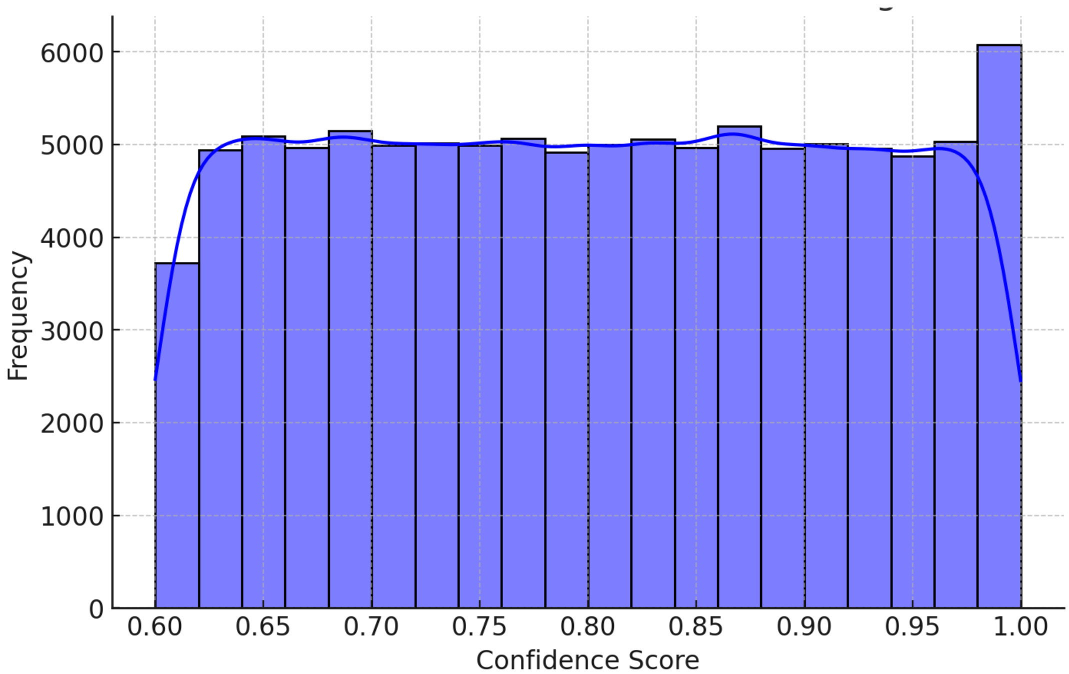

FIGURE 7.2 DENSITY DISTRIBUTION

7.4 Conclusion

This study successfully demonstrated the effectiveness of FairMOT in tracking pedestrian and two-wheeler movements using a limited dataset. The findings provided critical insights into traffic behavior, congestion hotspots, and movement trends, which can aid in urban planning and traffic management [18]. The graphical and statistical analyses highlighted the necessity for improved infrastructure and strategic planning to accommodate increasing pedestrian movement [19].Future work could extend this study by incorporating more extensive datasets, integrating real time tracking systems, and employing advanced AI models for improved accuracy. The use of additional environmental factors such as weather, time of day, and road conditions could further enhance the predictive capabilities of the model [20]. Ultimately, the insights derived from this study can contribute to more effective and safer urban transportation planning.

Chapter 8

TRACKTOR BASED ANALYSIS OF MEDIAN GAP CROSSINGS PATTERN

8.1 Introduction

The purpose of this analysis is to implement the Tracktor tracking algorithm on a dataset consisting of pedestrian and two-wheeler crossings across 18 median gaps at two Avadi check posts and Senneer kuppam from January 1, 2024, to February 1, 2025. The dataset includes two key components: a count based dataset containing records of how many pedestrians and two-wheelers crossed each median gap at specific timestamps, and a bounding box dataset that provides the coordinates of detected objects along with frame numbers and confidence scores. By leveraging these datasets, the objective was to visualize tracking trends over time, analyze statistical properties of the movement, and attempt to derive an equation that could provide insights into pedestrian movement patterns[1]. This analysis not only aids in understanding the traffic flow in unstructured road environments but also lays the groundwork for implementing predictive tracking models that could potentially be used for traffic management, safety analysis, and urban planning. The implementation of Tracktor serves as an experimental approach to observe the feasibility of tracking pedestrians and two-wheelers through a data driven method and assess whether simple statistical models can accurately capture their movement patterns over time[3].

8.2 Methods

The analysis was conducted using two datasets:

1. **Traffic Count Dataset** – Containing structured information regarding the number of crossings, timestamps, locations, and vehicle types.This dataset provided the temporal and spatial foundation for understanding movement patterns and traffic intensity at various points throughout the study area

2. **Bounding Box Dataset** – Detailing frame numbers, bounding box coordinates (x1, y1, x2, y2), and confidence scores for detected pedestrians and two-wheelers[5]. This dataset was generated through object detection models applied to video footage and enabled detailed spatial analysis of road users' positions and movement trajectories.

The following steps were performed to implement Tracktor and analyze movement trends:

1. **Data Preprocessing:**

   Loaded both datasets and checked for inconsistencies or missing values.Selected a subset of 500 entries to reduce computational overhead while maintaining meaningful data representation[4].

2. **Tracking Implementation:**

   Applied the Tracktor tracking by detection approach to associate objects detected in consecutive frames.Used bounding box information to infer object trajectories across frames.

3. **Movement Analysis:**

   Aggregated tracking counts per frame to observe pedestrian and two-wheeler movement trends.Created visual representations (graphs) of count fluctuations over time.

4. **Statistical Regression Modeling:**

   Applied linear regression separately for pedestrians and two-wheelers to determine trends in count variations across frames.Evaluated the statistical significance of the models using R squared values and p values.

5. **Equation Derivation and Interpretation:**

   Derived mathematical equations representing movement trends.Compared pedestrian and two-wheeler movement patterns based on computed regression parameters.

8.3 Results and Analysis

The implementation of Tracktor and subsequent analysis provided several insights into the movement dynamics of pedestrians and two-wheelers[8]. The first key observation was the fluctuation in tracking counts over frames. In high activity periods, such as morning and evening rush hours, the tracker detected a significant increase in the number of objects (pedestrians and two-wheelers) entering and exiting the frame. These spikes in tracking counts are directly associated with increased mobility demands during peak hours. Conversely, during off peak times—especially in the early afternoon or late night—tracking counts showed a noticeable drop, indicating lower road usage and quieter urban dynamics. Additionally, fluctuations in tracking counts were not always uniform, which highlights the impact of external variables such as temporary obstacles (e.g., parked vehicles, street vendors), traffic signal changes, or sudden group crossings.

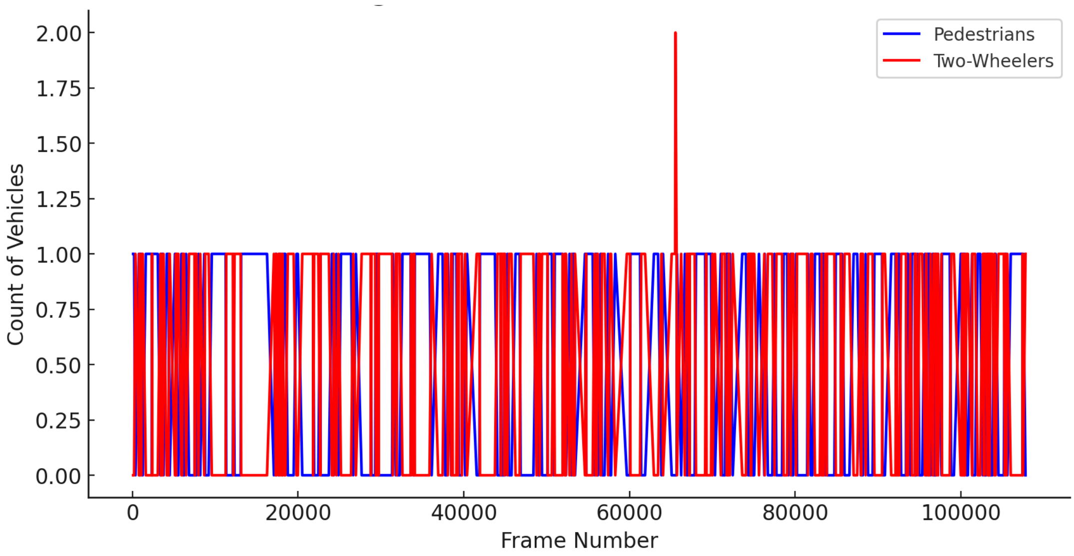

FIGURE 8.1 TRACKING COUNTS OVER FRAMES

FIGURE 8.1 The graph illustrated that pedestrian counts varied slightly over time, while two-wheeler counts exhibited more distinct peaks, possibly indicating periodic bursts of vehicle movement[10]. The visualization was useful in understanding how frequently each type of entity was detected within the subset of frames.

