Vehicle and Human Detection utilizing YOLOv8: A Case Study in Dhaka.

BY

Sajib Bormon
ID: 201-15-3773

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science and Engineering

Supervised By

Mr. Amir Sohel
Senior Scale
Department of CSE
Daffodil International University

Co-Supervised By

Ms. Aliza Ahmed Khan
Senior Scale
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

JANUARY 2024

APPROVAL

This Project titled **"Vehicle and Human Detection utilizing YOLOv8: A Case Study in Dhaka"**, submitted by **Sajib Bormon,** ID:201-15-3773 to the Department of Computer Science and Engineering, Daffodil International University, has been accepted as the satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 22 January, 2024.

BOARD OF EXAMINERS

Narayan Ranjan Chakraborty (NRC) — **Chairman**
Associate Professor and Associate Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Saiful Islam (SI) — **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Lamia

Lamia Rukhsara (LR) — **Internal Examiner**
Senior Lecturer
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Dr. Md. Zulfiker Mahmud (ZM) — **External Examiner**
Associate Professor
Department of Computer Science and Engineering
Jagannath University

DECLARATION

I hereby, declare that, I have done this project under the supervision of **Amir Sohel, Senior Lecturer, Department of CSE** Daffodil International University. I also declare that neither this project nor any part of this project elsewhere for the award of any degree or diploma.

Supervised by:

Mr. Amir Sohel
Senior Lecturer
Department of CSE
Daffodil International University

Co-Supervised by:

Ms. Aliza Ahmed Khan
Senior Lecturer
Department of CSE
Daffodil International University

Submitted by:

Sajib Bormon
ID: 201-15-3773
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First, I express our heartiest thanks and gratefulness to almighty God for His divine blessing, which made it possible to complete the final year project successfully.

I am really grateful and wish my profound indebtedness to **Amir Sohel**, **Senior Lecturer**, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & and keen interest of my supervisor in Computer Vision to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts and correcting them at all stages have made it possible to complete this project.

I also want to express my heartiest gratitude to **Dr. Sheak Rashed Haider Noori, Head of Department of CSE**, for his kind help to finish my project and to other faculty members and the staff of the CSE department of Daffodil International University.

I want to thank my entire Daffodil International University coursemates, who participated in this discussion while completing the course work.

Finally, I must acknowledge with due respect the constant support and patience of my parents.

ABSTRACT

Transportation problems happen in cities because they are so inherent in diversity and always changing. The result is that traffic congestion, accidents, and the economic burden of these things are extremely difficult to deal with in the modern day. The economy is slowed down due to these impediments, and the safety of individuals is put in jeopardy. As cities continue to grow and concerns about traffic and congestion continue to rise, detecting vehicles and humans can provide a potential route to a more effective solution. YOLOv8n, YOLOv8s, YOLOv8m, and YOLOv8l are the four variants of YOLOv8 that have been utilized in this study to investigate how well they function in detecting vehicles and persons. It has been accomplished by collecting a fresh dataset of Dhaka's streets. The purpose of this research is to address the problem mentioned above. The collection of data includes images of humans as well as six different types of vehicles. By gaining a high mAP of 0.909 at IoU 50, the results demonstrate that YOLOv8l is successful in detecting vehicles and humans in demanding traffic settings. Additionally, the results address the effectiveness of the dataset. Based on this accuracy, it appears that the model has the potential to be effectively utilized for a wide range of purposes, including the automation of traffic enforcement, the improvement of traffic flow, the reduction of congestion, the prevention of accidents, and the preparation of the groundwork for autonomous vehicles in Bangladesh.

Keywords: YOLOv8, Vehicle and Human Detection, Automated Traffic Enforcement, Autonomous Vehicles.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Introduction

Traffic congestion in urban places is a growing issue worldwide, with Dhaka, the capital of Bangladesh, illustrating this global crisis. Dhaka's streets are renowned for their continuous traffic congestion, resulting in substantial economic losses and severe safety dangers. In the previous ten years, average driving speeds of vehicles decreased from 21 kilometers to 6 kilometers per hour. If the current trend persists, it might decline to barely 4 kilometers per hour by 2035, slower than the average walking pace. This congestion destroys 3.2 million working hours daily and imposes billions of dollars in economic losses annually.

In addition, traffic congestion has been a major factor in many accidents, which have resulted in injuries and fatalities. This is a highly concerning development. According to some of the most recent figures, Bangladesh has a significant fatality rate of 15.3 per 100,000 persons owing to injuries sustained in cases of road traffic accidents. The majority of accidents being investigated at the moment suggest that 66.4% of people are hurt, and 33.6% of people are dead. This scenario highlights the urgent need for creative ideas to modernize traffic management and boost road safety in Dhaka and other similar urban settings.

When it comes to overcoming these challenges, especially vehicle detection, YOLO is absolutely essential. Due to human memory's unreliable nature, manual automobile inspections are complex and prone to errors. In order to solve this problem, it is possible to design a system that is entirely integrated, such as models based on computer vision. The fact that this system can be administered and modified from a central location makes it more efficient, lowering the number of accidents and the number of fatalities. In artificial intelligence (AI), computer vision and machine learning techniques offer promising solutions for enhancing traffic monitoring and increasing road safety. YOLOv8 (which has seen the most significant alterations since its beginning) demonstrates the best performance in terms of detection time, accuracy considering the frames that need to be processed to enable real-time usage, and ease of implementation. Furthermore, intersections based on vehicle identification can observe the traffic density by employing deep learning with the YOLOv8 model. Adopting AI (YOLOv8 models) technology is a move toward intelligent city solutions for a city like Dhaka, where traditional traffic control systems are increasingly ineffective.

These solutions simplify traffic and lessen traffic-related events' frequency and severity, saving lives and decreasing economic damage.

Targeting the intricate traffic difficulties faced by Dhaka, this research leverages advanced computer vision models, including YOLOv8n, YOLOv8s, YOLOv8m, and YOLOv7l, to develop more efficient and intelligent traffic management solutions. By collecting 710 images and then applying several augmentation techniques, the final dataset of 2045 images (the dataset went through augmentation, including brightness adjustment, artificial rain simulation, and horizontal flipping, to mimic real-world variability in traffic conditions.) from Dhaka's diverse traffic scenarios that analyze the effectiveness of YOLOv8's four variants in real-world situations. The dataset consists of 6 categories of vehicles and human images of Bangladesh. This study includes vehicle detection suited to the specific traffic environment of Dhaka. The findings offer crucial insights into the feasibility and benefits of deep learning-based traffic management systems, setting an expectation for further study and usage in intelligent urban transportation solutions.

1.2 Motivation

Adopting car and pedestrian detection systems in Dhaka city is vital to handle the multiplicity of issues and difficulties contributing to severe traffic congestion. Firstly, Dhaka's fast urbanization has led to an increase in the number of vehicles on its roadways, resulting in frequent gridlocks and longer commuting times. Secondly, a lack of proper traffic management and law enforcement has allowed for chaotic and frequently unsafe driving practices, worsening the traffic situation. Thirdly, the absence of reliable public transit alternatives has led a substantial percentage of the population to rely on private vehicles, further overtaking the roadways.

Moreover, poor road infrastructure, inadequate parking facilities, and encroachment on roadways increase traffic difficulties. Frequent road repairs and construction projects, often carried out without sufficient planning, lead to further bottlenecks and difficulties. Additionally, the uncontrolled rise of informal settlements and the accompanying haphazard town extension result in issues in urban planning and traffic flow management.

Pedestrian safety is also a key concern, as a lack of authorized pedestrian crossings and walkways puts the lives of pedestrians at risk. Furthermore, air pollution levels in Dhaka have skyrocketed due to traffic congestion, significantly harming the health and well-being of its population.

In this context, adopting vehicle and pedestrian detection systems can aid in controlling traffic, preventing accidents, enforcing traffic laws, identifying pollution sources, and enhancing urban planning. These tools can give data-driven insights and real-time solutions to address these challenges comprehensively, developing a safer and more efficient transportation environment for Dhaka's citizens. However, hurdles such as initial infrastructure expenses, technical adaptation, and public awareness must be overcome to apply these solutions effectively in Dhaka.

1.3 Rationale of the Study

Traffic Congestion: Dhaka city suffers significant traffic congestion due to fast urbanization, additional vehicles, and a lack of proper traffic management. This situation significantly affects the daily lives of its citizens.

Chaotic Driving Practices: The absence of sufficient law enforcement and chaotic driving practices add to the traffic problem and cause safety dangers on the road.

Dependency on Private vehicles: A considerable section of the population relies on personal vehicles due to poor public transit choices, leading to traffic congestion.

Infrastructure Challenges: Frequent road repairs and construction without sufficient planning generate further bottlenecks.

Urban Planning issues: The uncontrolled growth of informal settlements and unplanned city expansion offer urban planning and traffic flow management issues.

Pedestrian Safety: A lack of defined pedestrian crossings and sidewalks endangers the lives of pedestrians, making pedestrian safety an urgent issue.

Air Pollution: Traffic congestion has led to significant levels of air pollution in Dhaka, severely harming the health and well-being of its citizens.

Potential actions: By offering data-driven insights and real-time solutions, deploying vehicle and pedestrian detection systems can assist in solving these difficulties. These systems can offer traffic management, accident reduction, law enforcement, pollution control, and urban planning optimization.

1.4 Research Questions

Before we even begin with the study, there are questions that are important to this work. The following questions are the ones I noticed:

1.4.1 How does the degree of traffic congestion in Dhaka affect the everyday lives and productivity of its residents?

1.4.2 What are the primary factors that lead to chaotic and unsafe driving habits in Dhaka, and how may vehicle detection systems play a role in these issues?

1.4.3 To what degree does the absence of adequate public transit choices impact the dependence on private cars in Dhaka, and how may surveillance methods assist in lowering private vehicle usage?

