REAL-TIME VEHICLE FITNESS DETECTION USING YOLOV5

BY

Md. Mohiuddin Khan
ID: 201-15-13631

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

Supervised By

Dewan Mamun Raza
Senior Lecturer
Department of CSE
Daffodil International University

Co-Supervised By

Md. Sadekur Rahman
Assistant professor
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

DECEMBER 2022

APPROVAL

This Thesis titled "**Real-Time Vehicle fitness Detection Using YOLOv5**", submitted by **Md Mohiuddin Khan** to the Department of Computer Science and Engineering, Daffodil International University, has been accepted as 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 March 24,2022.

BOARD OF EXAMINERS

Dr. Touhid Bhuiyan **Chairman**
Professor and Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Subhenur Latif **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Tania Khatun **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Md. Safaet Hossain **External Examiner**
Associate Professor & Head
Department of Computer Science and Engineering
City University

DECLARATION

I hereby declare that, this project has been done by me under the supervision of **Dewan Mamun Raza, Senior Lecturer, Department of CSE** Daffodil International University. I also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:

Dewan Mamun Raza
Senior Lecturer
Department of CSE
Daffodil International University

Co-Supervised by:

Md. Sadekur Rahman
Assistant professor
Department of CSE
Daffodil International University

Submitted by:

Md. Mohiuddin Khan
ID: 201-15-13631
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First, I express My warmest regards and gratefulness to almighty God for His divine blessing making me possible to complete the final year research successfully.

I am truly grateful and wish My deep indebtedness to **Dewan Mamun Raza**, **Senior Lecturer**, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keep interest of my supervisor in the field of "*Deep Learning*" to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, useful guidance, proper mentorship, endless support to my work. Reading many poor-quality drafts and correcting them at all stages made this project possible to complete.

I would like to express my sincere admiration to Touhid Bhuiyan, Head, Department of CSE and Md. Sadekur Rahman assistant professor department of CSE, for their gracious assistance to finish my research and also to other faculty members and the staff of the CSE department of computer science at Daffodil International University.

I would like to thank our entire batch mate in Daffodil International University, who always helps me to complete my course.

Finally, I am dedicating all my work to my father and mother. All the time they provide continuous support from the childhood and all the time inspire me and appreciate my decision.

ABSTRACT

Road accidents, which are currently a big problem in Bangladesh, are largely caused by unsafe vehicles. Thousands of people are dying in Bangladesh due to road crushes. The huge number of unfit vehicles across the country has been increasing over the years, often leading to dangerous road accidents. Many lives are lost and thousands of people are injured in road accidents. According to the statistics, At least 5.42 lakh registered vehicles have been operating on the road without fitness clearance. Every year, 1.35 million people worldwide die in road accidents. In Bangladesh absence of proper monitoring of road transport authority, unfit vehicles can be operating on the road. To identify fit and unfit vehicles some major parts are checked on a vehicle such as the wheel system, braking system, steering system, body, tyres, etc. body parts are very important to identify the fitness status of the vehicle in the initial stage. Hence this paper proposes a vehicle fitness detection model by judging vehicle body parts. In this study only worked with bus which is major category of vehicle. Bus fitness detection using object detection techniques is a very challenging task because of the different colors, shapes, and sizes of vehicles. In this study deep learning-based YOLOv5 method is used for detecting unfit vehicles. YOLOv5 works based on CNN (Convolution Neural Network) architecture. It is a stable model of the YOLO family and has the ability to detect small objects. Data has been collected on the field. Performed data annotation, preprocessing, augmentation, and data labeling for the different classes of vehicles. Divide the dataset into three parts for the purposes of testing, validation, and training. After training the model based on YOLOv5 it has tested on different images and real-time footage and achieved the mAP of 97.3% for all classes. Overall, the approach can be used for vehicle fitness monitoring purposes. It will also help to open new research scope for more improvement in the vehicle fitness detection field.

TABLE OF CONTENTS

CONTENTS **PAGE**

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Introduction

The number of unfit vehicle is increasing every year in Bangladesh. Unfit vehicle is one of the major reasons behind road accidents. In the first eight months of this year, at least 3,186 people were killed and 3,500 people injured in 3776 road crashes [1]. At least 5.42 lakh registered vehicles have been operating on the road without fitness clearance. Thousands of unregistered vehicles there have no fitness results. According to Bangladesh Road Transport Authority (BRTA), only 5.64 lakh vehicles got fitness clearance in the last fiscal year. The number of unfit vehicles is increasing gradually shown by BRTA data [1]. until March 2019 The ten-year-old fitness certificates that are required for more than 70,000 vehicles were missing. And many of them are to blame for the serious accident risk. [2]. If we analyze the past three years' data then we can see In December 2019 the number of unfit vehicles was 2.92 lakh, it has increased to 4.81 lacks in December 2020. As of January 2022, it has increased to become at least 5.42 lakh according to BRTA (Bangladesh Road Transport Authority) every year number of unfit vehicles gradually increases [3]. In 2018 massive student protest was seen and they are demanding safer roads. Following the passing of two college students in a car accident, thousands of students demonstrated in the streets of the capital Dhaka and other cities, however despite their protests, the number of accidents did not diminish. A minimum of 4,439 persons were killed in 3,103 accidents in 2018, and one of the biggest contributing factors was defective vehicles. [2]. In addition, worldwide 1.35 million people perish in automobile accidents every year. Every day, 3,700 people die on the nation's highways. Road traffic accidents are the main cause of death for people under the age of 29. [4]. According to the Road safety status report from WHO for 2018, Bangladesh has one of the highest road crash fatality rates in the world. make an estimate of 24,954 people are killed every year [5]. Apart from this According to the study, Unfit vehicles are also responsible for air pollution in Dhaka city. Black smoke is released from unfit vehicles that is polluting the air. According to the study, not only unfit vehicles but even a good amount of fitness-certified vehicles also produce black smoke. It’s also harmful to the environment [6]. All the unfit vehicles running on the road due to poor monitoring of authority.

In recent years the government has taken so many steps to prevent unfit vehicles and monitor them as well but unfortunately, no effective result comes out. To overcome this problem and monitor unfit vehicles on road continuously without any human intervention a computer vision-based model has been proposed in this paper. In this project, we only work to identify bus fitness detection in real-time. Because in Bangladesh most of the people travel on the bus it is the least expensive form of public transport. Most of the road accidents that happen in Bangladesh are buses. And unfit buses have the highest number of vehicles. To identify fit and unfit vehicles some major parts are checked on the vehicle. For example, wheel system, braking system, steering system, body, communication, tyres, etc [7]. In the wheel system, the condition of tyres, suspension unit, wheel bearing, and alignment of the wheels are checked. In the braking system, test the braking efficiency for both service and parking brakes. In communication check headlamps, reflectors, rearview mirror, direction indicators, etc. In body parts check the body which includes seat belts, windscreen, body crack, etc. body parts are very important to identify the fitness status of the vehicle in the initial stage. Hence this paper proposes a bus fitness detection model by judging vehicle body parts. The model is a trained computer vision-based YOLOv5 object detection algorithm. YOLO models are used for high performance in object detection. As a first step, this model focus to identify bus moving on the road. The image dataset is used for identifying fit and unfit buses. Various kind of fit and unfit bus images has been taken. The model is trained based on 2 classes fit-bus and unfit-bus. As a result, the suggested paradigm is effective. to identify fit and unfit buses in real-time.