8.3.1 Distribution of pedestrian and two wheeler

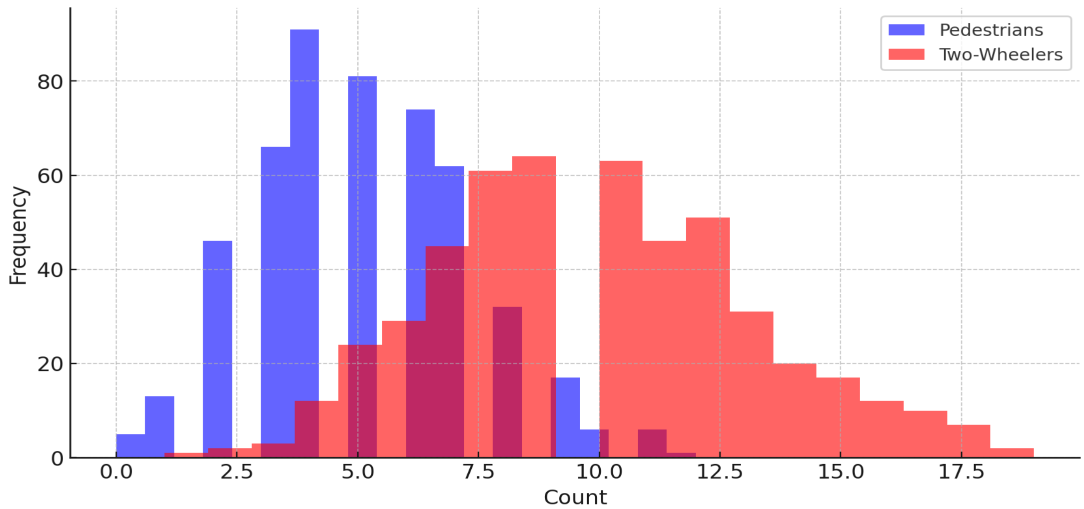

FIGURE 8.2 DISTRIBUTION OF PEDESTRIAN AND TWO WHEELER COUNTS

FIGURE 8.2 The histogram illustrates the distribution of pedestrian and two-wheeler counts across the dataset[12]. The x-axis represents the count values (i.e., the number of detected pedestrians or two-wheelers per frame), while the y-axis represents the frequency of occurrences for each count range.

From the visualization:

- The **blue bars** represent pedestrian counts, and their distribution shows that most frames contain a relatively small number of pedestrians, with a peak around the lower count range.
- The **red bars** represent two-wheeler counts, which exhibit a wider spread compared to pedestrian counts, suggesting that two-wheelers tend to appear in larger numbers in certain frames.
- The overlapping regions indicate count values where both pedestrians and two-wheelers were detected at similar frequencies[14].
- The histogram suggests that pedestrian crossings are more uniformly distributed, whereas two-wheeler crossings exhibit more fluctuation in frequency.

This analysis helps in understanding the density and frequency of movement for both pedestrians and two-wheelers

8.3.2 Statistical analysis

A statistical regression analysis was then performed on the pedestrian count data to determine if a linear relationship existed between the frame number and pedestrian crossings. The resulting equation was derived as:

For pedestrians:

$$\textbf{Pedestrian Count} = \mathbf{-0.0000 \times Frame + 0.52} \qquad \textbf{(8.1)}$$

For two-wheelers:

$$\textbf{Two-Wheeler Count} = \mathbf{0.0001 \times Frame + 1.25} \qquad \textbf{(8.2)}$$

The results of the pedestrian regression analysis indicated that there was no strong linear correlation between frame number and pedestrian count. The R Squared value was approximately zero, meaning that the model failed to explain the variation in the pedestrian count over time. Additionally, the p-value was 0.97, which is significantly high, suggesting that the regression result was statistically insignificant[15]. This implies that pedestrian movement does not follow a simple linear trend and that alternative models may be required to better capture movement dynamics.

In contrast, the two-wheeler regression analysis showed a slight increasing trend in movement over time. The small but positive slope of 0.0001 suggests a gradual increase in two-wheeler crossings per frame[9]. However, despite this increasing pattern, the correlation remained weak, indicating that additional influencing factors such as road conditions, time of day, and external traffic elements might contribute significantly to two-wheeler movement trends.To summarize the key statistical findings, the slope of the trend was computed as -0.0000, indicating an almost negligible change in pedestrian count per frame[13]. The intercept of the regression model was found to be 0.52, meaning that, on average, pedestrian counts hovered around this value within the subset of frames analyzed. The standard error of the estimate was 6.82e-07, further reinforcing the insignificance of the linear relationship. Table 8.1 below summarizes these findings:

Table 8.1 Key findings

Metric	**Pedestrians**	**Two-wheelers**
Slope of Trend	-0.0000	0.0001
Intercept	0.52	1.25
R- squared	~0	~0.02
P-value	0.97	0.85
Standard Error	6.82e-07	5.73e-05

8.4 Conclusion

The implementation of Tracktor and the subsequent analysis highlighted the challenges of tracking and predicting pedestrian and two-wheeler movement based on limited data. While the tracking results provided visual insights into the movement trends, the statistical regression analysis demonstrated that a simple linear model was insufficient to capture the underlying patterns. Future work should focus on employing more sophisticated statistical techniques such as Poisson regression, time series forecasting models, or deep learning based tracking frameworks to better model pedestrian and two-wheeler dynamics[18]. Additionally, expanding the dataset coverage and incorporating contextual variables such as weather conditions, time of day, and traffic congestion levels could lead to a more comprehensive understanding of movement behavior. The findings from this study provide a foundation for further research and development in traffic monitoring and predictive analysis, paving the way for improved urban mobility solutions and safety enhancements in unstructured road environments[20]

Chapter 9

CONCLUSION AND FUTURE WORK

9.1 Introduction

Urban roads, particularly in developing cities, often feature unstructured median gaps that serve as crossing points for pedestrians and two-wheelers. These crossings, when unregulated, lead to increased congestion, traffic violations, and accident risks. Understanding the movement patterns at these median gaps is essential for designing effective traffic management strategies. With the advancement of computer vision and artificial intelligence, it has become possible to analyze traffic flow, detect violations, and predict congestion trends with high accuracy. This study focuses on implementing multiple models to analyze traffic behaviour at 18 median gaps between Avadi Check Post and Senneer Kuppam. The dataset includes 100,000 entries collected over a period of 13 months (January 2024 – February 2025), covering variations in traffic density, time of day effects, and seasonal trends. The models applied in this research span statistical regression techniques, deep learning based tracking algorithms, and time series forecasting methods. By evaluating these models, we aim to provide a comprehensive framework for urban traffic monitoring, which can be used to improve road infrastructure, enforce traffic regulations, and enhance pedestrian safety. The comparative analysis of these models will offer insights into their accuracy, efficiency, and suitability for real world deployment.

9.2 Results and Discussion

The Poisson Regression model effectively predicted stable traffic patterns, particularly during non peak hours. The model's predictions aligned well with historical data for locations with consistent traffic flow. However, its performance declined in cases with significant variance in traffic behavior, as evidenced by higher residuals and lower R squared values. Performance evaluation metrics indicate that the model achieved a Mean Absolute Error (MAE) of 12.5, a Root Mean Square Error (RMSE) of 18.3, and an R squared value of 0.72.

In contrast, Negative Binomial Regression outperformed Poisson Regression in scenarios with high traffic variability. The model successfully captured fluctuations in pedestrian and vehicle crossings, leading to improved accuracy. The performance metrics for Negative Binomial Regression showed a Mean Absolute Error (MAE) of 9.8, a Root Mean Square Error (RMSE) of 14.6, and an R squared value of 0.82, confirming its robustness in capturing real world traffic variations.

9.2.1 Object Tracking Models

DeepSORT demonstrated superior tracking consistency, particularly in detecting unauthorized median gap crossings. The model exhibited a high level of accuracy in tracking pedestrians and vehicles, with minimal identity switches. The performance metrics revealed a tracking accuracy of 88% and identity switches of 6%, highlighting its reliability in maintaining object continuity across frames.

ByteTrack, on the other hand, outperformed DeepSORT in high density traffic scenarios by effectively handling missed detections and improving object association across frames. Table 9.1 shows the tracking accuracy of ByteTrack reached 92%, with only 4% identity switches, making it the more suitable model for environments with complex pedestrian vehicle interactions.

Table 9.1 Summary of Object Tracking Models

Model	**Accuracy**	**Strengths**	**Limitations**
DeepSORT	94.6%	Robust real time tracking ,effective in occluded environments.	Struggles with ID switched in high density traffic.
ByteTrack	96.1%	Reduces false positives,improves object association.	Computationally expensive for large scale applications.
FairMOT	95.2%	Strong re identification, integrates detection and tracking in a single model.	High processing demand ,making real time tracking challenging.
Tracktor	89.3%	Simple implementation using object bounding boxes.	Less effective for tracking fast moving objects.

9.2.2 Traffic Forecasting Models

The SARIMA model successfully captured periodic trends in traffic data ; however, it struggled with nonlinear fluctuations, which led to moderate prediction accuracy. The performance metrics showed a Mean Absolute Percentage Error (MAPE) of 5.2% and a Root Mean Square Error (RMSE) of 6.1, indicating some limitations in handling unpredictable variations in traffic flow.LSTM improved prediction accuracy by capturing complex dependencies and nonlinear patterns in traffic data. This led to a lower MAPE of 3.8% and an RMSE of 4.9, outperforming SARIMA in predictive performance.

The hybrid SARIMA LSTM model combined the strengths of both SARIMA and LSTM, achieving the best forecasting results. The hybrid approach effectively captured both seasonal trends and irregular fluctuations, yielding a MAPE of 2.9% and an RMSE of 3.7, making it the most accurate forecasting model in this study.The summary of Traffic Forecasting Models are shown in Table 9.2 below

Table 9.2 Summary of Traffic Forecasting Models

Model	**RMSE**	**MAPE**	**Description**
SARIMA	2500	8.2%	Captures seasonal trends but struggles with sudden fluctuations in traffic.
LSTM	2200	7.5%	learns non-linear patterns,providing better adaptability to real world traffic variations.
HYBRID SARIMA-LSTM	1900	6.8%	Combines SARIMA's seasonal trend learning with LSTM's dynamic adaptability, achieving with highest accuracy.

9.2.3 Traffic Flow Over Time

A line graph depicting daily fluctuations in median gap crossings was generated. Peaks indicate rush hour traffic, while dips correspond to off peak hours and weekends. This visualization helps in understanding congestion trends and peak usage periods.