1.4.4 How has traffic congestion influenced air pollution in Dhaka, and how may vehicle detection systems aid in detecting pollution sources and lowering pollution levels?

1.4.5 What are the difficulties linked to road infrastructure, parking facilities, and highway encroachments that worsen traffic congestion in Dhaka?

1.4.6 How do regular road repairs and construction projects, carried out without sufficient planning, contribute to traffic bottlenecks, and what role may vehicle detection systems play in controlling such disruptions?

1.4.7 What are the urban planning problems related to the uncontrolled growth of informal settlements in Dhaka, and how could pedestrian and vehicle detection methods operate in urban planning and traffic flow management?

1.4.8 What is the existing condition of pedestrian safety in Dhaka, and how might adopting pedestrian detection systems improve pedestrian safety?

This study of vehicle and human detection systems in Dhaka city would give significant insights and solutions to relieve the urgent difficulties faced by the city. It will help boost Dhaka's population's quality of life by proposing practical approaches that reduce traffic congestion, enhance road safety, optimize urban planning, and reduce air pollution. Furthermore, the research will give insight into the practical viability and possible challenges of adopting these systems, allowing policymakers and stakeholders to make informed choices for sustainable growth and better well-being of the city's population.

1.5 Expected Output

This study of vehicle and human detection systems in Dhaka city would give significant insights and solutions to relieve the urgent difficulties faced by the city. It will help boost Dhaka's population's quality of life by proposing practical approaches that reduce traffic congestion, enhance road safety, optimize urban planning, and reduce air pollution. Furthermore, the research will give insight into the practical viability and possible challenges of adopting these systems, allowing policymakers and stakeholders to make informed choices for sustainable growth and better well-being of the city's population.

1.6 Project Management and Finance

Project management and finance: Regarding this research, I have initially made a plan for how many steps will be, and how I will manage the time and what are the risks. At first, I studied this topic and then found some scope. Then, based on that I have talked to my supervisor, and he guided me in this whole project.

There is not funding yet for this project. I and my supervisor have done this by ourselves. If we get any funding in the future, we will practically apply this research outcome, which will be far more than the project-based research work.

1.7 Report Layout

In the **Chapter 1** I have discussed about introduction and motivation behind this study. I have mentioned rationale of this study and some question that might arise withing this research. Moreover, I have talked about research outcome and project manage and finance related discussion.

Chapter 2 is the background study of my work. In this chapter, we have discussed my thoughts on this research and why I have chosen this issue by studying related works. I also have presented a summary of these problems related to traffic congestion in Dhaka. Scope of these related summaries and challenges that I have faced.

Chapter 3 In this chapter, I have talked about data collection and statical analysis and my proposed methodology. I also have mentioned the implementation requirements that is essential for this research.

In the **Chapter 4** first of all, I have talked about experimental setup and experiment results. Furthermore, I have shown a comparative analysis and discussed overall result.

Chapter 5, Here, I have talked about the impact of this research on the society and environment. Ethical aspects and sustainability plan have also mentioned.

Chapter 6, In this chapter I have shown the summary of this study. And conclusion and implication for further study also given.

CHAPTER 2

BACKGROUND

2.1 Preliminaries

Dhaka, the capital city of Bangladesh, is well known for its vibrant culture and economic significance. However, considerable issues are controlling its traffic because of the rapid increase of both urbanization and population. The traffic problems in Dhaka city result from an intricate chain of interrelated issues. Research conducted by the Bangladesh University of Engineering and Technology's Accident and Research Institute estimates that the daily hours lost due to traffic congestion are worth roughly Tk 140 crore. Compared to 5 million working hours per day in 2017, almost 8 million working hours were lost on Dhaka roadways per day in 2022. The study estimates that around 2.5 million short- and long-distance vehicle trips are undertaken in Dhaka daily, with 44% of the passengers traveling to and from offices. The number of privately owned cars has grown dramatically over the last few years, which has led to more traffic congestion. Private vehicle and motorcycle ownership has increased dramatically.

As a result, traffic congestion is getting worse rather than getting better for public transportation. In this instance, the number of private vehicles and the decrease in the registration of new buses are cited as the primary causes of the increased traffic congestion. Data from the Bangladesh Road Transport Authority shows that during the same period in 2020–2021, sales of jeeps increased by 55%, private automobiles by 28%, and motorbikes by 27%, but new bus registrations decreased by 32.3%. Also, the road accident in Bangladesh is higher compared to other countries. Figure 2.1.1 shows the number of accidents in Bangladesh from 2015 to 2023, and Figure 2.1.2 shows the number of accidents in Dhaka city from 2015 to 2022.

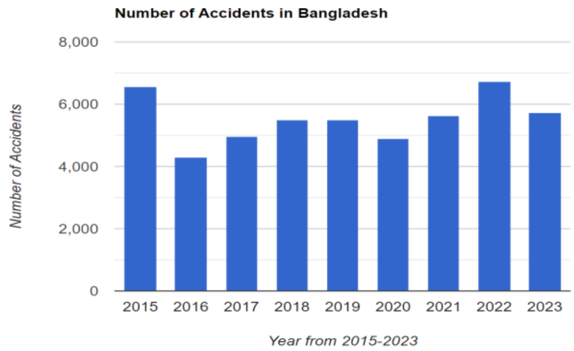

Figure 2.1.1 Number of Accidents from 2015 to 2023.

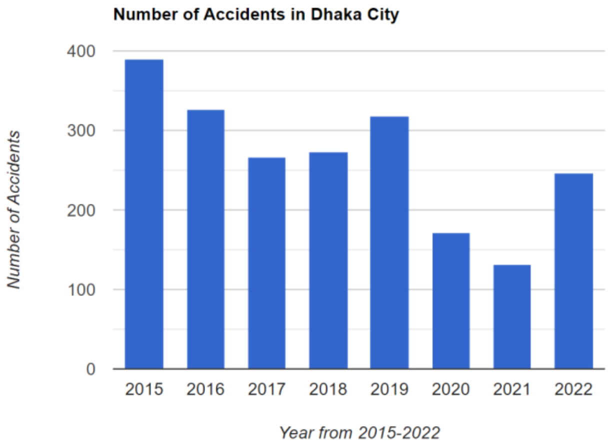

Figure 2.1.2 The number of accidents in Dhaka city from 2015 to 2022.

The annual financial losses resulting from traffic congestion in Dhaka as well as Bangladesh are significant, amounting to billions of dollars. In addition, road accidents cause tragedies involving people dying as well as additional financial loss due to increased medical expenses and missed productivity. In addition, these traffic challenges are not only necessary for the citizens of Dhaka, but also resolving the city's traffic problems is necessary for sustaining the country's economic growth.

2.2 Related Works

YOLO (You Only Look Once) is a sophisticated computer vision tool used for detecting cars. YOLO is a highly regarded detection technology widely recognized for accurately and quickly detecting numerous entities, such as vehicles, in images and videos. In this discussion, the research will examine vehicle detection utilizing the YOLO and other methodologies. The summary of the related work is provided below:

Telaumbanua et al. suggested a vehicle identification system utilizing the YOLOv8 algorithm, which reached a testing accuracy of 96% and a training accuracy of 77%. The teaching technique consisted of a dataset of 2042 images for training, 204 for validation, and 612 for testing. The dataset was meticulously chosen to capture the complexity of traffic conditions, drawing inspiration from the busy urban atmosphere of Medan City. Nevertheless, the research discovered challenges that limited identification accuracy, notably when dealing with occlusions and overlapping vehicles. The authors acknowledged these limitations and offered potential improvements for the future, including extending the dataset and employing sophisticated data augmentation methods. [1]

Maurya et al. focus on increasing autonomous vehicle perception by implementing YOLOv8 in a vehicle detection system. They used the Kaggle "Cars Detection" dataset, consisting of several vehicle sorts under varying driving scenarios. According to their study, YOLOv8 exceeded prior accuracy and processing speed models, including YOLOv5 and Faster R-CNN. The YOLOv8 model performed exceptionally well in essential performance criteria, with an accuracy of 0.491 mAP50. The research highlighted difficulties at complex urban junctions where the proximity of cars might hinder detection accuracy, highlighting this as a region for going model refining. Furthermore, it was highlighted that including bicycles and pedestrians in the model's detection repertoire was crucial to building a really effective perception system for autonomous vehicles in mixed-traffic metropolitan scenarios.[2]

Gozde and Ismail suggested using deep learning for road segmentation and pedestrian detection for advancing intelligent vehicle technology. Their method contains two subsystems: one for road segmentation using a Consecutive Triple Filter Size (CTFS) approach and another for pedestrian identification using the YOLOv7 network. They used the Cambridge-driving Labeled Video Database (CamVid) for road segmentation and the Pascal VOC dataset for pedestrian recognition. The CTFS approach gave a Jaccard index value of 95.84%, while the YOLOv7 network attained an average accuracy of 65.50% in pedestrian identification. The work suggests a viable research avenue in creating the system to identify road imperfections.[3]

Buitrón and Yoo focused on studying three object detection models: YOLOv8, RetinaNet, and SSD. The research collected 300 photos from diverse perspectives and conditions for real-time vehicle recognition and counting. It utilized the Research-Action conduct and performed a complete literature study for state-of-the-art analysis. YOLOv8 was praised for its top performance of .7901 mAP50 and precision of 0.6631 when detecting compared to SSD and RetinaNet, which were also assessed for accuracy and speed. The findings promote further research in blending models like RetinaNet with others for enhanced performance.[4]

Kumar and Muhammad worked on strengthening object detection for autonomous driving in severe weather settings. They applied two crucial datasets, ACDC and DAWN, and developed a new merged dataset, which was combined for better detection accuracy. They applied the YOLOv8 model, employing its increased object identification skills in several challenging weather circumstances. The research achieved an accuracy of 0.824 mAP50 on the combined dataset. However, the study demonstrated constraints in data gathering and the need for further refining of detection algorithms under diverse ambient situations. The research points to the likelihood of merging datasets with sophisticated models like YOLOv8 to enhance autonomous vehicle performance in harsh weather.[6]