The main The goal of this research project is to develop a model that can identify fit and unfit vehicles operating on the road in real-time. Reducing road accidents and death of human species which are mostly happened due to unfit vehicles.

1.2 Motivation

The huge number of unfit vehicles across the country has been increasing over the years, often leading to dangerous road accidents. Experts on road safety and road safety organization blamed unfit vehicles as the main reason behind accidents Many people are dying in road accidents every year. Thousands of people were injured in a road accident. According to the Bangladesh police statistic, more than five thousand were killed and a minimum of four thousand eight hundred people were injured in road accidents in 2021. If we analyze the statistic then we can see every year the death rate of people for road accidents and the number of accidents increase. Lack of oversight of improperly maintained vehicles is the primary cause of traffic accidents. Many unfit vehicles run without fitness certificates on road and it demonstrates that there is no control over them. The absence of proper monitoring of unfit vehicles causes dangerous road accidents. According to the accidents research institute, Bangladesh University of Engineering technology unfit vehicles not usually repair in workshops and the owner has no interest to repair unfit vehicles, as a result, these types of vehicles cannot control their balance during driving and easily cause accidents. If the vehicle owner can drive their vehicle on the road without fitness then they will repair their car and update the documents by spending money. Now it is clear that Lack of proper monitoring leads to unfit vehicles running on the road which is the main cause of road accidents. Many steps taken by the government to prevent road accidents but there was no effective result. In this work, I will illustrate a model that can accurately detect fit and unfit busses good accuracy in the first stage I worked with the bus. Because in Bangladesh busses are the highest number of vehicles among all vehicle. I expect that this model will help to monitor unfit vehicles running on the road which helps to prevent road crushes. This will open new research scope in the computer vision field.

1.3 Objectives

The following are the goals of the research article.

1. Monitoring fit and unfit buses running on the road in real time. Our min aims to detect unfit that cause most road accidents.

2. Prevent road crushes by detecting unfit vehicle

3. Helps traffic police to easily monitor unfit vehicles.

4. Helps with statistical analysis. During the monitoring, it is possible that how many unfit vehicles pass on the road will help to take further decisions. It is possible to implement but, in this project, I only work with fitness detection of busses.

5. Reducing life losses, and people injuries. Which happens in car accidents.

6. Prepare to optimize the bus fitness detection model with the help of deep learning methodology. That will be able to provide accurate detection results with a good confidence score.

7. Introducing a new research domain in the object detection field that will help to further research.

Thesis Outline

The outline of the rest of the report, in chapter 2 some related research papers I reviewed. Chapter 3 discusses how the YOLO model works and the basic idea behind the YOLOv5 model with a proper diagram. Explain about dataset and technique that I used. briefly discuss the proposed approach and Research methodology with proper images and diagrams. In chapter 4 Performance measure result and discussion part, I also illustrate different types of metrics and curve to evaluate the performance of the model.

Finally, I describe the overall research work in the conclusion part then I give acknowledgment who are helps me in my project and the platform that helps me to complete my project. In the end, I have included references.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

recent years there were published various research papers based on deep learning algorithms. In my research, one of the main parts is object detection. Algorithms that are able to detect objects in an accurate way those algorithms maximum probability to provide good results in vehicle fitness detection. In the case of bus fitness detection, there have not enough related papers. For that reason, I reviewed some papers that can help detect an unfit vehicle in an efficient way. Most of the papers about object detection. Discuss about object detection using different kinds of algorithms. Briefly discussed some of the papers in this chapter.

2.2 Related Work

Object detection in the Image processing and computer vision field increasing with great interest. Particularly object detection is used in traffic analysis and human observation. Object detection is one of the most important parts to detect buses accurately.

One of the approaches provided by *Sindhu et al*. is to detect vehicles from the real-time frame in conjunction with a computer vision algorithm. The data was collected from100 videos in mp4. Various vehicle categories can be seen moving on the roads in CCTV footage, four types of vehicles identified Car, Truck, Bus, and Motorcycle. The video was photographed under a variety of weather and lighting circumstances, such as daylight, rain, etc. video data has been converted into 1000 images. Here the main goal is to provide a faster way to identify. accurately identify various car categories on the road. To build this mode three tasks have been followed, Preparing data, mining features, and identifying vehicles. After completing training, model achieved a precision of 53%, an F1-score of 25% Average IoU of 37.78%, and a Mean average precision (mAP@0.50) of 27.29%. It also showed comparable results among some famous detection algorithms. However, it can be predicting vehicles but not all the time. Sometimes it’s missing some vehicles which indicate that the model is not perfect here is need more improvement [8].

Recently, *Rajput et al.* have proposed a model Automatic vehicle identification and classification model for toll management systems. The custom dataset is used for training. YOLOv3 deep learning algorithm has been used. In this study, various types of vehicle classes have been included that are required for the identification and classification. A total of seven classes were selected for the research. 160 images were collected for each class to keep the balance of the dataset. three steps have been followed for data preprocessing Augmentation, Annotation, and division. Complete the annotation process and divide the dataset in an 8:2 ratio. For testing purposes, the Python OpenCV library used and captured videos from toll plazas, highways, and urban areas. After training a custom dataset using the YOLOv3 algorithm. The accuracy result is very impressive although highway and urban areas are unable to detect some vehicles. The model also has some limitations. It takes a higher detection time when multiple vehicles are present in the frame. sometimes it’s unable to detect a vehicle that is far distance and shows a low confidence score. When the vehicle is near distance it shows a high confidence score [9].