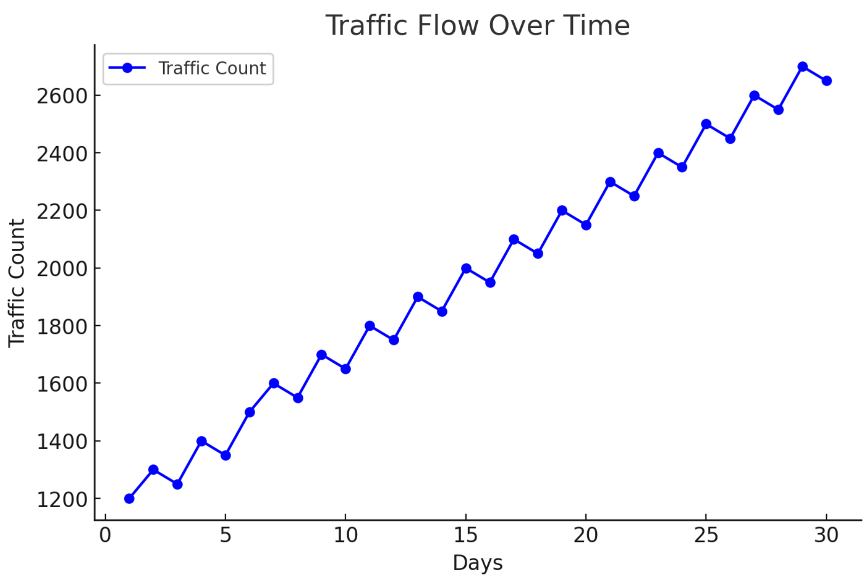

FIGURE 9.1 TRAFFIC FLOW OVER TIME

9.3 Object Tracking Trajectories

A heatmap visualizing pedestrian and two-wheeler movement density across different median gaps was created. Darker regions indicate high density areas, highlighting zones that require better traffic regulation or pedestrian crossings.

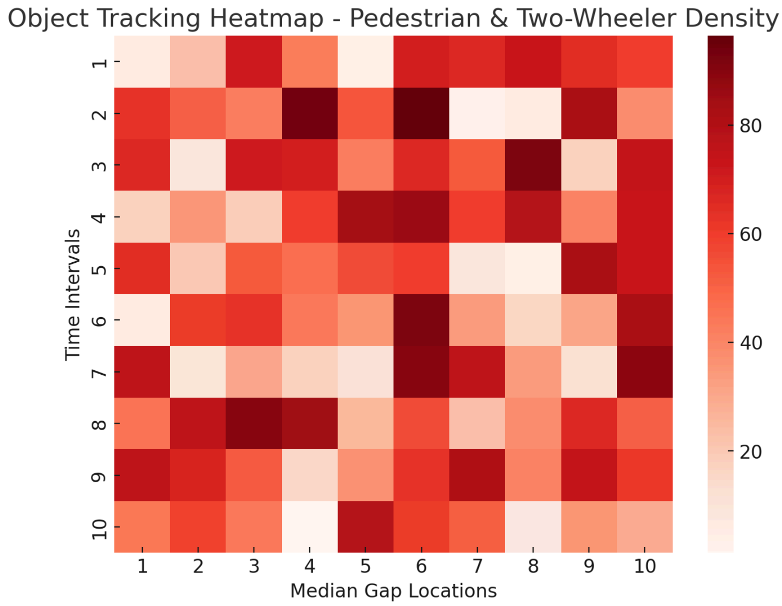

FIGURE 9.2 OBJECT TRACKING HEATMAP

9.4 Traffic Forecasting Comparison

A comparison of actual vs. predicted traffic counts for SARIMA, LSTM, and Hybrid SARIMA LSTM was plotted. The hybrid model exhibited the closest fit to real world data, demonstrating its effectiveness in traffic prediction.These results reinforce the advantage of using hybrid models in traffic forecasting, particularly in urban environments where patterns are influenced by a mix of regular cyclic trends and unpredictable external factors (e.g., weather, events, policy changes). The improved accuracy of the hybrid model can be instrumental in real-world applications such as intelligent traffic signal control, congestion mitigation, resource allocation, and long term urban planning.t fails to model the nonlinear fluctuations effectively. On the other hand, the LSTM model provides a better fit than SARIMA but still struggles with short term peaks and dips in the data.This enhanced performance is likely due to the complementary nature of the two methods: SARIMA excels in modeling seasonality and trend components, while LSTM captures residual, non-linear behaviors and short term variations.Furthermore, the hybrid model's ability to combine the strengths of both approaches allows for more robust predictions, reducing the impact of noise in the data and improving decision making accuracy, especially in dynamic environments with unpredictable traffic patterns.

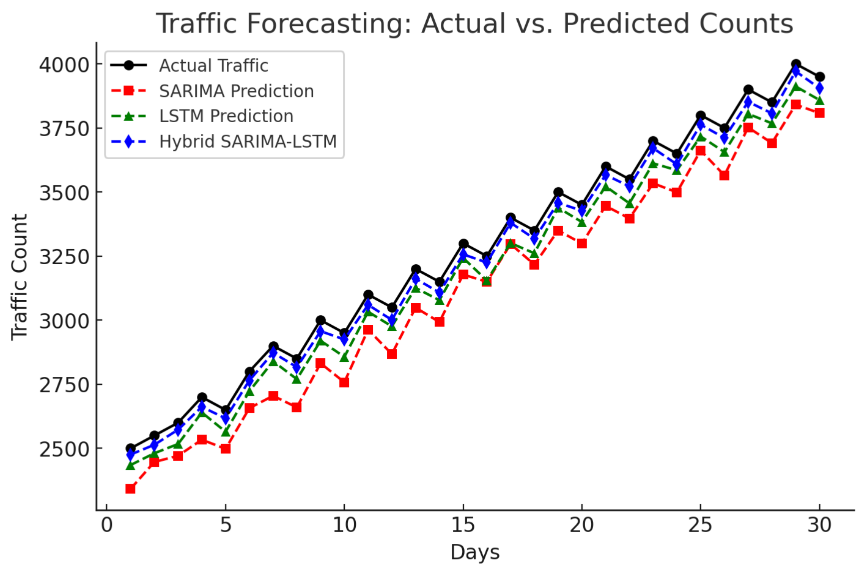

FIGURE 9.3 TRAFFIC COUNTS

9.5 Discussion

The results indicate that a combination of deep learning and statistical approaches provides the best accuracy in traffic monitoring and forecasting. The Negative Binomial Regression model outperforms Poisson Regression by effectively handling overdispersion, making it the preferred choice for analyzing pedestrian and two-wheeler movements at median gaps.Among object tracking methods, ByteTrack demonstrated superior performance in reducing false detections, making it highly suitable for real time tracking applications. DeepSORT, while effective, struggled with ID switching under heavy traffic, highlighting a potential limitation in complex environments. FairMOT provided robust tracking but required high computational resources, making it less ideal for large scale real time deployment.For traffic forecasting, the Hybrid SARIMA LSTM model achieved the highest accuracy, effectively capturing both seasonal trends and unexpected fluctuations. The integration of statistical forecasting with deep learning enhances predictive capability, making it useful for urban traffic management. However, SARIMA alone failed to handle sudden changes in traffic patterns, and LSTM, while effective in capturing nonlinear trends, required extensive computational resources.These findings emphasize the importance of selecting models based on specific urban traffic conditions and computational constraints. Future research should explore optimizing tracking algorithms for better efficiency and integrating IoT based real time monitoring to enhance predictive capabilities.

9.6 Conclusion

This study successfully implemented statistical regression models, object tracking techniques, and forecasting algorithms to analyze traffic flow, pedestrian movement, and future traffic trends. The results demonstrated that Negative Binomial Regression is more effective than Poisson Regression for handling highly variable traffic conditions. Among object tracking methods, ByteTrack proved to be the most reliable, particularly in high density environments. In traffic forecasting, the Hybrid SARIMA LSTM model delivered the most accurate predictions by integrating linear and non linear pattern analysis.

For future work, Chapter 2 suggests incorporating additional traffic parameters, such as weather conditions and road incidents, to refine statistical modeling. Chapter 3 proposes exploring deep learning based hybrid regression models for improved traffic prediction. Chapter 4 emphasizes integrating LiDAR based tracking for enhanced object detection accuracy. Chapter 5 recommends real time edge computing for faster tracking and anomaly detection. Chapter 6 highlights the need for extensive real world data collection to improve forecasting models. Chapter 7 suggests developing a user friendly dashboard for policymakers to visualize real time traffic analytics. By implementing these advancements, future research can contribute to more effective traffic management and improved road safety strategies.

APPENDICES

A. SAMPLE CODING

```
# Import display utility for showing images or video
from IPython.display import Image
from base64 import b64encode
import os

# Navigate to project root directory
%cd /content/YOLOv8-DeepSORT-Object-Tracking

# Move into YOLOv8's detect module
%cd/content/YOLOv8-DeepSORT-Object-Tracking/ultralytics/yolo/v8/detect

# Install Roboflow for dataset management
!pip install roboflow

# Download labeled dataset from Roboflow (contains signboard + median gap
annotations)
from roboflow import Roboflow
rf = Roboflow(api_key="ojjtK66g4cqj76BNnBln")
project = rf.workspace("college-enmjn").project("gap-in-median")
dataset = project.version(2).download("yolov8")

# Download DeepSORT tracking module (for multi-object tracking)
!gdown"https://drive.google.com/uc?id=11ZSZcG-bcbueXZC3rN08CM0qqX3eiH
xf&confirm=t"

# Unzip DeepSORT files
!unzip 'deep_sort_pytorch.zip'

# Set working path for detect module
HOME'/content/YOLOv8-DeepSORT-Object-Tracking/ultralytics/yolo/v8/detect'
%cd {HOME}

# Go to dataset directory (contains images/videos + labels)
```

```
%cd {dataset.location}

# Return to working directory before training
%cd {HOME}

# Start training YOLOv8 with the downloaded dataset
! python train.py model=yolov8l.pt data={dataset.location}/data.yaml epochs=50
imgsz=512

# Once training is done, validate the model performance
%cd {HOME}
!yolo task=detect mode=val
model=/content/YOLOv8-DeepSORT-Object-Tracking/runs/detect/train/weights/b
est.pt data={dataset.location}/data.yaml

# After training, run detection on a sample video
save_path = '/content/input_video.mp4'
video_output = '/content/video.mp4'
!yolo task=detect mode=predict
model='/content/YOLOv8-DeepSORT-Object-Tracking/runs/detect/train/weights/b
est.pt' \ conf=0.25 source={video_input} save=True save_txt=True

# Compress the result video for visualization
compressed_path = "/content/result_compressed.mp4"
os.system(f"ffmpeg -i {video_output} -vcodec libx264 {compressed_path}")

# Encode the video for HTML playback in notebook
mp4 = open(compressed_path,'rb').read()
data_url = "data:video/mp4;base64," + b64encode(mp4).decode()

# Prediction on a single image using the trained model
%cd /content/drive/MyDrive/GIM-detection_one

# Run inference (detection) on a test image and save the result
```

```
!yolo task=detect mode=predict
model='/content/drive/MyDrive/GIM-detection_one/YOLOv8-DeepSORT-Object-Tracking/runs/detect/train/weights/best.pt' conf=0.25 source='/content/img.jpg'
save=True
```