Mahmod et al. presented a deep learning-based framework for real-time traffic monitoring in Bangladesh. It utilizes quicker R-CNN to detect cars, examine their speed, and identify forbidden lane changes. The system achieves an accuracy of 85% for both vehicle recognition and lane change detection, compared to MobileNet SSD. Future work involves extending the dataset, studying lightweight models, and integrating with traffic control infrastructure. [7]

Wang et al. provided the SSB-YOLO algorithm, an improved version of YOLOv8, for vehicle identification. The work uses the Pascal VOC dataset for training and validation. The SSB-YOLO model integrates the Shuffle Attention mechanism with the spatial and channel reconstruction convolution (SCConv) approach to improve detection performance. They compared the SSB-YOLO model with Faster-RCNN, SSD, YOLOvXs, YOLOv5n, YOLOv6n, and YOLOv8n. The study shows a 1.6% improvement in mean Average Precision (mAP@50) compared to the basic YOLOv8 model, which overperforms other models by 87.5%, suggesting that it is helpful in vehicle identification. The study addressed the need for more robust and accurate vehicle recognition algorithms in autonomous driving and intelligent transportation systems.[8]

Alamgir et al. examined alternative YOLO-based concepts for vehicle detection in Bangladesh using a combined dataset of 7390 photos (DhakaAI dataset and personal gathered). Their analysis shows that YOLOv5x beats YOLOv3 and YOLOv5s by 0.287 mAP, suggesting its superiority for real-time vehicle recognition in Bangladeshi traffic. The models may be improved for several classes because they encountered misclassification. Future work includes broadening the dataset and researching lightweight models to make additional improvements.[9]

Tanzim Mostafa et al. dealt with the challenge of occluded object detection for autonomous cars by presenting a new dataset of occluded road sceneries from Bangladesh to evaluate the performance of three prominent object detection models: YOLOv5, YOLOX, and Faster R-CNN. YOLOX obtained the best performance, with mAP scores of 0.849 at 0.5 and 0.634 at 0.5:0.95, suggesting its potential for improved perception in autonomous cars. The finding sets the door for future breakthroughs in object identification for safer and more reliable autonomous driving.[10]

Rafi et al. suggested a unique method for detecting South Asian vehicles using YOLOv5 models. Focusing on overcoming data scarcity, they construct a complete dataset of 21 vehicle classifications, including uncommon South Asian vehicles like rickshaws. Their investigation reveals that the YOLOv5 Large model exceeds its smaller equivalents, obtaining the best accuracy in vehicle recognition. This study shows the potential of YOLOv5 models for considerably enhancing traffic management and safety in South Asia.[11]

Rouf et al. developed a real-time vehicle detection, tracking, and counting system utilizing the YOLOv7 algorithm leveraging backbone MSSA, targeted at enhancing highway traffic management with an accuracy of 0.57 mAP50, the highest compared to other suggested models. Their technology performs better at recognizing, monitoring, and categorizing distinct vehicle kinds in real-time recordings. They recognize errors in earlier studies, especially in real-world settings with different variables, and suggest improvements in detection accuracy and flexibility. The research results have potential uses in transportation management, traffic monitoring, and surveillance, with future work focused on combining characteristics like speed estimates and lane identification.[12]

The paper by Zhang and Hu presents the MSFFA-YOLO network, which seeks to enhance multiclass object identification in foggy situations. They deployed upgraded YOLOv7 for object detection, including features like multiscale feature fusion and attention processes. The approach was evaluated on synthetic foggy datasets (FC005, FC01, FC02) and real foggy datasets (RTTS, RW), exhibiting accuracies of 64.6% on FC005, 67.3% on FC01, 65.7% on FC02, 84.7% on RTTS, and 84.1% on RW. MSFFA-YOLO was incredibly efficient, with 37 frames per second. The paper suggests further work to increase accuracy and adapt the approach for diverse traffic situations. [13]

2.3 Comparative Analysis and Summary

The summary of the literature review is given below in Table 2.3.1:

Table 2.3.1: Summary of the Literature Reviews.

Author	Dataset	Accuracy	Limitation
[1] Telaumbanua et al. [2023]	CCTV images of the city of Medan	0.77 mAP50 on training (YOLOv8)	Accuracy could be higher, and there is no practical implementation.
[2] Maurya et al. [2023]	"Cars Detection" Kaggle	0.491 mAP50 (YOLOv8)	Accuracy could be more.
[3] Gozde and Ismail [2023]	Pascal VOC dataset	0.65 AP (YOLOv7)	Could have used a new dataset.
[4] Buitrón and Yoo [2023]	Custom datasets	0.8654 mAP50 (RetinaNet + ResNet50)	The dataset could have more images.

[6] Kumar and Muhammad [2023]	ACDC and DAWN, and merged dataset	0.824 mAP50 (YOLOv8)	There is no practical implementation.
[7] Mahmod et al. [2022]	Custom dataset	0.85 mAP50 (FR-CNN)	Could have been tried on multiple models with more images.
[8] Wang et al. [2023]	Pascal VOC dataset	0.875 mAP50 (SSB-YOLO)	Could use a different dataset.
[9] Alamgir et al. [2022]	DhakaAI dataset and custom dataset	0.287 mAP (YOLOv5 X)	Could have less misclassification.
[10] Mostafa et al. [2022]	Dhaka Occluded Objects Dataset	0.849 mAP (YOLOX)	Trucks, vans are identified as others vehicle.
[11] Rafi et al. [2022]	Secondary	0.6687 mAP (YOLOv5l)	Imbalance dataset and accuracy could be more.

[12] Rouf et al. [2023]	Highway Traffic	0.57 mAP50 (YOLOv7)	Challenges in diverse real-world scenarios.
[13] Zhang and Hu [2023]	FC005, FC01, FC02, RTTS, and RW	64.6%-84.7%. (MSFFA-YOLO)	Could have adapted the method for various traffic scenarios.

Based on this literature review, most studies have shown utilizing YOLO models for their research and have acquired better accuracy than any other algorithm using their own or secondary dataset. In this study, I have collected a new dataset and used YOLOv8 models for my research.

2.4 Scope of the Problem

The study aims to concentrate on developing and implementing YOLOv8-based object detection models. Specifically, the project's primary purpose is to detect vehicles and humans within the environment of Dhaka city. Collecting an updated dataset customized to Dhaka city's unique features and problems. This dataset will serve as a vital basis for training and assessing your YOLOv8 models. The research's primary goal is to contribute to solutions that may solve real-world difficulties in Dhaka, including traffic management, pedestrian safety, parking management, and urban planning.

2.5 Challenges

While doing this research, I have faced many challenges. They are:

2.5.1 Dataset Collection: The dataset must be good to get decent accuracy. YOLO models learn from different scenarios. So, I have taken images from different angles; while doing this, I had to be in the middle of the road; taking perfect images from different angles was challenging.

2.5.2 Implementation tools: I have used the free Colab version. In this version, the GPU and the time limit are limited, so I have faced multiple GPU time limit errors. While training with higher resolution (imgsz = 1280), GPU memory exceeded. As a result, I have to train the model using a resolution of 640.

CHAPTER 3

RESEARCH

METHODOLOGY

3.1 Research subject and instrumentation

This study aims to investigate how well YOLO-based vehicle detection might be used to enhance traffic management. This section gives a complete overview of the methodology used, which covers the following steps: gathering the dataset, annotating the images, expanding the number of images via augmentation, selecting models, training the models, and evaluating and comparing the performance of the models.

3.2 Data Collection Procedure

3.2.1. Data collection: In this study, the development of a comprehensive dataset is an essential aspect. Seven hundred and ten images were taken in a wide variety of settings and situations around the city of Dhaka. A more realistic representation of the ever-evolving nature of Dhaka's streets is achieved by including a variety of traffic conditions and vehicle classifications in the dataset. The collection includes several locations: Ashulia to Badda, Technical Road, Mirpur Sony Square, and Burulia Bridge to Mirpur 1. Table 3.2.1 provides the categories of vehicles. Creating a dataset that precisely identifies and classifies each significant item inside Dhaka's traffic environment was the primary purpose of this complete depiction.

Table 3.2.1: Overview of the dataset.

Category	Description
Car	Passenger cars, including sedans, hatchbacks, SUVs, etc.
Truck	Commercial trucks of any size, including pickups, vans, and heavy-duty vehicles.
Bus	Public transport buses, including single-decker and double-decker buses.

Rickshaw	Traditional cycle-based rickshaw is a standard mode of transportation.
CNG	Three-wheeler vehicles fueled by compressed natural gas are another popular transportation called CNG.
Two-Wheeler	Motorcycles and bicycles are grouped due to their similar size and traffic behaviour.
Person	Pedestrian or individual on a bike or rickshaw.
Total	710 images.

3.2.2 Annotation: I have done annotation using CVAT. Labeling each vehicle instance using CVAT is shown in Figure 3.2.1.

Figure 3.2.1: CVAT annotation tool.

Four parameters determine a bounding box in the YOLO framework: [x_center, y_center, width, height]. To normalize the coordinates, extraction of the x and y coordinates' pixel values indicates the bounding box's midpoint along the x and y axes. Next, by doing the operation of dividing the numerical value of x by the numerical value representing the width of the image, and in the same way, dividing the numerical value of y by the numerical value representing the height of the image. The terms "width" and "height" indicate the measurements of the bounding box in terms of its width and height, respectively.

3.3 Statistical Analysis

I have done the augmentation on 710 images. The augmentation techniques are given below.

3.3.1 Augmentation:

Several different augmentation strategies have been employed to improve the model's generalizability and resistance to differences in lighting, background, and picture quality further.