Yang et al. show the comparison between YOLOv5-DeepSort and YOLOv7-DeepSort. YOLOv7 is an object detection algorithm that is the recent member of the YOLO family. This version improves the detection accuracy and speeds than the previous version. On the other hand, DeepSort is an object-tracking algorithm. It tracks the object's movement in another way we can say its judge or estimate the position of the moving object. Actually, object tracking comes with an object detection process. This work tries to assess the performance of YOLOv7-DeepSort. In this comparison, they got multiple object tracking accuracies (MOTA) of 40.82% for YOLOv7-DeepSort and 40.77% for YOLOv5-DeepSort. In the same way, multiple object tracking precision (MOTP) and f1 scores got respectively 82.01% and 53.65% for YOLOv7-Deepsort and 81.96%, and 52.43% for YOLOv5-DeepSort. so it is clear that YOLOv7-deepSort is higher than YOLOv5-DeepSort. YOLOv7-DeepSort remarkably enhances the tracking accuracy than the YOLOv5-DeepSort. It proves that YOLOv7-deepSort is applicable for all kinds of tracking tasks for getting higher accuracy [10].

one of suggested models by *Ijjina et al.* Accident detection of car using CCTV surveillance camera and with the help framework for detecting objects using Mask R-CNN. The possibility of an accident is measured on the basis of speed and direction inconsistency. The work is completed in three steps Vehicle Detection, Tracking vehicles, recognizing features, and identifying accidents, at first detects the vehicle using the object detection framework Mask R-CNN. After completing the detection part, the next task was to keep track of each of the detection objects. Centroid Tracking algorithm has been used for object tracking. After accomplishing detection and tracking the last part was Accident Detection. Three new parameters have been used to observe an inconsistency in accident detection Acceleration Anomaly, α, Trajectory Anomaly, β, and Change in Angle Anomaly, γ. However, the model has a few limitations. Its incapability for high-density traffic due to inaccuracies detection and tracking of vehicles [11].

One of the models proposed by *Abdelhalim et al*. based on the YOLOv5 realtime, intelligent surveillance system. increase lab safety awareness. In this work, YOLOv7 and YOLOv5 model is trained using a custom novel dataset. The dataset was consisting of four classes. After training the model Based on evaluation parameters including precision, recall, F1 score, and mAP, the performance of the YOLOv5 and YOLOv7 versions was compared. Comparison result shows all the models' promising performances. The The model with the maximum mAP was YOLOv5 [12].

Another experiment shows proposed by *Liu et al*. In this experiment average accuracy values of YOLOv5, YOLOv4, and YOLOv3 for comparative study using the same dataset. The mAP of YOLOv4 is higher than the YOLOv3 and the YOLOv5 mAP value is more than YOLOv4 in value. Undoubtedly, the YOLOv5 model detection accuracy is stronger to the alternative model. This experiment also depicted that The YOLOv4 model requires more time than the YOLOv3, on average. in single-image detection. On the other hand, the YOLOv5 model takes less time compared to all other models [13].

Recently, a method utilizing multiple computer vision and deep learning ideas was revealed by *Shrestha et al.* They used a modified reidentification model of the original Deep SORT tracking system for vehicle tracking and the object detection method YOLOv3, a precursor to YOLOv4, for vehicle recognition to train the network using the label vehicle dataset. Due to the employment of a similarly quick object detector, they were able to track objects quickly and accurately. However, if the features contain too many information captured in the background, causing due to the object detector's bounding boxes being too large, the performance of model is hampered. However, this model does not provide significant results in low light [14].

CNNs were incredibly successful at detecting objects in moving cars. In addition to performing a number of related tasks, such as segmentation and bounding box regression, CNNs are effective for learning pictorial features [15].

The typical computer vision approach can detect the car more rapidly, but it struggles when there are slow-moving cars around, complicated scenes, photos with varying brightness or recurring background movement. Although advanced CNN has produced positive outcomes in object detection, it is sensitive to differences in scale [16].

CHAPTER 3

METHODOLOGY

3.1 Proposed Method

This section goes through our proposed framework's methodology. We have used the YOLOv5 object detection algorithm for detecting fit and unfit bus in real-time. Recognize the object from the real-time footage possible to achieve using the YOLO algorithm. YOLO divides the picture into several grids and inside each grid tries to identify the object. The abbreviation of YOLO is You Only Look Once. YOLO has several versions the latest version of YOLO is YOLOv7. I am using YOLO version 5 which is a stable version at the present time. This algorithm can identify objects very accurately. With the help of a neural network classifier, this method processes the entire image and divides them into several grids, and predicts bounding boxes for each image. It produces forward propagation with the help of a neural network approach.

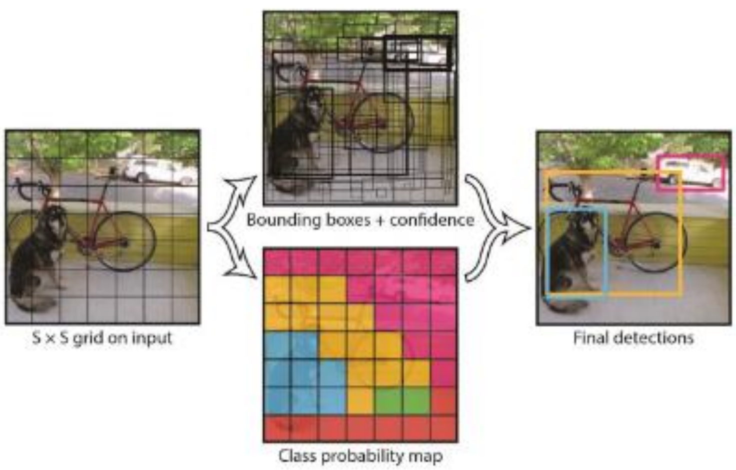


Fig: 3.1: YOLOv4 segments the image into S × S areas and forecasts bounding boxes.

YOLO is a sophisticated, real-time object detecting system. The primary accomplishment of this new approach is that the region proposal step was entirely eliminated which marginally increased the speed and detection effectiveness. The stable version of the method proposed YOLOv5 is used in this paper for vehicle detection.

Figure 4.1 shows the bus fitness detection experimental workflow.