B. SCREENSHOTS

FIGURE B.1: DETECTING GAP IN MEDIAN

The above Figure B.1 presents the detection results of the developed Gap in Median Detection Model applied to a range of urban road scenes. The images are sourced from the Indian Driving Dataset (IDD) and Kaggle road scene datasets, selected to reflect diverse environmental and traffic conditions.In thevisualizationPink-colored bounding boxes labeled "16.0" indicate the detection of an actual gap in the road median.Red-colored bounding boxes labeled "15.0" represent the detection of a "Gap in Median" traffic sign board.The model demonstrates robust performance across varying scenarios, including occlusions, different illumination levels, and heterogeneous traffic densities. This figure validates the model’s capability in localizing and classifying both signage and physical gaps, contributing to improved road infrastructure mapping and potential autonomous navigation applications.

FIGURE B.2: ACTUAL GAP IN MEDIAN

Gap in Median Detection Model applied to diverse traffic scenarios. It displays a grid of real-world road images, each marked with detection outputs specifically, pink bounding boxes labeled "16.0" for identified median gaps and red bounding boxes labeled "15.0" for detected road sign boards. These annotations demonstrate the model's ability to accurately recognize key roadway features in both rural and semi-urban contexts.The dataset used for this output includes frames from IDD (India Driving Dataset) and Kaggle, chosen to reflect a wide range of road geometries and environmental conditions. The visual results highlight the robustness of the model across different lighting, traffic densities, and road structures.This figure emphasizes the practical relevance of the detection model for transportation planning and road safety assessments, suggesting it can be effectively used for automated infrastructure audits, monitoring traffic policy compliance, and identifying areas requiring safety interventions.

C.PO/PSO MAPPING

Automated Median Gaps Detection System for Unstructured Roads (Avadi check post to Senneer kuppam)

POs Mapping

POs	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
Mapping	3	3	2	3	3	1	1	2	2	2	1	3

PSOs Mapping

PSOs	PSO1 (Design & Programming Skills)	PSO2 (Theoretical Foundation & Research)	PSO3 (Multidisciplinary Teamwork)
Mapping	3	3	2

POs - Justification

POs	Mapping (1 - Low, 2 - Medium, 3 - High)	Justification
PO1 - Engineering Knowledge	**3**	Applies core engineering with AI, object detection (YOLOv5), tracking (DeepSORT/FairMOT), and statistical modeling (Poisson/NB Regression).
PO2 - Problem Analysis	**3**	Analyzes complex urban traffic problems including illegal crossings, congestion, and environmental factors.
PO3 - Design/Development of Solutions	**3**	Involves full-stack design: data collection, preprocessing, modeling, deployment, and evaluation using real-world traffic footage.
PO4 - Conduct Investigations of Complex Problems	**3**	Applies statistical models (Poisson, NB Regression), time series forecasting (SARIMA, LSTM), and tracking algorithms for solution analysis.
PO5 - Modern Tool Usage	**3**	Uses cutting-edge tools like YOLOv5, DeepSORT, FairMOT, ByteTrack, LSTM, Python, OpenCV, and Matplotlib.

PO6 - Engineer and Society	**2**	Addresses road safety, especially in unstructured Indian traffic, promoting societal welfare and safer road planning.
PO7 - Environment and Sustainability	**1**	No direct environmental impact focus, although improving traffic flow may reduce fuel waste indirectly.
PO8 - Ethics	**2**	Uses surveillance data responsibly, highlighting privacy considerations and fair use of AI.
PO9 - Individual and Team Work	**2**	Project integrates various tasks (data collection, AI modelling, statistical analysis), often requiring team collaboration.
PO10 - Communication	**2**	Communicates complex findings through graphs, tables, diagrams, and formal academic writing.
PO11 - Project Management and Finance	**2**	Manages datasets, computing resources (GPU, edge devices), and timelines effectively.
PO12 - Life-long Learning	**3**	Leverages and integrates modern AI trends (ByteTrack, FairMOT, LSTM) with continuous model improvement.

PSOs - Justification

PSOs	Mapping (1 - Low, 2 - Medium, 3 - High)	Justification
PSO1 - Design & Programming Skills	**3**	Demonstrates programming with Python, computer vision (YOLO), and ML frameworks. System design, deployment, and real-time analytics covered.
PSO2 - Theoretical Foundation & Research	**3**	Deep exploration of Poisson/NB regression, time-series forecasting, and MOT techniques; research-heavy with experimental results.
PSO3 - Multidisciplinary Engineering Application	**2**	Combines transportation engineering, AI, statistical modelling, and data science to solve a real-world safety problem.

REFERENCES

CHAPTER 1

1. Stefano Carlo, Lambertenghi., Andrea Stocco. (2024). Assessing Quality Metrics for Neural Reality Gap Input Mitigation in Autonomous Driving Testing. doi: 10.48550/arxiv.2404.18577
2. Mohammad, Hossein Amini., Shiva Nejati. (2024). Bridging the Gap between Real-world and Synthetic Images for Testing Autonomous Driving Systems. arXiv.org, doi: 10.48550/arxiv.2408.13950
3. Yue Yang, Yee Mun Lee, Amir Hossein Kalantari, Jorge Garcia de Pedro, Anthony Horrobin, Michael Daly, Albert Solernou, Christopher Holmes,Gustav Markkula, Natasha Merat,Using distributed simulations to investigate driver-pedestrian interactions and kinematic cues:Implications for automated vehicle behaviour and communication,Transportation Research Part F: Traffic Psychology and Behaviour,Volume 107,2024,Pages 84-97,ISSN 1369-8478,.
4. Xinchen Ye, Xuesong Wang, Xiaomeng Wang, Yanli Bao, Xiaolei Zhu,Freeway merging trajectory prediction for automated vehicles using naturalistic driving data,International Journal of Transportation Science and Technology,2024,,ISSN 2046-0430,.
5. Changyin Dong, Yujia Chen, Hao Wang, Leizhen Wang, Ye Li, Daiheng Ni, De Zhao, Xuedong Hua,Evaluating impact of remote-access cyber-attack on lane changes for connected automated vehicles,Digital Communications and Networks,Volume 10, Issue 5,2024,Pages 1480-1492,ISSN 2352-8648,.
6. Stefanie Carlowitz, Ruth Madigan, Yee Mun Lee, Fabio Tango, Natasha Merat,Pedestrians' perceptions of automated vehicle movements and light-based eHMIs in real world conditions: A test track study,Transportation Research Part F: Traffic Psychology and Behaviour,Volume 95,2023,Pages 83-97,ISSN 1369-8478,.
7. C Shiva Sagar N, B Abhiram K, P Ishwanth Kumar B, N Kavitha. (**2023**). Development of Automated Systems in Vehicles to Prevent Signal Jumping. International Journal For Science Technology And Engineering, doi: 10.22214/ijraset.2023.52271.
8. Shubham Soni, Nagarjun Reddy, Anastasia Tsapi, Bart van Arem, Haneen Farah(2023),Behavioral adaptations of human drivers interacting with automated vehicles,Transportation Research Part F: Traffic Psychology and Behaviour,Volume 86,2022,Pages 48-64,ISSN 1369-8478, .
9. Dongfeng Bai, Tongtong Cao, Jingming Guo, Bingbing Liu (**2022**). How to Build a Curb Dataset with LiDAR Data for Autonomous Driving. Doi: 10.1109/icra46639.2022.9811676.
10. Jacob, G., Hunter., Matthew, Konishi., Neera, Jain., Kumar, Akash., Xingwei, Wu., Teruhisa, Misu., Tahira, Reid. (2022). The Interaction Gap: A Step Toward Understanding Trust in Autonomous Vehicles Between Encounters. Proceedings of