3.3.1.1 Brightness Augmentation: To imitate the broad range of lighting scenarios observed in real-world traffic, we intentionally adjusted the brightness of our training images. We have increased the brightness by 30% and decreased the brightness by 50%. This boosted vehicle detection accuracy in varied conditions (Figure 3.2.1).

a) Original. b) Brighter. c) Darker.

Figure 3.3.1: Brightness augmentation.

3.3.1.2 Flipping: Images have been turned horizontally to enhance the variety of training data and improve the model's resistance to mirrored objects. Flipping also significantly spreads the training data, strengthening the model's generalizability. Figure 3.3.2 illustrates the before and after flipping.

a) Original image. b) Flipped image.

Figure 3.3.2: Augmentation by horizontal flipping.

3.3.1.3 Rain Effect: Adding rain effect to images may effectively replicate genuine weather conditions. We have added rain using Albumentation, where the rain is a kind of drizzle. (Figure 3.3.3)

a) Before rain effect. b) After rain effect.

Figure 3.3.3: Augmentation by adding rain effect.

Applying these multiple data augmentation procedures, the dataset is expanded from 710 to 2045. These strategies increased the model's capacity to generalize to unknown data, resulting in a more reliable and accurate vehicle recognition system performing consistently during dry weather in Dhaka.

3.3.2 Training, Testing data, and Validation: From 2045 images, 185 images have been chosen for testing purposes. And 1860 images have been taken for training and validation

(80% training, 20% validation). The number of objects in the training and validation dataset is shown below in figure 3.3.4:

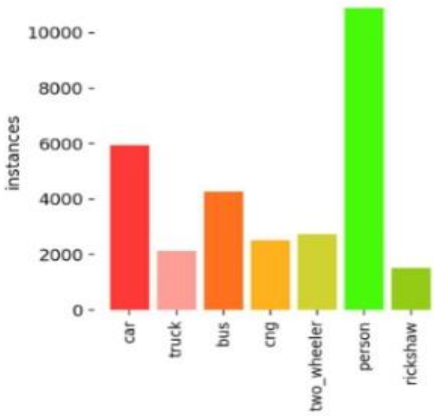

Figure 3.3.4: Number of labels per instance on train and validation dataset.

3.4 Proposed Methodology

For this study, a set of steps have been followed, and the followed method is given below in figure 3.4.1:

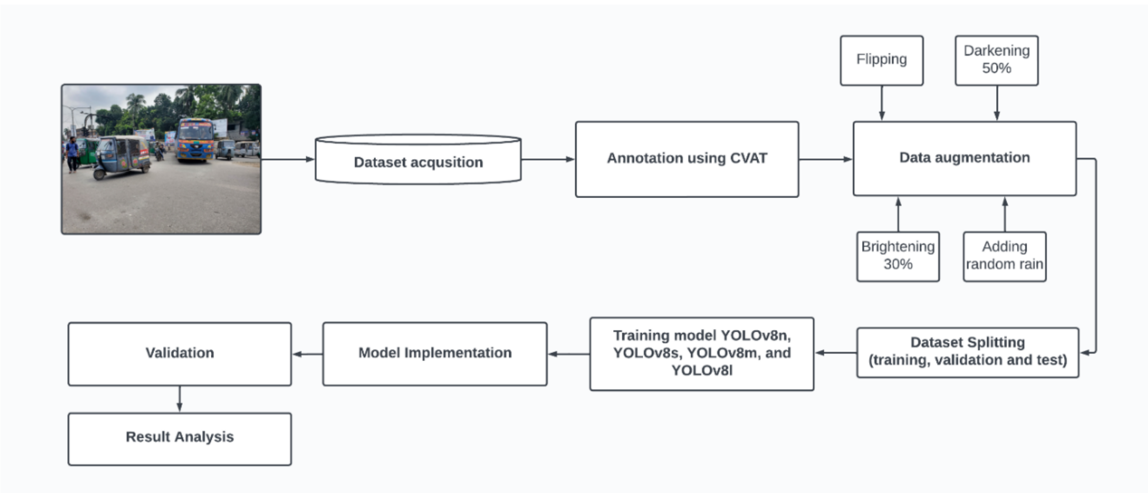

Figure 3.4.1: Proposed methodology.

3.5 Implementation Requirements

3.5.1 Model Selection:

This study has been carried out using four (YOLOv8 variations) models, YOLOv8n, YOLO8s, YOLOv8m, and YOLOv8l, because of their fantastic performance in object detection tasks, especially their high accuracy real-time capabilities, and open-source availability.

YOLOv8 employs a convolutional neural network separated into two primary parts: the backbone and the head. The backbone is designed on a modified version of the CSPDarknet53 architecture, containing 53 convolutional layers with cross-stage partial connections. The head includes multiple convolutional and fully connected layers that predict bounding boxes, objectness scores, and class probabilities. The architecture is depicted below in figure 3.5.1.

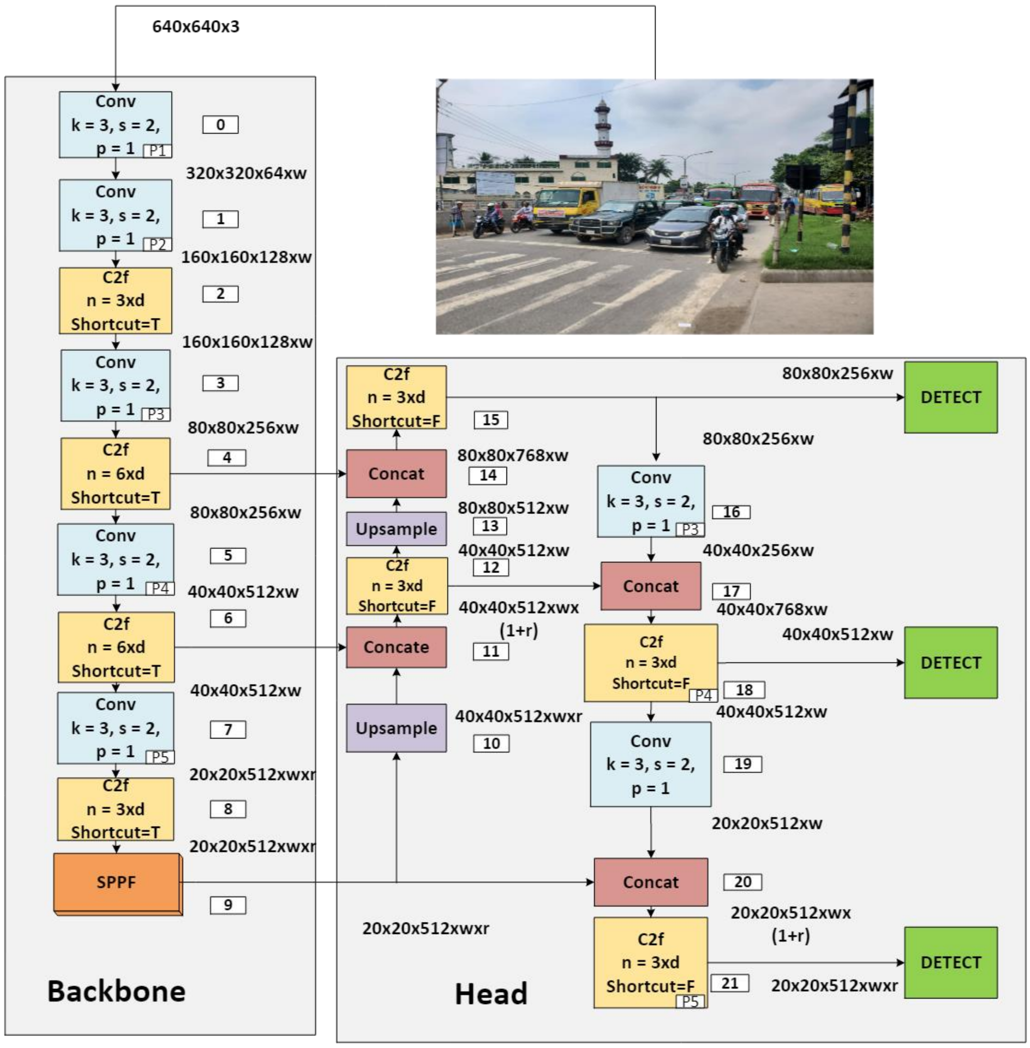

Figure 3.5.1: YOLOv8 architecture.

The architecture uses new convolutions, such as replacing the stem's first 6x6 convolution with a 3x3 one. The basic building block was also modified, with C2f replacing C3. In C2f, each output from the Bottleneck (two 3x3 convolutions with residual connections) is concatenated, unlike in C3, where only the output of the last Bottleneck was used. SPPF layer in YOLOv7 plays a significant role in boosting the accuracy and robustness of object detection, especially for objects of variable sizes in images. Moreover, Multi-scaled object detection featured the pyramid network, Self-Attention Mechanism, Anchor-Free Detection, and Mosaic Augmentation, making the YOLOv8 faster with better accuracy than the previous version of YOLO models. The difference between YOLOv8n, YOLOv8s, YOLOv8m, and YOLOv8l is in the depth and width of layers shown in Table 3.5.1.

Table 3.5.1: Difference between YOLOv8n, YOLOv8s, YOLOv8m, and YOLOv8l

Model	Depth Multiple (d)	Width Multiple (w)
n	0.33	0.25
s	0.33	0.5
m	0.67	0.75
l	1	1

YOLOv8n: This lightest model combines a small size and high inference speed. It was developed to deploy resource-constrained devices or applications demanding real-time performance, such as mobile phone apps or embedded systems.

YOLOv8s: This variation provides an acceptable mix of size and performance, delivering greater accuracy than YOLOv8n while preserving a reasonably swift inference speed. It is a good solution in situations that balance accuracy and computational efficiency.

YOLOv8m: This medium-sized variation emphasizes superior precision over speed. It delivers dramatically enhanced detection accuracy compared to minor variations while preserving appropriate inference performance on contemporary hardware. It is suited for high-detection accuracy applications like traffic monitoring systems or driverless cars.