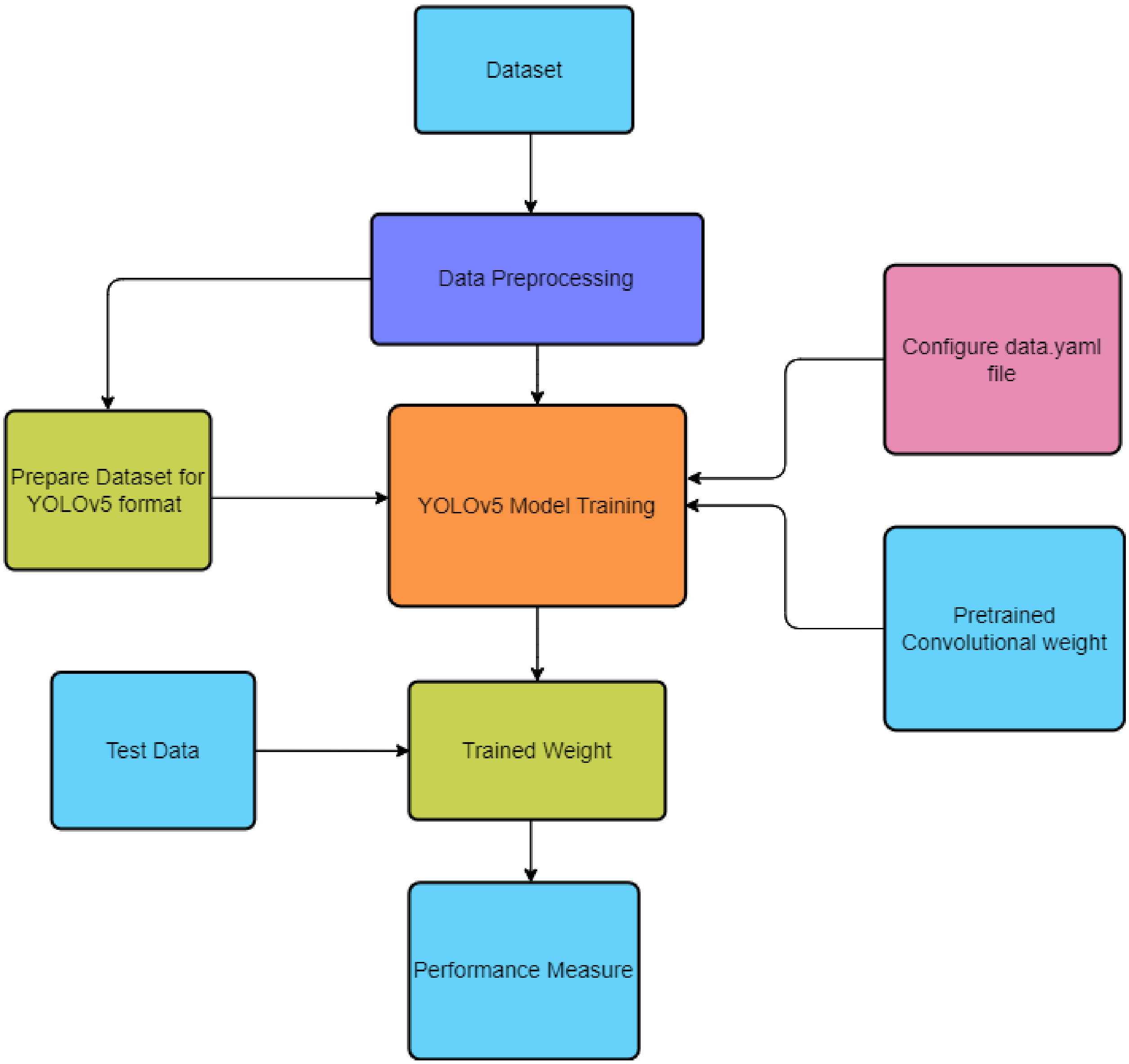


Fig: 3.2. Bus fitness detection experimental workflow

3.2 Basic idea behind YOLOv5

In YOLOv5 detected image is filtered through an input layer before being forwarded to the skeleton to extract features. The backbone has different sizes feature maps after combines the using the feature fusion network, properties are formed into the three maps P3, P4, and P5. (80 × 80,40×40 and 20×20 sizes dimensions used in YOLOv5) to identify large, medium, and small objects in the images. Since sending the prediction head receives three feature maps., bounding-box regression, and confidence calculation using the preset prior anchor execute every single dot on the feature map. Multi-dimensional array contains information about box coordinates, class confidence, object class width, and height. Using corresponding thresholds filters the array's meaningless information then performs a non-maximum suppression process. This technique of converting input images into a multi-dimensional array is called the method of inference. Upcoming threshold and non-maximum suppression process are called post-processing. The network structure is not included in the post-processing. YOLOv5 standard inference procedure can be elaborate in figure 3.3.

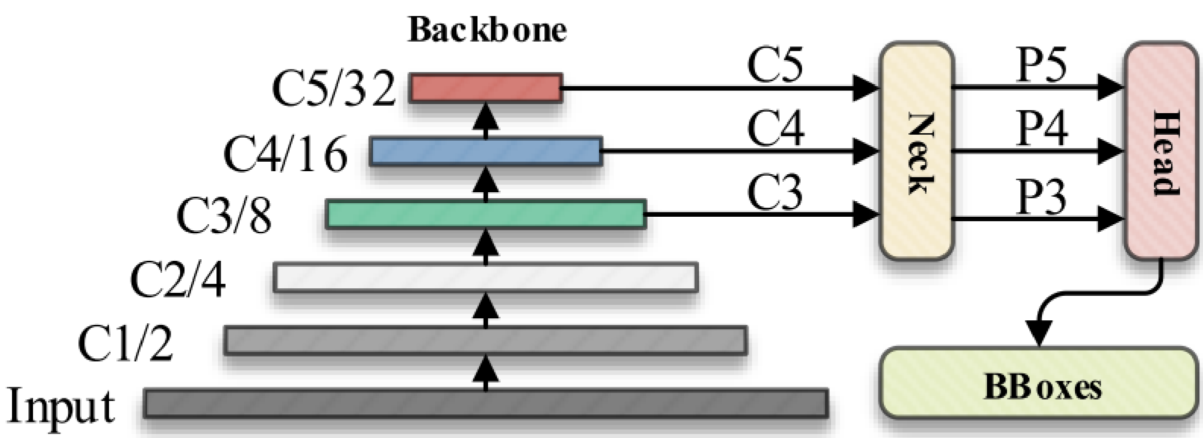


Fig: 3.3. The YOLOv5 standard inference diagram [17]

All the experiments were performed on Google Collaboratory. Which is used 12 GB main memory (RAM) and 78 GB disk space and provide 12 GB NVIDA Tesla k80 GPU. Python 3.7 was used to written all the program.

3.3 Dataset

Image data set used in this project. All the image data was collected from various places in Dhaka city and outside of Dhaka city. Except few images Collected all the data by physically present at the place. Data is collected in various bus terminal one of them popular bus stands in Bangladesh name Gabtoli in Dhaka city and another of them is the Sayedabad bus terminal. This image has various types of buses of different heights and widths. There has fit and unfit various kinds of mini and large buses. Images were shot in various sunlight conditions. Many photos were taken while the bus was on the road and many were taken while the bus was parked. Bus images have been taken from different angles. A total of 260 images were collected for each class. To keep the balance of the dataset equal number of data is considered for each class. For 2 class total of 520 data were considered without augmentation.

A dataset of 1000 images has been created after augmentation. For detecting the bus, I split my dataset 70% for training, 20% for validating and 10% for testing [18].

A dataset sample are showing in Figure 3.4.

Fig. 3.4. Fit and unfit bus in various sunlight condition.

3.4 Technique Used

To capture the images, I have used a 16 MP (megapixel pixel) main camera and a 2 MP deep sensor. The annotation and labeling of several bus images in a frame roboflow Annotation tolls are used manually for each frame. roboflow is a computer vision developer framework used for data preprocessing. It helps image annotation, augmentation, and division of datasets in an accurate and easy way. YOLOv5 An algorithm of object detection is utilized to train and evaluate the model.