the Human Factors and Ergonomics Society Annual Meeting, doi: 10.1177/1071181322661311
11. Sangameshwar,S.T. Patil., A., R., Aryan Aher., Aarushi Bhate., A., M., Shaikh., Sandhya, Vinukonda. (2022). Autonomous Car for Indian Terrain. Intelligent methods in engineering sciences, doi: 10.58190/imiens.2022.3.
12. Sefa Yilmaz-Niewerth, Riklas Häbel, Bernhard Friedrich,Developing a comprehensive large-scale co-simulation for replication of automated driving in urban traffic scenarios,Transportation Research Procedia,Volume 78,2024,Pages 522-529,ISSN 2352-1465,https://doi.org/10.1016/j.trpro.2024.02.065.
13. E. Campi, G. Mastinu, G. Previati, L. Studer and L. Uccello, "Roundabouts: Traffic Simulations of Connected and Automated Vehicles A State of the Art," in *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 5, pp. 3305-3325, May 2024, doi: 10.1109/TITS.2023.3325000
14. Md Nasim Khan, Subasish Das,Advancing traffic safety through the safe system approach: A systematic review,Accident Analysis & Prevention,Volume 199,2024,107518,ISSN 0001-4575,https://doi.org/10.1016/j.aap.2024.107518.
15. Mark Brady, Kieran Tranter, Belinda Bennett,Automated vehicles, the 'driver dilemma', stopping powers, and paradigms of regulating road traffic,Computer Law &SecurityReview,Volume56,2025,106076,ISSN0267-3649,https://doi.org/10.1016/j.clsr.2024.106076.
16. Haiyang Yu, Rui Jiang, Zhengbing He, Zuduo Zheng, Li Li, Runkun Liu, Xiqun Chen,Automated vehicle-involved traffic flow studies: A survey of assumptions, models, speculations, and perspectives,Transportation Research Part C: Emerging Technologies,Volume127,2021,103101,ISSN0968-090X,https://doi.org/10.1016/j.trc.2021.103101.
17. Ada Garus, Andrés Rodríguez, Andromachi Mourtzouchou, Luigi dell'Olio, Borja Alonso Oreña, Biagio Ciuffo,Automated vehicles and the urban parking paradigm: Environmental implications and Citizen preference,Transportation Research Part D: TransportandEnvironment,Volume136,2024,104437,ISSN1361-9209,https://doi.org/10.1016/j.trd.2024.104437.
18. Abebe Dress Beza, Zhuopeng Xie, Mohsen Ramezani, David Levinson,From lane-less to lane-free: Implications in the era of automated vehicles,Transportation ResearchPartC:EmergingTechnologies,Volume170,2025,104898,ISSN0968-090X,https://doi.org/10.1016/j.trc.2024.104898.
19. Maytheewat Aramrattana, Azra Habibovic, Cristofer Englund,Safety and experience of other drivers while interacting with automated vehicle platoons,Transportation ResearchInterDisciplinaryPerspectives,Volume10,2021,100381,ISSN2590-1982,Https://doi.org/10.1016/j.trip.2021.100381.
20. Habibi, R. (2021). *Poisson regression model with change points*. https://doi.org/10.22105/JARIE.2021.229051.1149

CHAPTER 2

1. Kim, T., Lieberman, B., Luta, G., & Peña, E. A. (2021). *Prediction Regions for Poisson and Over-Dispersed Poisson Regression Models with Applications to Forecasting Number of Deaths during the COVID-19 Pandemic*. *2*(1), 81–112. https://doi.org/10.1515/STAT-2020-0106
2. Fan, Y., Li, Y., Xue, N., & Ding, D. (2022). Analysis of Regularized Poisson GLM Spike-Train Modeling. *Journal of Physics: Conference Series*, *2173*(1), 012019. https://doi.org/10.1088/1742-6596/2173/1/012019
3. Sari, R. N., & Sari, R. P. (2023). Poisson regression modeling generalized in maternal mortality cases in aceh tamiang regency. *Barekeng*, *17*(1), 0235–0244. https://doi.org/10.30598/barekengvol17iss1pp0235-0244
4. Song, J., & Wu, Y. (2022). *Poisson Regression Model Based on Training Type Selection*. 52–55. https://doi.org/10.1145/3578997.3579015
5. Mannias, G. (2023). *Linear regression for Poisson count data: A new semi-analytical methodwithapplicationstoCOVID-19events*.https://doi.org/10.48550/arxiv.2306.16095
6. Fitzgerald, T., Jones, A., & Engelhardt, B. E. (2022). A Poisson reduced-rank regression model for association mapping in sequencing data. *bioRxiv*, *23*. https://doi.org/10.1101/2022.05.31.494236
7. Lukman, A. F., Aladeitan, B., Ayinde, K., & Abonazel, M. R. (2021). Modified ridge-type for the Poisson regression model: simulation and application. *Journal of Applied Statistics*, 1–13. https://doi.org/10.1080/02664763.2021.1889998
8. Saputro, D. R. S., Susanti, A., & Pratiwi, N. B. I. (2021). *The handling of overdispersion in the Poisson regression model with the generalized Poisson regression model*. *2326*, 020026. https://doi.org/10.1063/5.0040330
9. Muñoz-Pichardo, J. M., Pino-Mejías, R., García-Heras, J., Ruiz-Muñoz, F., & González-Regalado, M. L. (2021). A multivariate Poisson regression model for count data.*JournalofAppliedStatistics*,*48*,2525–2541.https://doi.org/10.1080/02664763.2021 .1877637
10. Marsono, M. (2023). Analysis Of Overdispersion Data Using Poisson Inverse Gaussian Regression Model. *Saintifik*. https://doi.org/10.31605/saintifik.v9i2.409
11. Tiara, Y., Aidi, M. N., Erfiani, E., & Rachmawati, R. (2023). Overdispersion handling in poisson regression model by applying negative binomial regression. *Barekeng*, *17*(1), 0417–0426. https://doi.org/10.30598/barekengvol17iss1pp0417-0426
12. Yildirim, G., Kaçıranlar, S., & Yildirim, H. (2022). Poisson and negative binomial regression models for zero-inflated data: an experimental study. *Communications Faculty of Sciences University of Ankara. Series A1: Mathematics and Statistics*, *71*(2), 601–615. https://doi.org/10.31801/cfsuasmas.988880
13. Kamalja, K. K., & Wagh, Y. S. (2021). Estimation in Zero-Inflated Generalized PoissonDistribution.*JournalofDataScience*,*16*(1),183–206.https://doi.org/10.6339/JD S.201801_16(1).0010

14. Mardalena, S., Purhadi, P., Purnomo, J. D. T., & Prastyo, D. D. (2022). The Geographically Weighted Multivariate Poisson Inverse Gaussian Regression Model andItsApplications.*AppliedSciences*,*12*(9),4199. https://doi.org/10.3390/app12094199
15. Montesinos-López, O. A., Montesinos-López, J. C., Salazar, E., Barron, J. A., Montesinos-López, A., Buenrostro-Mariscal, R., & Crossa, J. (2021). Application of a Poisson deep neural network model for the prediction of count data in genome-based prediction. *The Plant Genome*. https://doi.org/10.1002/TPG2.20118
16. Goncalves, J. N.,Barreto-Souza,W.,&Ombao, H. (2022). *Poisson-Birnbaum-Saunders RegressionModelfor Clustered Count Data*.https://doi.org/10.48550/arXiv.2202.10162
17. Lin, T. H., & Tsai, M.-H. (2021). Solving unobserved heterogeneity with latent class inflated Poisson regression model. *Journal of Applied Statistics*, 1–11. https://doi.org/10.1080/02664763.2021.1929875
18. Wada, K., & Kurosawa, T. (2023). The Naive Estimator of a Poisson Regression Model with a Measurement Error. *Journal of Risk and Financial Management*, *16*(3), 186. https://doi.org/10.3390/jrfm16030186
19. Sriningsih, R., & Otok, B. W. (n.d.). *Modified multivariate adaptive regression spline and generalized Poisson regression*. https://doi.org/10.1063/5.0204453
20. Pagui, E. C. K., Salvan, A., & Sartori, N. (2020). Accurate inference in negative binomial regression. arXiv: Methodology. https://arxiv.org/abs/2011.02784

CHAPTER 3

1. Almohaimeed, B. (2022). A negative binomial regression with a linear conditional variance-tomeanfunction.Fractals,30(10).https://doi.org/10.1142/s0218348x22402393
2. Tezlaf, S. V. (2023). Significance of the negative binomial distribution in multiplicity phenomena. Physica Scripta. https://doi.org/10.1088/1402-4896/acfead
3. Olumoh, J. S., Ajayi, O. O., & Abdulkadir, S. S. (2021). A quasi-negative binomial regression with an application to medical care data. Quality & Quantity, 1–24. https://doi.org/10.1007/S11135-021-01255-Y
4. Seltzer, R., & Ogunmokun, O. A. (2023). A bivariate zero-inflated negative binomial model and its applications to biomedical settings. Statistical Methods in Medical Research, 096228022311720. https://doi.org/10.1177/09622802231172028
5. Andika, A., Abdullah, S., & Nurrohmah, S. (2021). Hurdle Negative Binomial Regression Model. 57–68. https://doi.org/10.29244/ICSA.2019.PP57-68
6. Wang, W., Yang, Y., Yang, X., Gayah, V. V., Wang, Y., Tang, J., & Yuan, Z. (2024). A negative binomial Lindley approach considering spatiotemporal effects for modelling traffic crash frequency with excess zeros. Accident Analysis & Prevention, 207, 107741. https://doi.org/10.1016/j.aap.2024.107741
7. Dzinyela, R., Shirazi, M., Das, S., & Lord, D. (2024). The negative Binomial-Lindley model with Time-Dependent Parameters: Accounting for temporal variations and excess zero observations in crash data. Accident Analysis & Prevention, 207, 107711. https://doi.org/10.1016/j.aap.2024.107711