YOLOv8l: This substantial modification emphasizes getting the best possible accuracy, sacrificing some inference speed. It delivers the highest prospective detection accuracy but needs processing resources for real-time inference.

3.5.2 Model Performance Metrics:

The metrics generally utilized in the area of object detection include precision, recall, intersection over union (IoU), average precision (AP), and Mean Average Precision (mAP) values. The IoU evaluation measures the amount of the intersection between the bounding box predicted by the model and the ground truth bounding box in the original picture. Figure 3.3.3 shows the IoU.

$$IoU = \frac{Area\ of\ Intersection}{Area\ of\ Union} \quad \text{-------------------------------}(1)$$

a) Area of Intersection. b) Area of Union.

Figure 3.3.3: Intersection over Union.

The detection Result was identified as a True Positive (TP) when the IOU value calculated between the Detection Result and the Ground Truth exceeded the noted threshold value.

Precision is a metric that evaluates the ratio of correctly predicted positive objects (True Positives) to the total number of objects that are predicted as positive (True Positives + False Positives).

$$Precision = \frac{TP}{TP+FP} \quad \text{-------------------------------}(2)$$

The recall metric is the ratio of accurately detected objects to all positive objects in the test set.

$$Recall = \frac{TP}{TP+FN} \quad \text{-------------------------------}(3)$$

The precision-recall (PR) curve is the graphical representation that displays precision on the vertical axis and recall on the horizontal axis. This visualization demonstrates the relationship between the classifier's accuracy in correctly finding positive instances and its capability to capture all positive instances.

The average precision (AP) is a scalar measure that predicts the area enclosed by the precision-recall (PR) curve. A higher AP value signifies the classifier's enhanced performance.

$$AP = \int_0^1 Precision(Recall)dRecall \quad \text{-------------------------------(4)}$$

The mAP is a numeric measure utilized to evaluate the accuracy of object detection models across all classes in a specific dataset. The formula for mAP tells us that, for a given class, k, we need to determine its corresponding AP.

$$mAP = \frac{1}{n}\sum_{k=1}^{k} AP_k \quad \text{-------------------------------(5)}$$

The harmonic mean of precision and recall is identified as F1-score.

$$F1 - score = \frac{2*(Precision * Recall)}{Precision + Recall} \quad \text{-------------------------------(6)}$$

3.5.3 Necessary Tools:

The tools and modules I used for this work is given below:

CVAT: For annotation I have used CVAT.

Goolge Colab: I have used the Google Colab IDE.

Ultralytics: Ultralytics provides cutting-edge, state-of-the-art computer vision models for applications including image classification, object detection, image segmentation, and posture estimation.

Necessary libraries and modules: CV2, Matplotlib, Numpy, OS, Shutil, Google drive etc.

CHAPTER 4

EXPERIMENTAL RESULT AND DISCUSSION

4.1 Experimental Setup

In this study, the devices that I have used for data collection is shown in Table 4.1.1.

Table 4.1.1: Specification of Devices for Image Acquisition.

Device Name	Camera Resolution	RAM	Processor
Redmi Note 8 Pro	64MP main + 8MP ultrawide + 2MP macro + 2MP depth	6GB	MediaTek Helio G90T, Octa-core Max 2.05 GHz
Samsung Galaxy S10	16MP main + 12MP ultrawide + 12MP telephoto	6GB	Exynos 9820 (8 nm), Octa-core 2.73 GHz

Moreover, four variations of the YOLOv8 model have been trained on Google Colab with a Tesla T4 GPU. The training procedure contains 50 epochs with a batch size of 8, the size of picture 640x640, momentum: 0.93, weight_decay: 0.0005, workers: 8, pretrained: true, optimizer: AdamW, and IoU: 0.7. All the models have been trained using the same parameters. This argument helps by using the GPU's computing power for quick training.

4.2 Experimental Result & Analysis

YOLOv8n: YOLOv8n is the nano variant and reached the 0.799 mAP50. Table 4.2.1 shows overall result, and the confusion matrix and precision-recall curve is also given below (Figure 4.2.1 & 4.2.2):

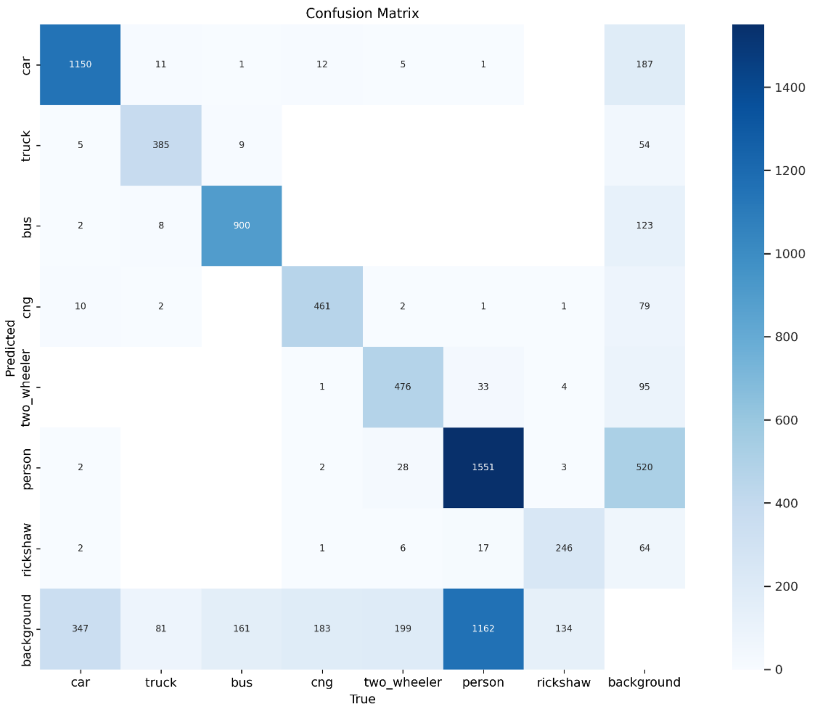

Figure 4.2.1: Confusion matrix of YOLOv8n

There are many wrong predictions of all classes in this confusion matrix. The accurate prediction of the car class is 1150, the truck class is 385, the bus class is 900, the cng class is 461, the two-wheeler class is 476, the person is 1551, and the rickshaw class is 246. It has mostly miss predicted of the person class with background.

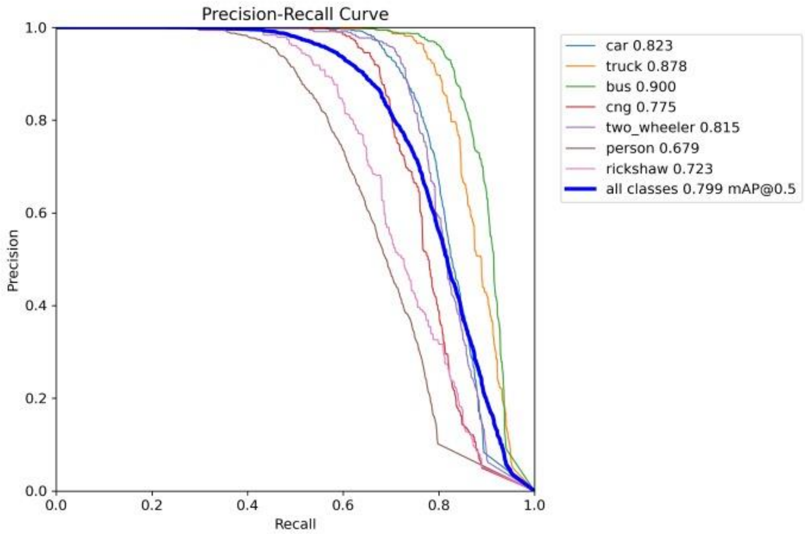

Fig. 4.2.2: Precision-Recall curve of YOLOv8n.

Here, it is shown that in the Fig. 4.2.2 the person class has the lowest AP of 0.679 shown as the left most curve, and the bus class has the highest AP of 0.900 as it has the right most curve.

Table 4.2.1: YOLOv8n result.

Class	Precision	Recall	mAP50	mAP50:95	F1-score	Miss Classification Error Rate(IoU:50)
all	0.902	0.693	0.799	0.603	0.784	0.201
car	0.912	0.717	0.823	0.639	0.803	0.177
truck	0.921	0.791	0.878	0.714	0.854	0.122
bus	0.93	0.822	0.9	0.726	0.875	0.1
cng	0.904	0.676	0.775	0.607	0.777	0.225
two_wheeler	0.948	0.71	0.815	0.613	0.817	0.185
person	0.863	0.528	0.679	0.408	0.657	0.321
rickshaw	0.835	0.603	0.723	0.516	0.705	0.277

YOLOv8s: YOLOv8s has performed better than YOLOv8n (Table 4.2.2).