This segment discusses the process behind our recommended approach. The YOLOv5 object detection technique was employed for detecting fit and unfit bus in real-time.

My final goal is to provide an easy and fast method to detect various types of fit and unfit buses running on the road with good efficiency. In bus fitness detection to get the expected outcome and use the detection algorithm in a proper way for getting better accuracy, the process consists of some steps.

A: Data Pre-processing
B: Training
C: Feature Extraction.
D: Bus fitness Identification
Proposed approach is possible to achieve by following stages

3.5 Data Pre-processing

The data preprocessing phase comprised of annotation, augmentation, and the dataset is divided into training and validation groups. In the annotation phase, I annotated each image and leveled the fit and unfit bus parking on the terminal or traversing on the road in various sunlight conditions. For labeling, I have used the labeling tool provide by the roboflow. In addition, labeling the bus as a rectangle shape. In frame labeling tool creates a .txt file for each image that contains information about the image. In text file contains a class identifier, x-center coordinate, y-center coordinate, rectangular box width, and rectangular box Table 1 displays the height for the bounding box made over the bus.

Table 1: Contains sample on a text file.

Class	Coordinate(x)	Coordinate(y)	Width(w)	Height(h)
0	0.396875	0.42265625	0.5578125	0.35625

In data pre-processing phase auto orientation is also applied. After complete annotation split the dataset 70:20:10 ratio for training validating and testing. The Augmentation techniques were applied to the image dataset. In the augmentation image flip horizontally, rotate clockwise and anticlockwise between -15 degrees and + 15 degrees. Increase and decrease brightness between -25% and +25%. A sample of the augmentation image is shown in Fig 3.5.

Fig: 3.5 augmented image sample

3.6 Training

In this phase, I first download the PyTorch framework from the GitHub repository which is automatically downloaded by using the command import torch in google colab notebook. After that cloned the YOLOv5 GitHub repository for training our model. I also import the os module and install requirement.txt for some dependencies. For uploading our dataset import roboflow because roboflow has collaboration with google colab and it helps to upload datasets directly roboflow to google collab with the help of the API key provided by roboflow. API key sample and dataset upload command in google colab are shown in fig:3.6

```
!pip install roboflow

from roboflow import Roboflow
rf = Roboflow(api_key="[redacted]")
project = rf.workspace("objectdetection-gsgdp").project("buss-fitness-detection")
dataset = project.version(3).download("yolov5")
```

Fig. 3.6. Dataset upload command and api key sample

After successfully executing the command all data will download inside the yolov5 folder. Now I make configuration on data.yaml file. this is new in YOLOv5. In a previous version of the YOLO family for example YOLOv3, YOLOv4 works in a different way. data.yaml file contains the path of train images, validation images, number of classes, and name of the classes. In my case, the number of classes is 2 and the names of the class are fit bus and unfit bus. This file is very important for the successful training and testing of my model using the YOLOv5 algorithm.

Now our dataset is ready for training. For training my model using the YOLOv5 algorithm I have to make some changes according to our 2 classes the following changes are img=416, batch=16, epochs=150, weights=yolov5s.pt it is pre-trained weights so initially, I am starting training on the pre-trained weights that is by default located inside yolov5 folder, trian.py file is responsible for train model on the algorithm.

3.7 Feature Mining

With the help of PyTorch framework using the YOLOv5 algorithm, I train our model for 150 epochs 32 iterations are required for each epoch. when I finish my model training at the end, I get best.pt file which is my final model. This file I can use for testing my model and can be used to identify the fitness of a bus in Realtime.

3.8 Bus fitness Identification

With the best.pt file I got my train model. Now I take a few images that were not trained and run prediction on the test images with the help of detect.py this file is responsible for detecting the fit and unfit bus. It is able to predict fit and unfit bus framewise. The bounding box is comprised of the output. around the buses and shows the buses' class between two specified classes, i.e., fit bus unfit bus in top left corner of the bounding boxes. It also contains a confidence score that the model is able to predict.

CHAPTER 4

PERFORMANCE ANALYSIS AND DISCUSSION

4.1 Introduction

There have different kinds of techniques to measure the accuracy of the fitness detection model. Similarly, I measure the accuracy and performance using various techniques some of which are Precision, Recall, mean average precision, F1 score, and many more.

Precision: It measures how accurately our model is able to predict our expected result. It calculates the predictions of the correct percentage. Precision measures the number of exclusively positive predictions that fall under this category.

$$Precision = \frac{TP}{TP+FP}$$

Recall: The overall number of accurate forecasts across all accurate predictions is measured by recall. It gauges how well the model is able to identify all positive value.

Recall = $\frac{TP}{TP+FN}$

F1 Score: A harmonic mean of recall and precision is the F1 score. The arithmetic mean is replaced by the harmonic mean. It measures a good combination of recall and precision. In F1 score we calculate average of precision and recall. Maximum value of f1 score indicate high accuracy. On the other hand, less value of F1 score indicate higher disparity of the precision and recall.

$$F1 = 2 \times \frac{Precesion \times Recall}{Precision+Recall}$$

Where,

TP = True Positive

TN = True Negative

FP = False Positive

FN = False Negative

4.2 Performance Analysis

Mean Average precision (mAp): It is used in various object detection model such as YOLO, Mask R-NN, R-CNN etc. With the help of mAp value we can measure how well our model is performing.

mAp calculation formula define as follows:

$$Mean\ Average\ Precision\ (mAP) = 1/M + \sum_{J=1}^{M}(APj)$$

Bounding box loss regressing formula as follows:

$$l_{box} = \lambda_{crd} \sum_{i=0}^{S^2} \sum_{j=0}^{B} I_{i,j}^{ojb} b j(2 - w_i \times h_i) \left[\left(x_i - \hat{x}_i^j\right)^2 + \left(y_i - \hat{y}_i^j\right)^2 + \left(w_i - \hat{w}_i^j\right)^2 + \left(h_i - \hat{h}_i^j\right)^2 \right]$$

4.3 Result and discussion

This section discusses the model performance and result. The model is trained by the YOLOv5 algorithm which is responsible to detect fit and unfit vehicles. In my experiment, I used 1000 datasets 70% data for training 20% for validating, and 10% for testing. Yolov5 Able to detect distant vehicles and provide a good confidence score with an accurate bounding box surrounding the objects.