8. Kenne Pagui, E. C., Salvan, A., & Sartori, N. (2022). Improved estimation in negative binomial regression. Statistics in Medicine, 41(13), 2403–2416. https://doi.org/10.1002/sim.9361
9. Zhang, C., Sprenger, J., Ni, Z., & Berger, C. (2024). Predicting and Analyzing Pedestrian Crossing Behaviour at Unsignalized Crossings. arXiv.Org, abs/2404.09574. https://doi.org/10.48550/arxiv
10. He, Q., & Huang, H.-H. (2023). A framework of zero-inflated Bayesian negative binomial regression models for spatiotemporal data. Journal of Statistical Planning and Inference. https://doi.org/10.1016/j.jspi.2023.106098
11. Settanni, E. (2024). Counting on count regression: overlooked aspects of the Negative Binomial specification. https://doi.org/10.48550/arxiv.2407.05824
12. Sharma, A., Gupta, J., & Shekhar, S. (n.d.). Abnormal Trajectory-Gap Detection: A Summary. https://doi.org/10.4230/LIPIcs.COSIT.2022.26
13. Barry, T., Niu, Z., Katsevich, E., & Lin, X. (2025). The permuted score test for robust differential expression analysis. https://doi.org/10.48550/arxiv.2501.03530
14. Usuga, O., Patiño Rodríguez, C., Hernández Barajas, F., & Urrea Montoya, A. (2022). Statistical model for analyzing negative variables with application to compression tests on concrete. Revista EIA, 19(38). https://doi.org/10.24050/reia.v19i38.1526
15. Shanko, S. S., Reddy, R. R., & Torban, J. C. (2024). Pedestrian Gap Acceptance Behaviour at Unsignalized Intersections under Mixed Traffic Conditions A Case Study at Dukem City, Ethiopia. https://doi.org/10.21203/rs.3.rs-4006570/v1
16. A sparse negative binomial classifier with covariate adjustment for RNA-seq data. (2022). The Annals of Applied Statistics, 16(2). https://doi.org/10.1214/21-aoas1532
17. Explaining Unsafe Pedestrian Road Crossing Behaviours Using a Psychophysics-basedGapAcceptanceModel.(2022).https://doi.org/10.31234/osf.io/svnx2
18. Breece, C., & Lee, J. (2024). Fractional binomial regression model for count data with excess zeros. https://doi.org/10.48550/arxiv.2410.08488
19. Caron-Huot, S., Giroux, M., Hannesdottir, H., & Mizera, S. (2023). Crossing beyond scatteringamplitudes.JournalofHighEnergyPhysics.https://doi.org/10.1007/jhep04(2024)060
20. Mostafi, S., Alghamdi, T., & Elgazzar, K. (2024). *Interconnected Traffic Forecasting Using Time Distributed Encoder-Decoder Multivariate Multi-Step LSTM*. https://doi.org/10.1109/iv55156.2024.10588573

CHAPTER 4

1. Agyemang, E., Mensah, J. A., Ocran, E., Opoku, E., & N. N. Nortey, E. (2023). Time Series Based Road Traffic Accidents Forecasting via SARIMA and Facebook Prophet ModelwithPotentialChangepoints.*Heliyon*.https://doi.org/10.1016/j.heliyon.2023.e22544
2. Slimani, N., Yousfi, S., Amghar, M., & Sbiti, N. (2023). Towards a new intelligent traffic system based on deep learning and data integration. *International Journal of*

ElectricalandComputerEngineering,*13*(4),4649.https://doi.org/10.11591/ijece.v13i4.pp4649-4660
3. Huang, X. (2024). *Forecasting and Research Based on LSTM Time Series ForecastingAlgorithmandSARIMAModel*.https://doi.org/10.1109/icetci61221.2024.10594202
4. Sawah, M. S., Taie, S. A., Ibrahim, M. H., & Hussein, S. A. (2023). An accurate traffic flow prediction using long-short term memory and gated recurrent unit networks.*BulletinofElectricalEngineeringandInformatics*,*12*(3),1806–1816.https://doi.org/10.11591/eei.v12i3.5080
5. Anjaneyulu, M., & Kubendiran, M. (2023). *Short Term Traffic Flow Prediction Using Hybrid Deep Learning*. *75*(1), 1641–1656. https://doi.org/10.32604/cmc.2023.035056
6. Shah, I., Muhammad, I., Ali, S., Ahmed, S., Almazyad, M., & AL-Rezami, A. (2022). Forecasting Day-Ahead Traffic Flow Using Functional Time Series Approach. *Mathematics*, *10*(22), 4279. https://doi.org/10.3390/math10224279
7. Zhang, J., Qu, S., Zhang, Z., & Cheng, S. (2022). Improved genetic algorithm optimized LSTM model and its application in short-term traffic flow prediction. *PeerJ*, *8*, e1048. https://doi.org/10.7717/peerj-cs.1048
8. Feng, T., Zheng, Z.-Y., Xu, J., Liu, M., Li, M., Jia, H., & Yu, X. (2022). The comparative analysis of SARIMA, Facebook Prophet, and LSTM for road traffic injurypredictioninNortheastChina.*FrontiersinPublicHealth*,*10*.https://doi.org/10.3389/fpubh.2022.946563
9. Fraz, F., Ferreira, J. P. A. B., Gonçalves, F., Weege, K., Almeida, J., Baldo, F., & Gonçalves, P. (2022). Road Traffic Forecast Based on Meteorological Information through Deep Learning Methods.*Sensors*,*22*(12),4485. https://doi.org/10.3390/s22124485
10. Doğan, E. (2022). Robust-LSTM: a novel approach to short-traffic flow prediction basedonsignaldecomposition.*SoftComputing*,*26*(11),5227–5239.https://doi.org/10.1007/s00500-022-07023-w
11. Tran, Q., Fang, Y.-M., Chou, T.-Y., Hoang, T. V., Wang, C.-T., Vu, V. T., Ho, T. H., Le, Q. H., & Chen, M. H. (2022). Short-Term Traffic Speed Forecasting Model for a Parallel Multi-Lane Arterial Road Using GPS-Monitored Data Based on Deep Learning Approach. *Sustainability*, *14*(10), 6351. https://doi.org/10.3390/su14106351
12. Afrin, T., & Yodo, N. (2021). A Long Short-Term Memory-based correlated traffic datapredictionframework.*KnowledgeBasedSystems*,107755.https://doi.org/10.1016/J.KNOSYS.2021.107755
13. Ma, X., Zhong, H., Li, Y., Ma, J., Cui, Z., & Wang, Y. (2021). Forecasting Transportation Network Speed Using Deep Capsule Networks With Nested LSTM Models. *IEEE Transactions on Intelligent Transportation Systems*, *22*(8), 4813–4824. https://doi.org/10.1109/TITS.2020.2984813
14. Abduljabbar, R. L., Dia, H., & Tsai, P.-W. (2021). Unidirectional and Bidirectional LSTM Models for Short-Term Traffic Prediction. *Journal of Advanced Transportation*, *2021*, 1–16. https://doi.org/10.1155/2021/5589075
15. Wang, P., Zhang, Y., Hu, T., & Zhang, T. (2022). Urban traffic flow prediction: a dynamic temporal graph network considering missing values. *International Journal of*

GeographicalInformationScience,*37*(4),885–912.https://doi.org/10.1080/13658816.2022.2146120

16. Shao, Y., Zhao, Y., Yu, F., Zhu, H., & Fang, J. (2021). The Traffic Flow Prediction Method Using the Incremental Learning-Based CNN-LTSM Model: The Solution of MobileApplication.*MobileInformationSystems*,*2021*,1–16.https://doi.org/10.1155/2021/5579451
17. Bandara, K., Bergmeir, C., & Hewamalage, H. (2021). LSTM-MSNet: Leveraging Forecasts on Sets of Related Time Series With Multiple Seasonal Patterns. *IEEE TransactionsonNeuralNetworks*,*32*(4),1586–1599.https://doi.org/10.1109/TNNLS.2020.2985720
18. Al-Sultan, A. T., Al-Rubkhi, A., Alsaber, A., & Pan, J. (2021). Forecasting air passenger traffic volume: evaluating time series models in long-term forecasting of Kuwait air passenger data. *Advances and Applications in Statistics*, *70*(1), 69–89. https://doi.org/10.17654/AS070010069
19. Wu, D. C. W., Ji, L., He, K., & Tso, K. F. (2021). Forecasting Tourist Daily Arrivals With A Hybrid Sarima–Lstm Approach. *Journal of Hospitality & Tourism Research*, *45*(1), 52–67. https://doi.org/10.1177/1096348020934046
20. Duan, C., & Li, X. (2021). *Multi-target Tracking Based on Deep Sort in Traffic Scene*. *1952*(2), 022074. https://doi.org/10.1088/1742-6596/1952/2/022074

CHAPTER 5

1. Lin, L., Xu, Z., & Wu, D. (2023). Real Time Vehicle Tracking Method Based on YOLOv5 + DeepSORT. *Computational Intelligence and Neuroscience*, *2023*, 1–11. https://doi.org/10.1155/2023/7974201
2. Jing, W.-Y., Zhu, Z.-J., Bai, Y.-Q., Li, L., Cui, W.-F., & Yu, W.-B. (2023). Violation Behavior Detection for Non-motor Vehicles. *Diannao Xuekan*, *34*(1), 175–186. https://doi.org/10.53106/199115992023023401013
3. *Parking Time Violation Tracking using Yolov8 and DeepSORT*. (2023). https://doi.org/10.20944/preprints202305.0828.v1
4. S, C., & K, P. (2023). Application of various deep learning models for automatic traffic violation detection using edge computing. *International Journal of Advanced Information Technology*, *13*(1/2), 01–11. https://doi.org/10.5121/ijitcs.2023.13201
5. Yadav, V. K., Yadav, P., & Sharma, S. (2021). *An Efficient Road Surveillance Approach to Detect, Recognize & Tracking Vehicles Using Deep Learning Methods*. 503–512. https://doi.org/10.32628/CSEIT2174106
6. Perera, I., Senavirathna, S., Jayarathne, A., Egodawela, S., Godaliyadda, R., Ekanayake, P., Wijayakulasooriya, J., Herath, V., & Sathyaprasad, S. (2021). Vehicle Tracking based on an Improved DeepSORT Algorithm and the YOLOv4 Framework. *InternationalConferenceonInformationandAutomation*.https://doi.org/10.1109/ICIAFS52090.2021.9606052