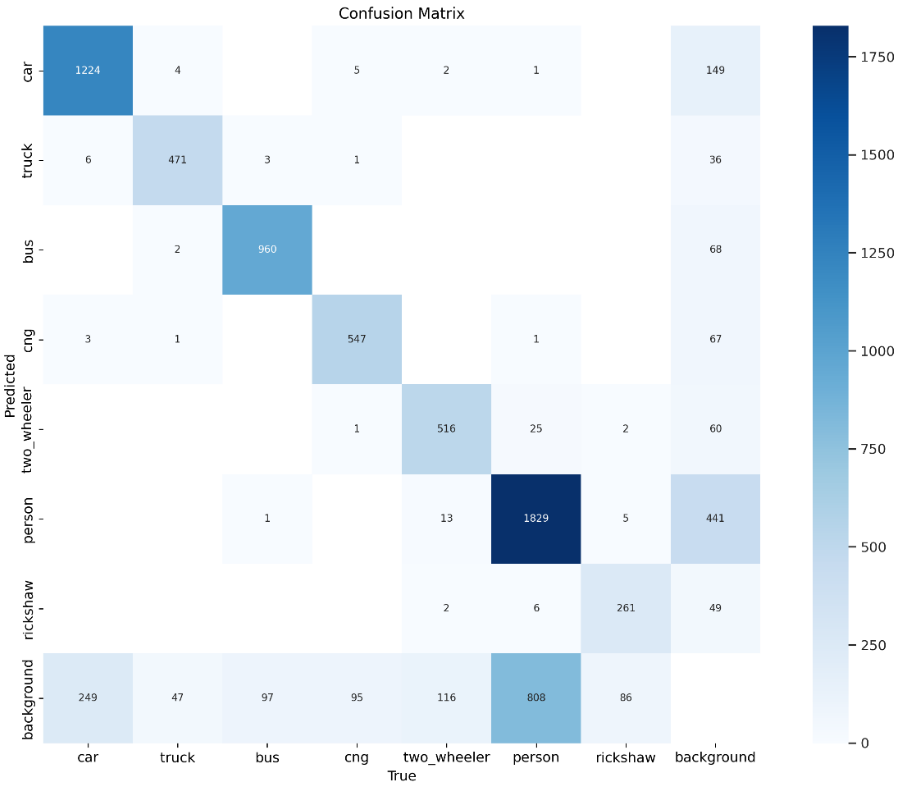

Figure 4.2.3: Confusion matrix of YOLOv8s.

There are still many wrong predictions of all classes in this confusion matrix (figure 4.2.3) as YOLOv8n, the accurate prediction of the car class is 1224, the truck class is 471, the bus class is 960, the cng class is 547, the two-wheeler class is 516, the person class is 1829, and the rickshaw class is 261.

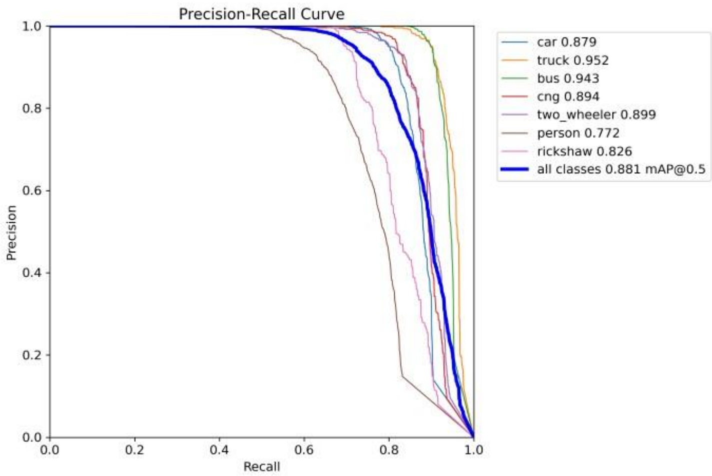

Figure 4.2.4: Precision-Recall curve of YOLOv8s.

Here, it is shown that in Figure 4.2.4, the person class has the lowest AP of 0.772 as it is the highest left curve, the truck class has the highest AP of 0.952 as it is the right most curve, and all classes' mean Average Precision is 0.881 at IoU 0.5.

Table 4.2.2: YOLOv8s result.

Class	Precision	Recall	mAP50	mAP50:95	F1-score	Miss Classification Error Rate(IoU:50)
all	0.944	0.797	0.881	0.711	0.865	0.119
car	0.948	0.801	0.879	0.73	0.875	0.121
truck	0.963	0.895	0.952	0.812	0.927	0.048
bus	0.962	0.892	0.943	0.815	0.925	0.057
cng	0.945	0.82	0.894	0.75	0.882	0.106
two_wheeler	0.938	0.82	0.899	0.721	0.883	0.101
person	0.915	0.634	0.772	0.514	0.754	0.228
rickshaw	0.937	0.715	0.826	0.635	0.821	0.174

YOLOv8m: As YOLOv8m is good at both speed and accuracy, it has performed well. The result is shown below:

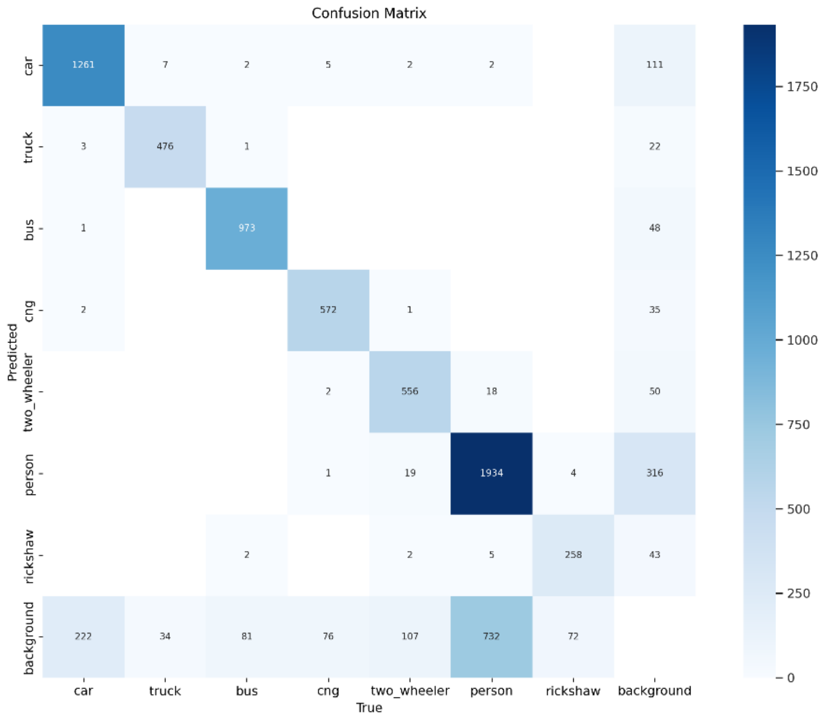

Figure 4.2.5: Confusion matrix of YOLOv8m

This confusion matrix has many wrong predictions of all classes (figure 4.2.5). The accurate prediction of the car class is 1261, the truck class is 476, the bus class is 973, the CNG class is 572, the two-wheeler class is 556, the person class is 1834, and the rickshaw class is 258.

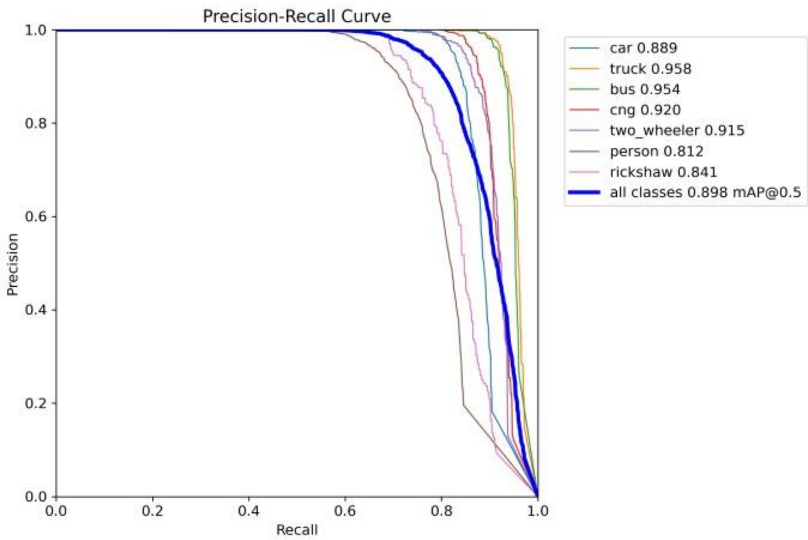

Figure 4.2.6: Precision-Recall curve of YOLOv8m.

This curve (figure 4.2.6) shows that the person class has the lowest AP of 0.812 as it has the highest left curve, the truck class has the highest AP of 0.958 as it has the highest right curve, and all class's mean Average Precision is 0.898 at IoU 0.5.

Table 4.2.3: YOLOv8m result.

Class	Precision	Recall	mAP50	mAP50:95	F1-score	Miss Classification Error Rate (IoU: 50)
all	0.954	0.822	0.898	0.746	0.885	0.102
car	0.95	0.826	0.889	0.753	0.885	0.111
truck	0.975	0.914	0.958	0.849	0.944	0.042
bus	0.974	0.906	0.954	0.85	0.939	0.046
cng	0.969	0.854	0.92	0.787	0.905	0.08
two_wheeler	0.966	0.836	0.915	0.756	0.897	0.085
person	0.93	0.687	0.812	0.58	0.795	0.188
rickshaw	0.915	0.731	0.841	0.647	0.817	0.159

YOLOv8l: It is the large model among previous three models, it has shown the best mAP among all (Table 4.2.4).

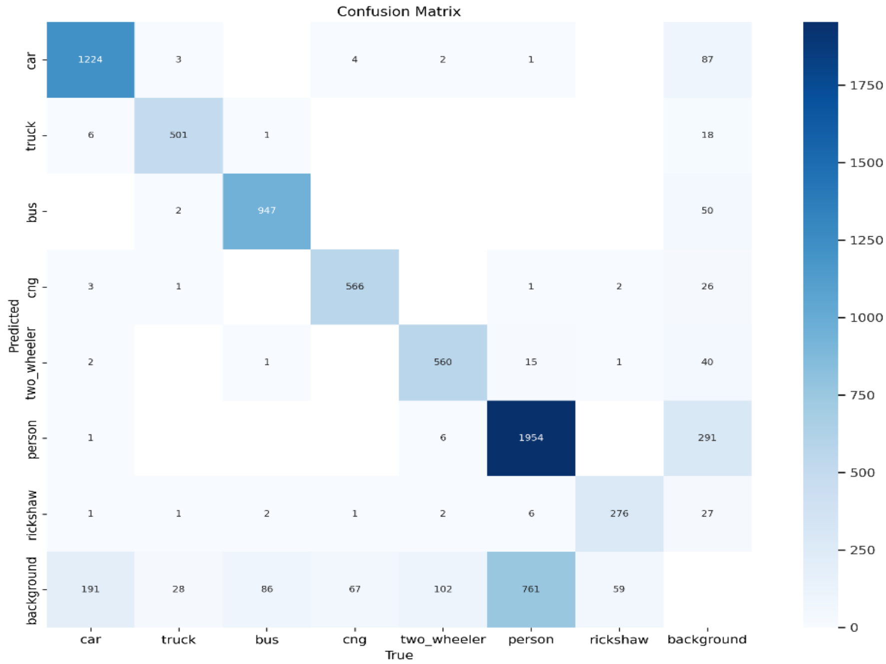

Figure 4.2.7: Confusion matrix of YOLOv8l

It is the large model of all the previous three models but also shows wrong predictions in all classes in this confusion matrix (figure 4.2.5). The accurate prediction of the car class is 1224, the truck class is 501, the bus class is 947, the CNG class is 566, the two-wheeler class is 560, the person class is 1954, and the rickshaw class is 276.