4.4 Training and Validation Result

A stable version of YOLOv5 was used. A variety of models are available for testing and validation. Various version of the YOLO family is used for getting high accuracy. My experiment shows that the YOLOv5 performed better than the others depicted in table 2. I also used a higher version of YOLO than YOLOv5 but still, I get higher accuracy in YOLOv5. For testing and validation, I employed 50,100,150 epochs for all other models. At first, I used 50 epochs in YOLOv7 show a poor level of precision. Finally, I run 150 epochs in the YOLOv5 object detection model, which shows a good result, and the model performance and accuracy result was superior. Figure 4.1 illustrates the training box losses, figure 4.2 shows validation losses. and figure 4.3 illustrates the precision, recall and mAP metrics. Following all the figures helps to evaluate model performance.

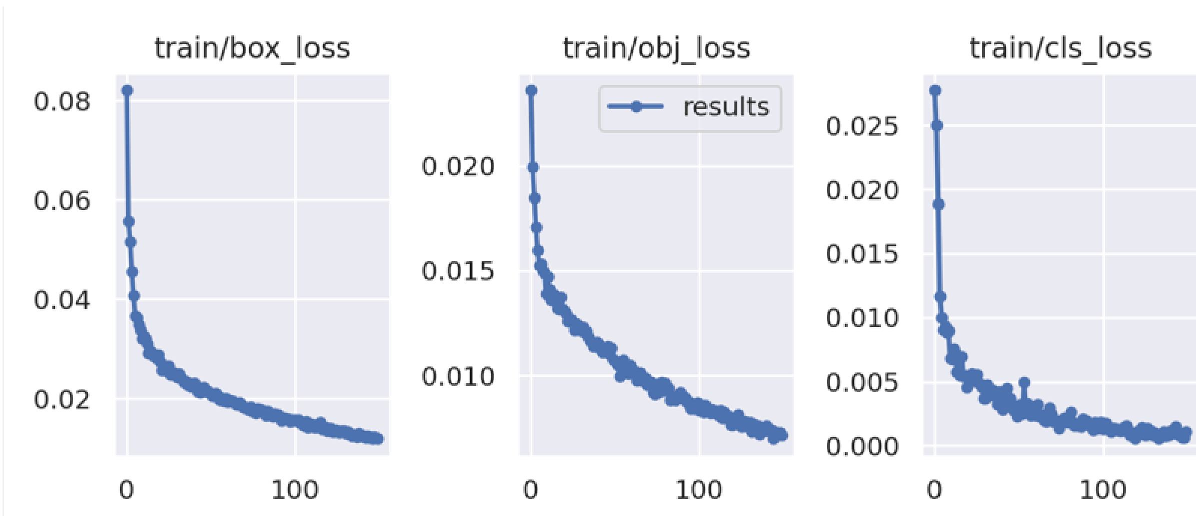


Fig: 4.1 Training losses

Fig: 4.2- Validation losses

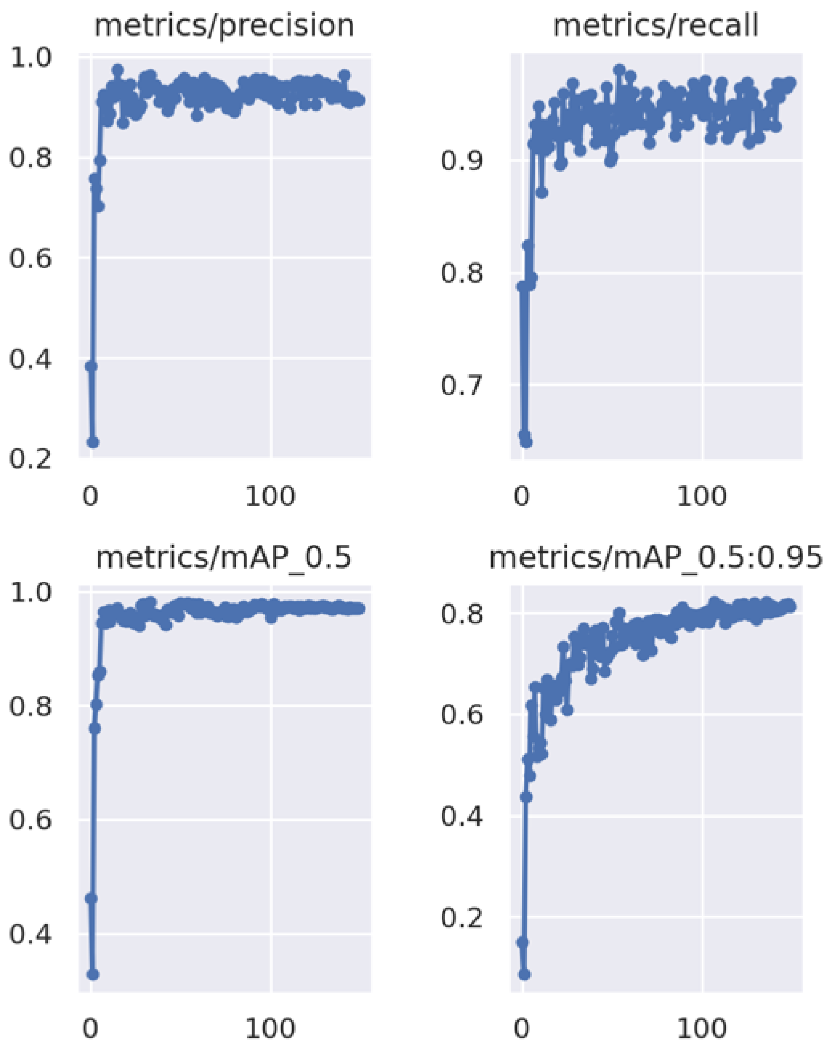


Fig: 4.3- Accuracy, Precision, Recall and mAP metrics.

In figure 4.3 If we look then we can see precision, recall, mAP metrics most of the data point relay above 0.9 this is proved that my model achieved good accuracy. In this metrics identify each data point which is relevant with class.

4.5 Evaluation on fitness detection result

Various measurement techniques are performed to predict the results in the object detection phase. My model provides detection results which are measured with the help of the Intersection over Union (IoU) approach. I used this model on images and real-time footage and some result samples shown below that is able to detect my model.

Fig: 4.4- YOLOv5 bus fitness detection result

In this figure, one is an unfit bus, and the other is a fit bus. My model has been able to detect unfit bus with a 96% confidence score and it's successfully able to predict the bounding box surrounding the bus. In the same way, fit bus is also able to detect with a 94% confidence score. The bounding box is also predicted in an accurate way.

Fig: 4.5- YOLOv5 bus fitness detection result.

In this figure exactly able to identify correctly fit and unfit bus with very good confidence score. In the first image there is a motorcycle in front and it is able to detect it correctly. This proves that my model is capable of detecting vehicles in real-time no matter what is in front or behind.

Fig: 4.6- YOLOv5 bus fitness detection result

In the prediction, we can see fit and unfit vehicles can identify with a good confidence score and it's able to predict bounding boxes surrounding the vehicle very accurately.