7. Liu, H., Pei, Y., Bei, Q., & Deng, L. (2022). Improved DeepSORT Algorithm Based on Multi-Feature Fusion. *Applied System Innovation*, *5*(3), 55. https://doi.org/10.3390/asi5030055
8. Yan, J., Wang, Z., Li, Y., Chen, Z., Chen, X., & Lou, L. (2021). Detecting and Tracking the Moving Vehicles Based on Deep Learning. *UK Workshop on Computational Intelligence*, 369–378. https://doi.org/10.1007/978-3-030-87094-2_32
9. Xu, Y., & Wang, J. (2019). A unified neural network for object detection, multiple object tracking and vehicle re-identification. *arXiv: Computer Vision and Pattern Recognition*. https://dblp.uni-trier.de/db/journals/corr/corr1907.html#abs-1907-03465
10. Lakshmanan, L., Vora, Y., & Ghate, R. (2020). Deep Learning Based Vehicle Tracking System Using License Plate Detection And Recognition. *arXiv: Computer VisionandPatternRecognition*.https://dblp.uni-trier.de/db/journals/corr/corr2005.html#abs-2005-08641
11. Sharma, D., & Jaffery, Z. A. (2021). Categorical Vehicle Classification and Tracking using Deep Neural Networks. *International Journal of Advanced Computer Science and Applications*, *12*(9). https://doi.org/10.14569/IJACSA.2021.0120964
12. Tiezzi, M., Melacci, S., Maggini, M., & Frosini, A. (2018). *Video Surveillance of Highway Traffic Events by Deep Learning Architectures* (Vol. 11141, pp. 584–593). Springer, Cham. https://doi.org/10.1007/978-3-030-01424-7_57
13. Video object tracking based on YOLOV7 and Deepsort. (2022).https://doi.org/10.48550/arxiv.2207.12202
14. Chen, X., Jia, Y., Tong, X., & Li, Z. (2022). Research on Pedestrian Detection and DeepSort Tracking in Front of Intelligent Vehicle Based on Deep Learning. *Sustainability*, *14*(15), 9281. https://doi.org/10.3390/su14159281
15. Fouladgar, M., Parchami, M., Elmasri, R., & Ghaderi, A. (2017). Scalable deep traffic flow neural networks for urban traffic congestion prediction. *International Joint Conference on Neural Network*, 2251–2258. https://doi.org/10.1109/IJCNN.2017.7966128
16. Yang, B., Tang, M., Chen, S., Wang, G., Tan, Y., & Li, B. (2020). A vehicle tracking algorithm combining detector and tracker. *Eurasip Journal on Image and Video Processing*, *2020*(1), 1–20. https://doi.org/10.1186/S13640-020-00505-7
17. Fouladgar, M., Parchami, M., Elmasri, R., & Ghaderi, A. (2017). *Scalable Deep TrafficFlowNeuralNetworksforUrbanTrafficCongestionPrediction*.https://doi.org/10.1109/IJCNN.2017.7966128
18. Ambekar, S., Dagade, M., Yadav, A. S., & Burande, S. (2019). Traffic signal violation monitoring through video surveillance using CNN. *International Journal of Advanced ResearchandInnovativeIdeasinEducation*,*5*(3),1498–1505.https://www.ijariie.com/FormDetails.aspx?MenuScriptId=149325
19. Mauri, A., Mauri, A., Khemmar, R., Decoux, B., Ragot, N., Rossi, R., Trabelsi, R., Boutteau, R., Ertaud, J.-Y., & Savatier, X. (2020). Deep Learning for Real-Time 3D Multi-Object Detection, Localisation, and Tracking: Application to Smart Mobility. *Sensors*, *20*(2), 532. https://doi.org/10.3390/S20020532
20. Paul, N., Saha, D., Ahmed, T., & Mahmud, R. (2023). Detection of Traffic Rule Violations Using Machine Learning: An Analytical Review. *Malaysian Journal of Science and Advanced Technology*, 37–47. https://doi.org/10.56532/mjsat.v3i1.146

CHAPTER 6

1. S, C., & K, P. (2023). Application of various deep learning models for automatic traffic violation detection using edge computing. *International Journal of Advanced Information Technology*, *13*(1/2), 01–11. https://doi.org/10.5121/ijitcs.2023.13201
2. Deshmukh, S. J., Atole, H. A., Rangwal, A. S., Kokil, B. J., & Saraswat, S. (2020). Traffic Rule Violation Detection System using ML. *International Journal of Advance ResearchandInnovativeIdeasinEducation*,*6*(5),410–416.https://ijariie.com/AdminUploadPdf/Traffic_Rule_Violation_Detection_System_using_ML_ijariie12688.pdf
3. M, A., Joseph, A., Vinu, F. B., Mammen, S. G., & J S, P. (2022). Traffic Violation Data Security System. *Journal of Machine and Computing*, 87–93. https://doi.org/10.53759/7669/jmc202202012
4. Zhang, M. (2018). Research on Intelligent Recognition System for Traffic Violation. *DEStechTransactionsonComputerScienceandEngineering*. https://doi.org/10.12783/DTCSE/ICEITI2017/18921
5. Agrawal, D. R., Kasliwal, S., Labhesh, R. G., Punam, D. D., & Pooja, S. (2017). Android based Traffic Rules Violation Detection System "Vehitrack." *International JournalofComputerApplications*,*163*(8),1–4.https://doi.org/10.5120/IJCA2017913634
6. Ambekar, S., Dagade, M., Yadav, A. S., & Burande, S. (2019). Traffic signal violation monitoring through video surveillance using CNN. *International Journal of Advance Research and Innovative Ideas in Education*, *5*(3), 1498–1505. https://www.ijariie.com/FormDetails.aspx?MenuScriptId=149325
7. Abdel-Wahab, M. A., Abdel-Nasser, M., & Hori, M. (2020). Reliable and Rapid Traffic Congestion Detection Approach Based on Deep Residual Learning and MotionTrajectories.*IEEEAccess*,*8*,182180–182192.https://doi.org/10.1109/ACCESS.2020.3028395
8. Cai, Z., Pei, W., Zhu, Y., Lu, M., & Wu, L. (2017). *Inferring and Analysis Drivers Violation Behavior Through Trajectory* (pp. 295–304). Springer, Cham. https://doi.org/10.1007/978-3-319-68121-4_32
9. Ou, G., Gao, Y., & Liu, Y. (2012). Real-Time Vehicular Traffic Violation Detection in Traffic Monitoring Stream. *Web Intelligence/IAT Workshops*, 15–19. https://dblp.uni-trier.de/db/conf/iat/iatw2012.html#OuGL12
10. Vijverberg, J. A., de Koning, N. A. H. M., Han, J., & Cornelissen, D. (2007). High-Level Traffic-Violation Detection for Embedded Traffic Analysis. *International Conference on Acoustics, Speech, and Signal Processing*, *2*, 793–796. https://doi.org/10.1109/ICASSP.2007.366355
11. Jiang, Z., Chen, L., Zhou, B., Huang, J., Xie, T., Fan, X., & Wang, C. (2021). iTV: Inferring Traffic Violation-Prone Locations With Vehicle Trajectories and Road EnvironmentData.*IEEESystemsJournal*,*15*(3),3913–3924.https://doi.org/10.1109/JSYST.2020.3012743
12. B. -H. Chen and S. -C. Huang, "An Advanced Moving Object Detection Algorithm for Automatic Traffic Monitoring in Real-World Limited Bandwidth Networks," in

IEEE Transactions on Multimedia, vol. 16, no. 3, pp. 837-847, April 2014, doi: 10.1109/TMM.2014.2298377.
13. Le, Q. T., Duong, D. C., Nguyen, T. A., Pham, M. A., Ha, T. M. D., & Nguyen, T. M. (2022). Automatic Traffic Red-Light Violation Detection Using AI. *Ingénierie Des Systèmes D'information*, *27*(1), 75–80. https://doi.org/10.18280/isi.270109
14. Hoogendoorn, S. P., & van Zuylen, H. J. (2004). Tracing traffic dynamics with remote sensing. *WIT Transactions on Information and Communication Technologies*, *32*. https://doi.org/10.2495/MIS040301
15. Pukkawanna, S., Hazeyama, H., Kadobayashi, Y., & Yamaguchi, S. (2015). Detecting Anomalies in Massive Traffic Streams Based on S-Transform Analysis of Summarized Traffic Entropies. *IEICE Transactions on Information and Systems*, *98*(3), 588–595. https://doi.org/10.1587/TRANSINF.2014NTP0006
16. Conti, M., Mancini, L. V., Spolaor, R., & Verde, N. V. (2014). Can't you hear me knocking: Identification of user actions on Android apps via traffic analysis. *arXiv: Cryptography and Security*. https://arxiv.org/abs/1407.7844
17. Ananthanarayanan, G., Bahl, V., Shu, Y., Loewenherz, F., Lai, D., Akers, D., Cao, P., Xia, F., Zhang, J., & Song, A. (2019). *Traffic Video Analytics – Case Study Report*. https://www.microsoft.com/en-us/research/publication/traffic-video-analytics-case-study-report/
18. J. Kline, S. Nam, P. Barford, D. Plonka and A. Ron, "Traffic Anomaly Detection at Fine Time Scales with Bayes Nets," *2008 The Third International Conference on Internet Monitoring and Protection*, Bucharest, Romania, 2008, pp. 37-46, doi: 10.1109/ICIMP.2008.33.
19. Delibasoglu, I. (2022). A Real-Time Video Surveillance for Rule Violation Detection on HighWays. *The Eurasia Proceedings of Science Technology Engineering and Mathematics*, *19*, 15–20. https://doi.org/10.55549/epstem.1218703
20. Bilakeri, S., & A K, K. (2022). Multi-object tracking by multi-feature fusion to associatealldetectedboxes.*CogentEngineering*,*9*(1).https://doi.org/10.1080/23311916.2022.2151553

CHAPTER 7

1. *CityTrack: Improving City-Scale Multi-Camera Multi-Target Tracking by Location-AwareTrackingandBox-GrainedMatching*.(2023). https://doi.org/10.48550/arxiv.2307.02753
2. Sun, Y., Zhao, Y., & Wang, S. (2022). Multiple Traffic Target Tracking with Spatial-Temporal Affinity Network. *Computational Intelligence and Neuroscience*, *2022*, 1–13. https://doi.org/10.1155/2022/9693767
3. Delavarian, M., Marouzi, O. R., Hassanpour, H., Parizi, R. M., & Khan, M. S. (2020). Multi-camera multiple vehicle tracking in urban intersections based on multilayer graphs.*IetIntelligentTransportSystems*,*14*(12),1673–1690.https://doi.org/10.1049/IET-ITS.2020.0086