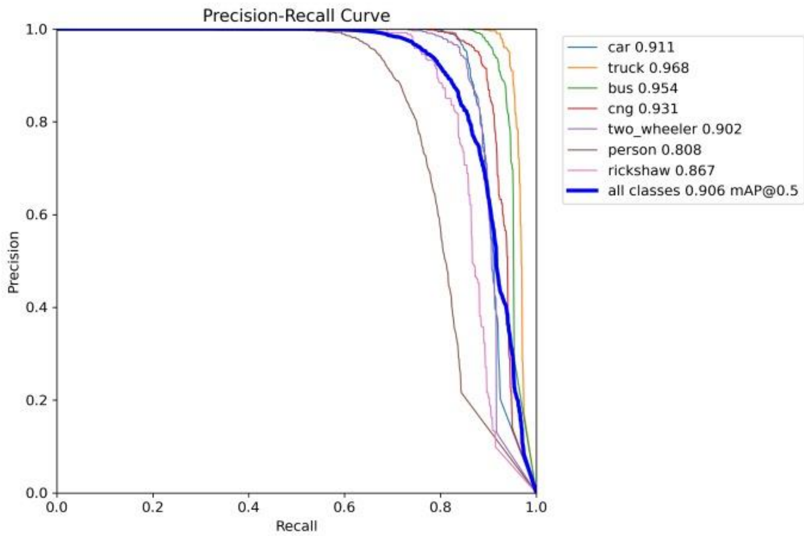

Figure 4.2.8: Precision-Recall curve of YOLOv8l.

In this curve (figure 4.28), it shows that the person class has the lowest AP of 0.808 as it is the most left curve, and for the truck class, it has the highest AP of 0.968 as it is the right most curve, and all classes mean Average Precision is 0.906 at IoU 0.5.

Table 4.2.4: YOLOv8l result.

Class	Precision	Recall	mAP50	mAP50:95	F1-score	Miss Classification Error Rate (IoU: 50)
all	0.958	0.832	0.906	0.766	0.892	0.094
car	0.97	0.843	0.911	0.789	0.895	0.089
truck	0.972	0.929	0.968	0.869	0.95	0.032
bus	0.969	0.9	0.954	0.86	0.935	0.046
cng	0.961	0.868	0.931	0.802	0.904	0.069
two_wheeler	0.969	0.827	0.902	0.76	0.893	0.098
person	0.932	0.681	0.808	0.584	0.794	0.192
rickshaw	0.933	0.777	0.867	0.697	0.852	0.133

4.3 Discussion

This research thoroughly evaluated the findings of the four models. Now, the research analyzes four models depending on how they performed. The test result of the best model (YOLOv8l) of this study is shown in Figure 4.3.1.

Figure 4.3.1: Test result using YOLOv8l.

This study has utilized mAP at 50 and mAP at 50:95 and accuracy as the parameter for comparison. Additional measurements are also shown (Table 4.3.1).

Table 4.3.1: Models comparison.

Models	Precision	Recall	F1-score	Miss Classification Error Rate	mAP50	mAP50:95	mAP50(%)
YOLOv8l	0.958	0.832	0.892	0.094	0.906	0.776	90.6%
YOLOv8m	0.954	0.822	0.885	0.102	0.898	0.746	89.8%
YOLOv8s	0.944	0.797	0.865	0.119	0.881	0.711	88.1%
YOLOv8n	0.902	0.693	0.784	0.201	0.799	0.603	79.9%

According to the metrics in Table 8, YOLOv8l, YOLOv8m, and YOLOv8s have demonstrated acceptable results. In contrast, YOLOv8n demonstrates the lowest mAP, 79.9%, at IoU 50. Regarding accuracy, YOLOv8l has obtained mAP 90.6% at IoU 50. These findings suggest that YOLOv8l beats all other models on my collected dataset.

CHAPTER 5

IMPACT ON SOCIETY, ENVIRONMENT AND SUSTAINABILITY

5.1 Impact on Society

The effect of deploying the project outcome may be enormous, particularly in a heavily populated and traffic-congested metropolis like Dhaka. Here are some possible impacts:

5.1.1 Traffic Control and Safety: By precisely recognizing vehicles and pedestrians, this technology will considerably enhance traffic control in Dhaka. It may aid in decreasing traffic congestion, improving traffic signals, and creating better urban transportation systems. Enhanced safety for pedestrians and drivers may be predicted since the system can help spot hazardous actions or dangerous circumstances in real-time.

5.1.2 Accident Prevention and Analysis: The capacity to recognize cars and individuals in real-time may play a vital role in averting accidents. The technology may inform drivers or traffic control officials about possible risks. Moreover, in the case of an accident, the data gathered may be analyzed to understand its origins and create measures to avoid such accidents in the future.

5.1.3 Urban Planning and Development: Long-term data acquired from such a system may be essential for urban planners. Understanding traffic flow, peak hours, pedestrian mobility, and vehicle types may guide infrastructure development, public transit demands, and urban design choices.

5.1.4. Law Enforcement and Security: The technology may enable law enforcement to monitor and control traffic law compliance, discover stolen cars, or even perform more extensive security operations by tracking suspicious behaviors or persons.

5.2 Impact on Environment

Various beneficial environmental consequences may exist, especially in a heavily populated metropolitan region like Dhaka. Here's how this technology might contribute to environmental benefits:

5.2.1 Reduction in Traffic Congestion: Efficient detection and control of traffic may lead to smoother traffic flow, lowering idle time for cars. Less congestion means cars spend less time on the road generating emissions, which directly contributes to reducing air pollution levels.

5.2.2 Decreased Vehicle Emissions: Improved traffic management equals lower fuel use since cars aren't trapped in traffic for long periods. This decrease in fuel use correlates directly with reducing carbon emissions and other hazardous pollutants.

5.2.3 Promotion of Eco-friendly Transportation Modes: By improving safety and efficiency in traffic management, cities may promote the use of bicycles and walking, further reducing automobile traffic and pollution.

5.3 Ethical Aspects

This study poses various ethical problems that must be addressed to ensure the technology is utilized ethically and fairly. Here are some of the critical ethical aspects:

5.3.1 Privacy Concerns: AI for detecting cars and persons inherently requires collecting enormous amounts of data, including images or recordings of individuals without their knowledge. It's vital to guarantee that this data is managed to respect people's privacy and conform to data protection rules.

5.3.2 Bias and Fairness: AI systems, like YOLOv8, may have built-in biases that depend on the data they were trained on. It's crucial to guarantee that the system is not prejudiced against specific groups of people in general, such as people who belong to a certain race, gender, or socioeconomic position.

5.3.3 Openness and Accountability: There should be openness in how the system functions and how choices are made. For any accidents or misidentifications, it's crucial to establish transparent accountability processes to assess culpability and eliminate any damage caused.

5.3.4 Surveillance and Civil Liberties: The deployment of such technology might be viewed as a type of surveillance, which raises issues about civil liberties. Balancing technology's advantages with people's rights to freedom and autonomy is necessary.

5.3.5 Data Security: With the collection and storage of vast volumes of data, there's a danger of data breaches and illegal access. Ensuring adequate data security procedures is vital to secure sensitive information.

5.4 Sustainability Plan

Such a strategy should cover environmental, economic, and social components to line with sustainable development goals. Here's a summary for a thorough sustainability plan:

5.4.1 Environmental Sustainability: Optimizing the AI and sensing systems for energy efficiency, lowering the carbon footprint related to operating the technology. Moreover, YOLOv8 based optimized software or application could be beneficial for lowering the carbon emission of the system.

5.4.2 Economic Sustainability: Exploring money creation possibilities, such as collaborations with private businesses or data monetization, while maintaining privacy laws and ethical principles. Establishing a financial plan for the future to sustain with YOLO-based system continuing operation, maintenance, and system improvements that can sustain. Investing in local workforce development by establishing new employment associated with operating and maintaining the system and offering training programs make progress.

5.4.3 Social Sustainability: Actively including the community in the planning and implementation process to ensure the technology reflects the demands of varied groups and respects the local context. Conducting education and awareness initiatives to educate the public on the advantages and workings of the system, establishing public trust and acceptance in YOLO-based system. Adhering to the highest ethical standards and privacy protection procedures to guarantee the respectful and fair use of the technology.

CHAPTER 6

SUMMERY, CONCLUSION, RECOMMENDATION, AND IMPLECATION FOR FUTURE RESEARCH

6.1 Summary of the Study

The study includes the analysis of the performance of four distinct YOLOv8 models YOLOv8n, YOLOv8s, YOLOv8m, and YOLOv8l - for detecting vehicles and people in the urban environment of Dhaka. The research focuses on the usefulness of these models in terms of accuracy and speed, especially pertinent in Dhaka's complex and congested metropolitan environment. The article details the technique adopted, which involves deploying these models in several real-world settings in Dhaka to identify vehicles and people. The performance these models has been evaluate based on the mean Average Precision (mAP) measure at an Intersection over Union (IoU) threshold of 0.50. While intriguing, this study's vehicle detection model has drawbacks. The small dataset, missing specific Bangladeshi vehilces, may lead to low real-world accuracy. Due to a lack of solid computing resources and restricted GPU on the Colab free edition, the picture size during training has been set to 640, resulting in more misclassifications of distant levelled objects and worse accuracy. This research exhibits an outstanding mAP of 0.906 at IoU 0.50 (YOLOv8l) on this freshly constructed dataset employing YOLOv8 models.