After calculating in above results the following values of various metrics (depicted in figure 4.3) used to measure the model performance and detect busses belonging to 2 classes are concluded.

For all classes: Precision is 94% and recall is 92%.

For Fit-bus: precision is 94% and recall is 93%.

For Unfit-bus: precision is 94% and recall is 92%

F1 score for all classes is: 94%

And finally, I got a mean average precision (mAp) of 97% for all classes, for the fit bus 97%, and for the unfit bus, I got a mean average precision (mAP) of 96%. If we notice all the percentage values all most similar for all classes. Because my dataset was balanced, an equal number of images were selected for both classes. Figures 4.5 and 4.6 demonstrate the Precision, recall confidence curve, and 4.7 illustrate the precision and recall curve related to my model. 4.8 demonstrate the F1 confidence curve. This curve helps to understand how well our model is performing in a visual format.

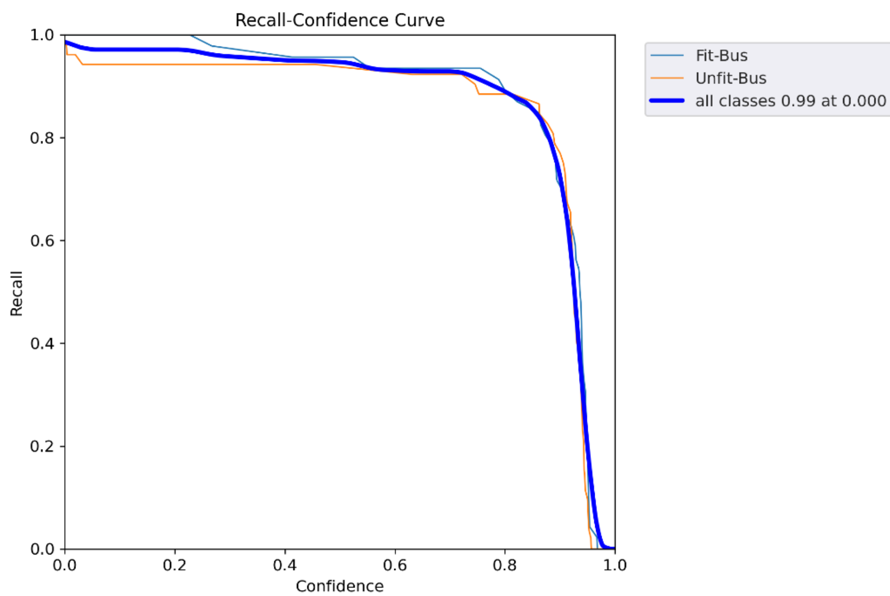


Fig: 4.7- Recall confidence curve

In this fig shows that most of the area under the curve indicates that my model achieves high recall. The curve starts with high recall with low confidence and exponentially increases confidence with good recall. At a point the curve hit the maximum and goes down.

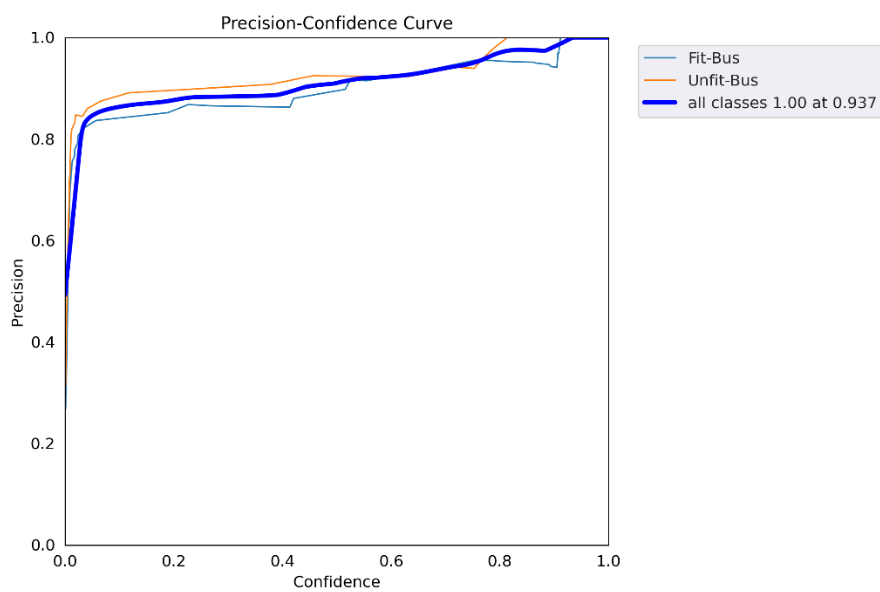


Fig: 4.8- Precision confidence curve

In fig 4.6 Precision confidence curve shows most of the area inside the curve. So it is clear that my model is able to give high precision. The curve starts at 0.5 precision with 0 confidence and exponentially increases precision value from 0.5 to above 0.8 with a low confidence score. After that confidence, the value started to increase exponentially and the precision value increased gradually at the point the curve hit the maximum point.

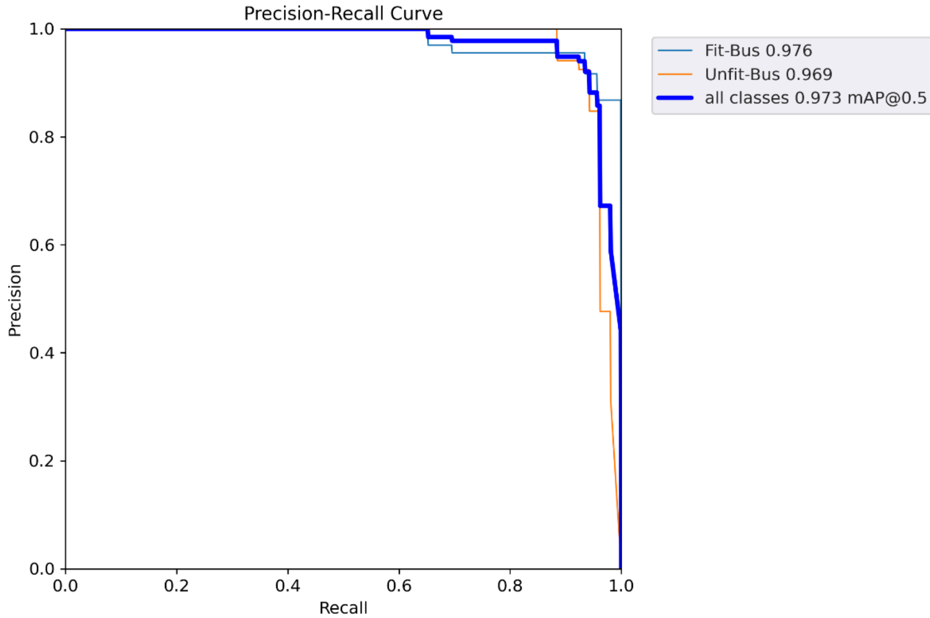
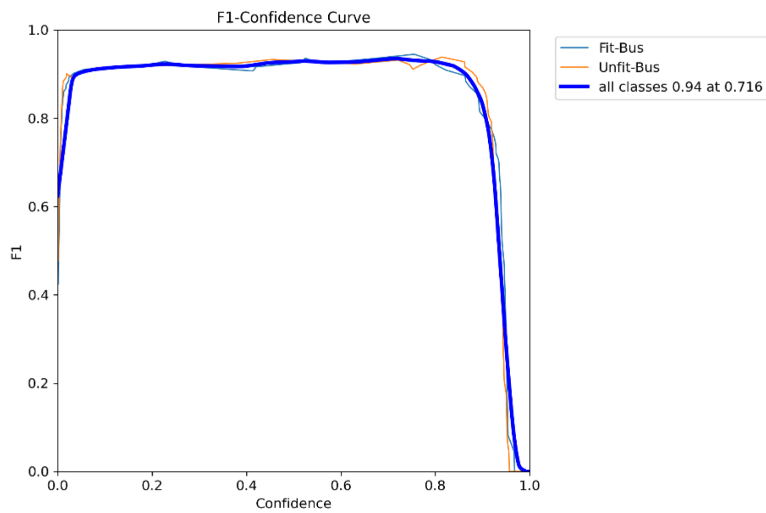