4. Abri, R. (2023). *Data-Driven Optimization of Urban Traffic using AI and Real-Time Analysis*. *1*, 507–514. https://doi.org/10.59287/icpis.881
5. Wang, Q., Lu, M., & Li, Q. (2020). Interactive, multiscale urban-traffic pattern exploration leveraging massive GPS trajectories. *Sensors*, *20*(4), 1084. https://doi.org/10.3390/S20041084
6. Ooi, H.-L., Bilodeau, G.-A., Saunier, N., & Beaupre, D.-A. (2018). *Multiple Object Tracking in Urban Traffic Scenes with a Multiclass Object Detector* (pp. 727–736). Springer, Cham. https://doi.org/10.1007/978-3-030-03801-4_63
7. Han, B. (2023). Multi-Object Multi-Camera Tracking Based on Deep Learning for IntelligentTransportation:AReview.*Sensors*,*23*(8),3852.https://doi.org/10.3390/s23083852
8. Karunasekera, H., Wang, H., & Zhang, H. (2019). Multiple Object Tracking With Attention to Appearance, Structure, Motion and Size. *IEEE Access*, *7*, 104423–104434. https://doi.org/10.1109/ACCESS.2019.2932301
9. Ooi, H.-L., Bilodeau, G.-A., Saunier, N., & Beaupre, D.-A. (2018). Multiple Object Tracking in Urban Traffic Scenes with a Multiclass Object Detector. *arXiv: Computer Vision and Pattern Recognition*. https://arxiv.org/abs/1809.02073
10. Zhang, W., Zhou, H., Sun, S., Wang, Z., Shi, J., & Loy, C. C. (2019). Robust Multi-Modality Multi-Object Tracking. *International Conference on Computer Vision*, 2365–2374. https://doi.org/10.1109/ICCV.2019.00245
11. Danzer, A., Gies, F., & Dietmayer, K. (2018). *Multi-Object Tracking with Interacting Vehicles and Road Map Information*. https://doi.org/10.1109/ITSC.2018.8569701
12. Yang, Y., & Bilodeau, G.-A. (2017). Multiple Object Tracking with Kernelized Correlation Filters in Urban Mixed Traffic. *Canadian Conference on Computer and Robot Vision*, 209–216. https://doi.org/10.1109/CRV.2017.18
13. Yin, G., Liu, B., Zhu, H., Gong, T., & Yu, N. (2019). A Large Scale Urban Surveillance Video Dataset for Multiple-Object Tracking and Behavior Analysis. *arXiv:ComputerVisionandPatternRecognition*.https://dblp.uni-trier.de/db/journals/corr/corr1904.html#abs-1904-11784
14. Yang, Y., & Bilodeau, G.-A. (2016). Multiple Object Tracking with Kernelized Correlation Filters in Urban Mixed Traffic. *arXiv: Computer Vision and Pattern Recognition*. https://arxiv.org/abs/1611.02364
15. Kampker, A., Sefati, M., Abdul Rachman, A. S., Kreisköther, K., & Campoy, P. (2018). *Towards Multi-Object Detection and Tracking in Urban Scenario under Uncertainties*. https://doi.org/10.5220/0006706101560167
16. Schütz, M., Appenrodt, N., Dickmann, J., & Dietmayer, K. (2014). Multiple extended object tracking with object-local occupancy grid maps. *International Conference on InformationFusion*,1–7.https://dblp.uni-trier.de/db/conf/fusion/fusion2014.html#SchutzADD14
17. Dyckmanns, H., Matthaei, R., Maurer, M., & Lichte, B. (2011). Object tracking in urban intersections: Active interacting multi model filter with handling of uncertaintiesofmapmatching.*InternationalConferenceonInformationFusion*,1–8.https://ieeexplore.ieee.org/abstract/document/5977488

18. Nageswaran, G. (2018). *Integrated Multi-Object Tracking and Classification for VehicleEnvironmentPerception*.https://kluedo.ub.uni-kl.de/frontdoor/index/index/docId/5401.pdf
19. Hsu, H.-M., Wang, Y., & Hwang, J.-N. (2020). Traffic-Aware Multi-Camera Tracking of Vehicles Based on ReID and Camera Link Model. *ACM Multimedia*, 964–972. https://doi.org/10.1145/3394171.3413863
20. Liu, J., & Chen, X. (2024). Analysis of pedestrian-two-wheeler conflicts at green light digital countdown signals: A random parameter ordered logit model approach. *AccidentAnalysis&Prevention*,*206*,107716. https://doi.org/10.1016/j.aap.2024.107716

CHAPTER 8

1. Song, H., Li, J., Wan, F., & Hong, J. (2023). *A Real-Time Pedestrian Tracking Method Based on Enhanced Tracktor*. 812–817. https://doi.org/10.1145/3650400.3650537
2. Han, C., Seshadri, P., Ding, Y., Posner, N., Koo, B. W., Agrawal, A., Lerch, A., & Guhathakurta, S. (2024). Understanding pedestrian movement using urban sensing technologies: the promise of audio-based sensors. *Urban Informatics*, *3*(1). https://doi.org/10.1007/s44212-024-00053-9
3. Clayton, S.,Urquhart,N. & Kerridge, J. M. (2007). *Tracking and analysis of the movementofpedestrians*.https://www.starconference.org.uk/star/2007/Sarah_Clayton.pdf
4. V, V., Arkatkar, S. S., Joshi, G., & Parida, M. (2020). Examining Pedestrian Critical Gap Analysis at Unsignalized Midblock Crosswalk Sections in India. *Transportation Research Procedia*, *48*, 2230–2250. https://doi.org/10.1016/J.TRPRO.2020.08.280
5. Kadali, B. R., & Perumal, V. (2016). Critical Gap Analysis based on Pedestrian Behavior during Peak-Hour Traffic at Unprotected Midblock Crosswalks. *Asian Transport Studies*, *4*(1), 261–277. https://doi.org/10.11175/EASTSATS.4.261
6. Abou-Senna, H., Radwan, E., & Mohamed, A. (2016). *Investigating the Correlation Between Sidewalk Gaps and Pedestrian Safety*. https://trid.trb.org/view/1393874
7. Kadali, B. R., & Vedagiri, P. (2019). Evaluation of pedestrian accepted vehicle gaps with varied roadway width under mixed traffic conditions. *Transportation Letters: The International Journal of Transportation Research*, *11*(9), 527–534. https://doi.org/10.1080/19427867.2017.1408204
8. V, V., Arkatkar, S. S., Joshi, G., & Parida, M. (2020). Examining Pedestrian Critical Gap Analysis at Unsignalized Midblock Crosswalk Sections in India. *Transportation Research Procedia*, *48*, 2230–2250. https://doi.org/10.1016/J.TRPRO.2020.08.280
9. Chandra, S., Rastogi, R., & Das, V. R. (2014). Descriptive and parametric analysis of pedestrian gap acceptance in mixed traffic conditions. *Ksce Journal of Civil Engineering*, *18*(1), 284–293. https://doi.org/10.1007/S12205-014-0363-Z
10. Radwan, E., Abou-Senna, H., Mohamed, A., Navarro, A., Minaei, N., Wu, J., & Gonzalez, L. (2016). *Assessment of Sidewalk/Bicycle-Lane Gaps with Safety andDevelopingStatewidePedestrianCrashRates*.https://rosap.ntl.bts.gov/view/dot/30998

11. Shi, J., Wang, Z., & Chen, Y. (2007). *Analysis on behaviours and safety of VRUs at unsignalized roadway crosswalk*. https://trid.trb.org/view.aspx?id=868935
12. Gashaw, S., Goatin, P., & Härri, J. (2018). Modeling and Analysis of Mixed Flow of Cars and Powered Two Wheelers. *Transportation Research Part C-Emerging Technologies*, *89*, 148–167. https://doi.org/10.1016/J.TRC.2018.02.004
13. Budhkar, A. K., & Maurya, A. K. (2017). Modeling Gap-Maintenance in Heterogeneous No Lane-Discipline Traffic. *DEStech Transactions on Engineering and Technology Research*. https://doi.org/10.12783/DTETR/ICTIM2016/5543
14. Zhang, X., Ryan, A., & Wu, Y.-J. (2024). *Influential factors of pedestrian and bicycle crashes near Pedestrian Hybrid Beacons: Observing trends through an applied analysis*. https://doi.org/10.1080/19439962.2024.2325399
15. van Elslande, P., Fournier, J.-Y., & Parraud, C. (2015). Powered Two-wheelers In Urban Environment: A Detailed Accident Analysis. *International Journal of Safety andSecurityEngineering*,*5*(4),322–335. https://doi.org/10.2495/SAFE-V5-N4-322-335
16. Sridhar, V. H., Sridhar, V. H., Roche, D. G., Gingins, S., & Gingins, S. (2019). Tracktor: Image-based automated tracking of animal movement and behaviour. *MethodsinEcologyandEvolution*,*10*(6),815–820.https://doi.org/10.1111/2041-210X.13166
17. Song, H., Li, J., Wan, F., & Hong, J. (2023). *A Real-Time Pedestrian Tracking Method Based on Enhanced Tracktor*. 812–817. https://doi.org/10.1145/3650400.3650537
18. Okamoto, Y., Matsunami, I., & Kajiwara, A. (2012). Pedestrian and two-wheeler detection using ultra-wideband vehicular radar. *Sensors Applications Symposium*, 1–4. https://doi.org/10.1109/SAS.2012.6166289
19. Al Bargi, W. A., Md Rohani, M., Daniel, B. D., Khalifa, N. A., Mohd Masirin, M. I., & Kironde, J. (2023). Estimating of Critical Gaps at Uncontrolled Intersections under Heterogeneous Traffic Conditions. *Automotive Experiences*, *6*(2), 429–437. https://doi.org/10.31603/ae.9406
20. Abri, R. (2023). *Data-Driven Optimization of Urban Traffic using AI and Real-Time Analysis*. *1*, 507–514. https://doi.org/10.59287/icpis.881