6.2 Conclusion

This study has gone to the heart of Dhaka's traffic chaos, utilizing the power of YOLO to predict a brighter, more efficient, and safer future. By properly training and assessing YOLOv8n, YOLOv8s, YOLO8m, and YOLOv8l models for vehicle detection, this research has set the path for a dramatic change in the city's traffic management environment. Our investigation has effectively proved the performance of YOLOv8-powered vehicle detection in Dhaka's unique traffic circumstances. The precision and mAP of the trained models reveal the potential to transform traffic control systems. Among the examined models, YOLOv8l emerged as the leader, establishing an ideal balance between accuracy and speed. Its capacity to correctly identify vehicles in real-time, even in demanding circumstances, makes it a perfect option for deployment in resource-constrained contexts. Integrating these YOLO-based technologies into traffic signals and infrastructure may yield several advantages, including decreased congestion, increased safety, and directing paths for Autonomous Vehicles.

6.3 Implication of Further Study

Many major implementation stages might be explored for future studies to increase the efficacy of car and pedestrian detection in Dhaka using YOLOv8. First, I will increase the dataset by gathering real-world data unique to Dhaka's metropolitan setting and employ data augmentation methods for improved model generalization. Experiment with hyperparameter adjustment to optimize model performance and look into different IoU thresholds to understand precision-recall trade-offs. Develop a real-time deployment mechanism for YOLOv8 models, merging object tracking for traffic flow analysis. Collaborate with local governments to incorporate the technology into city infrastructure while resolving privacy concerns. Monitor and evaluate system performance, assuring scalability and energy economy. Assess the environmental impact and undertake a cost-benefit analysis. Lastly, by the support research of partnerships to develop object detection and urban monitoring technology. These stages together seek to increase the accuracy, efficiency, and ethical issues of urban management solutions in Dhaka, helping safer and more effective traffic management and urban planning in the city.

REFERENCES

[1] Telaumbanua, A.P.H., Larosa, T.P., Pratama, P.D., Fauza, R.U.H. and Husein, A.M., 2023. Vehicle Detection and Identification Using Computer Vision Technology with the Utilization of the YOLOv8 Deep Learning Method. Sinkron: jurnal dan penelitian teknik informatika, 8(4), pp.2150-2157.

[2] Maurya, A.S., VEHCLE DETECTION IN AUTONOMOUS VEHICLES USING YOLOV8.

[3] ÖZTEL, G.Y. and ÖZTEL, İ., 2023. Deep Learning-based Road Segmentation & Pedestrian Detection System for Intelligent Vehicles. Sakarya University Journal of Computer and Information Sciences, 6(1), pp.22-31.

[4] Buitrón, I.A. and Yoo, S.G., 2023, October. Performance Analysis of You Only Look Once, RetinaNet, and Single Shot Detector Applied to Vehicle Detection and Counting. In Conference on Information and Communication Technologies of Ecuador (pp. 248-262). Cham: Springer Nature Switzerland.

[5] Salekin, S.U., Ullah, M.H., Khan, A.A.A., Jalal, M.S., Nguyen, H.H. and Farid, D.M., 2023, October. Bangladeshi Native Vehicle Classification Employing YOLOv8. In International Conference on Intelligent Systems and Data Science (pp. 185-199). Singapore: Springer Nature Singapore.

[6] Kumar, D. and Muhammad, N., 2023. Object detection in adverse weather for autonomous driving through data merging and YOLOv8. Sensors, 23(20), p.8471.

[7] Usmani, S.A., Salam, M.A., Foyjul Haque Somrat, M. and Shamim Kaiser, M., 2022. Toward Deep Learning-Based Automated Speed and Line Change Detection System in Perspective of Bangladesh. In Proceedings of Trends in Electronics and Health Informatics: TEHI 2021 (pp. 351-361). Singapore: Springer Nature Singapore.

[8] Wang, M. and Ren, L., 2023. SSB-YOLO: A vehicle object detection algorithm based on improved YOLOv8.

[9] Alamgir, R.M., Shuvro, A.A., Al Mushabbir, M., Raiyan, M.A., Rani, N.J., Rahman, M.M., Kabir, M.H. and Ahmed, S., 2022, December. Performance analysis of yolo-based architectures for vehicle detection from traffic images in bangladesh. In 2022 25th International Conference on Computer and Information Technology (ICCIT) (pp. 982-987). IEEE.

[10] Mostafa, T., Chowdhury, S.J., Rhaman, M.K. and Alam, M.G.R., 2022, October. Occluded object detection for autonomous vehicles employing YOLOv5, YOLOX and Faster R-CNN. In 2022 IEEE 13th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON) (pp. 0405-0410). IEEE.

[11] Rafi, M.M., Chakma, S., Mahmud, A., Rozario, R.X., Munna, R.U., Wohra, M.A.A., Joy, R.H., Mahmud, K.R. and Bijan, P., 2022. Performance Analysis of Deep Learning YOLO Models for South Asian Regional Vehicle Recognition. International Journal of Advanced Computer Science and Applications, 13(9).

[12] Rouf, M.A., Wu, Q., Yu, X., Iwahori, Y., Wu, H. and Wang, A., 2023. Real-time Vehicle Detection, Tracking and Counting System Based on YOLOv7. Embedded Selforganising Systems, 10(7), pp.4-8.

[13] Zhang, Q. and Hu, X., 2023. MSFFA-YOLO Network: Multi-Class Object Detection for Traffic Investigations in Foggy Weather. IEEE Transactions on Instrumentation and Measurement.

[14] Soylu, E. and Soylu, T., 2023. A performance comparison of YOLOv8 models for traffic sign detection in the Robotaxi-full scale autonomous vehicle competition. Multimedia Tools and Applications, pp.1-31.

[15] Wang, G., Chen, Y., An, P., Hong, H., Hu, J. and Huang, T., 2023. UAV-YOLOv8: A Small-Object-Detection Model Based on Improved YOLOv8 for UAV Aerial Photography Scenarios. Sensors, 23(16), p.7190.

[16] Hussain, M., 2023. YOLO-v1 to YOLO-v8, the Rise of YOLO and Its Complementary Nature toward Digital Manufacturing and Industrial Defect Detection. Machines, 11(7), p.677.

[17] Li, Y., Fan, Q., Huang, H., Han, Z. and Gu, Q., 2023. A Modified YOLOv8 Detection Network for UAV Aerial Image Recognition. Drones, 7(5), p.304.

[18] Farid, Annam, Farhan Hussain, Khurram Khan, Mohsin Shahzad, Uzair Khan, and Zahid Mahmood. "A Fast and Accurate Real-Time Vehicle Detection Method Using Deep Learning for Unconstrained Environments." Applied Sciences 13, no. 5 (2023): 3059.

[19] Omari Alaoui, A., El Bahi, O., Oumoulylte, M., El Youssefi, A., Farhaoui, Y. and El Allaoui, A., 2023, May. Optimizing Emergency Vehicle Detection for Safer and Smoother Passages. In Proceedings of the 6th International Conference on Networking, Intelligent Systems & Security (pp. 1-6).

[20] Ma, M. and Pang, H., 2023. SP-YOLOv8s: An Improved YOLOv8s Model for Remote Sensing Image Tiny Object Detection. Applied Sciences, 13(14), p.8161.

[21] Arefin, N.J., 2021. Green City Development: Readiness and Challenges of Bangladesh Government. Dynamics of Public Administration, 38(2), pp.158-172.

[22] Ali, Y., Rafay, M., Khan, R.D.A., Sorn, M.K. and Jiang, H., 2023. Traffic Problems in Dhaka City: Causes, Effects, and Solutions (Case Study to Develop a Business Model). Open Access Library Journal, 10(5), pp.1-15.

[23] Islam, B.Z., Tune, S.N.B.K., Naher, N. and Ahmed, S.M., 2023. Trauma care scenarios following road traffic crashes in Bangladesh: a scoping review. Global Health: Science and Practice, 11(2).

[24] Surendar, V., Raghavendran, P.S., Pillai, A.A., Rakhav, K.A. and Kavin, V., 2023, March. Vehicle Collision Avoidance using Raspberry Pi. In 2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS) (pp. 1653-1657). IEEE.

[25] Huu-Huy, N., 2023. Vehicle-detection-based traffic density estimation at road intersections. International Journal of Open Information Technologies, 11(7), pp.39-46.

[26] Zuraimi, M.A.B. and Zaman, F.H.K., 2021, April. Vehicle detection and tracking using YOLO and DeepSORT. In 2021 IEEE 11th IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE) (pp. 23-29). IEEE.

Plagiarism Report

Plagiarism report Sajib

ORIGINALITY REPORT

18%	15%	6%	9%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	6%
2	Submitted to Daffodil International University Student Paper	4%
3	www.mdpi.com Internet Source	1%
4	Lwando Nkuzo, Malusi Sibiya, Elisha Didam Markus. "A Comprehensive Analysis of Real-Time Car Safety Belt Detection Using the YOLOv7 Algorithm", Algorithms, 2023 Publication	1%
5	Iván Andrés Buitrón, Sang Guun Yoo. "Chapter 17 Performance Analysis of You Only Look Once, RetinaNet, and Single Shot Detector Applied to Vehicle Detection and Counting", Springer Science and Business Media LLC, 2023 Publication	<1%
6	link.springer.com Internet Source	<1%