Fig: 4.9- Precision Recall Curve

In fig 4.7 shows precision recall curve. Here we can see maximum area under the curve. The balance of precision and recall for different thresholds is shows by the precision-recall curve. High precision indicate low false positive rate, on the other hand high recall is refer to low false negative. A large area under the curve denotes both high recall and high precision.



Fit: 4.10- F1-confidence curve

A high F1 score demonstrates your ability to recognize serious risks. and are unphased by false alarms. Additionally, it shows that your false positive and false negative rates are low. A model is considered to be completely unsuccessful if its F1 score is 0, whereas a score of 1 is considered excellent. In this curve f1 score shows 0.94 which is very close to 1. It is also indicating my model has minimum false positive and minimum false negative.

4.6 Comparative Study

I tested my dataset another version of the YOLO family one of them is YOLOv7. Trained results of some standard algorithms are shown below for comparison purposes.

Table 2: The standard result of various algorithms for comparison purposes.

Algorithm	Class	Precision	Recall	Mean Average Precision
YOLOv5	All	94	92.9	97.3
	Fit Bus	94.3	93.5	97.6
	Unfit Bus	93.7	92.3	96.9
YOLOv7	All	82.1	79.5	82.7
	Fit Bus	81.8	78.3	81.4
	Unfit Bus	82.4	80.8	83.9

Table 2 depicted YOLOv5 and YOLOv7 model performance based on my custom vehicle dataset. In this comparison, YOLOv5 shows high performance than YOLOv7. YOLOv5 is stable and optimize version of YOLO family that is able to provide good accuracy with no suspicion.

CHAPTER 5

CONCLUSION AND FUTURE WORK

5.1 Conclusion

The study begun by analyzing object detection methods. Using the object detection concept can possibly solve the road crushes problem in which a country has poor monitoring of authority in the transport system. After completing the literature analysis, I prepared my dataset and completed image annotation then used various data augmentation techniques to improve model performance. I trained my dataset to perform validation and testing on my dataset, which included 1000 images. In this research I used 70% of the data for model training, 20% for model validation, and 10% for model testing. I used various models in the YOLO family for experiment purposes. throughout all models, YOLOv5 performs well. YOLOv5 shows great performance with low training and validation loss. In addition, YOLOv5 performed well in the testing phase in comparison to other models. The model is able to predict accurate bounding boxes around the objects and shows a good confidence score. My approach helps in bus fitness detection, The overall performance I got 97%.

5.2 Limitation and Future work

Although my model is performing very well there have some limitations. Sometimes it fails to detect multiple vehicles. This model is able to detect in daylight conditions but at night it's unable to detect because this project only worked with daylight data.

In the future, I have a plan to use a large-size dataset with large number higher-resolution images that will capture in day-night conditions. Here I worked only with buses but in the future possible to increase the number of classes such as trucks, cars, jeeps, motorcycles, and many more.

In the future, I would work with an alert and tracking system by using this model which will provide an alert signal when the unfit vehicle will detect and keep information about the vehicle for tracking and statistic analysis.

5.3 Acknowledgements

I want to thank Google Collaboratory for giving support for GPU. And I am more thankful for roboflow which is helping me in data preprocessing, annotation, and augmentation.

REFERENCES

[1] Staff Correspondent, "Alarm as vehicle fitness clearance at 7-year low," The daily star, p. 2, Oct 20, 2022. [Online]. Available: https://www.thedailystar.net/. [Accessed: Nov 02, 2022].

[2] T. S. Adhikary, " No fitness docs, yet running," The daily star, p. 1, Mar 23, 2019. [Online]. Available: https://www.thedailystar.net/. [Accessed: Nov 02, 2022].

[3] S. Mamun, "half million unfit vehicles threatening road safety," DhakaTribune, p. 1, April 28, 2022. [Online]. Available: https://www.dhakatribune.com/. [Accessed: Oct 15, 2022].

[4] Road Safety facts, available at <<https://www.asirt.org/safe-travel/road-safety-facts/>>, last accessed on 14-10-2022 at 9:00 AM.

[5] A. N. Hussain, "An urgent call to pass Road Transport Rules," DhakaTribune, p. 7, Oct 22, 2022. [Online]. Available: https://www.dhakatribune.com/. [Accessed: Nov 20, 2022].

[6] N. F. Antara, " Unfit vehicles responsible for 15% of air pollution in Dhaka," DhakaTribune, p. 1, Sep 24, 2021. [Online]. Available: https://www.dhakatribune.com/. [Accessed: Oct 02, 2022].

[7] Government of India, “Ministry of Road Transport & Highways,” Annu. Rep., vol. 2020, no. 1, pp. 264–267, 2020.

[8] V. S. Sindhu, “Vehicle identification from traffic video surveillance using YOLOv4,” *Proc. - 5th Int. Conf. Intell. Comput. Control Syst. ICICCS 2021*, no. Iciccs, pp. 1768–1775, 2021, doi: 10.1109/ICICCS51141.2021.9432144.

[9] S. K. Rajput *et al.*, “Automatic Vehicle Identification and Classification Model Using the YOLOv3 Algorithm for a Toll Management System,” *Sustain.*, vol. 14, no. 15, pp. 1–15, 2022, doi: 10.3390/su14159163.

[10] F. Yang, X. Zhang, and B. Liu, “Video object tracking based on YOLOv7 and DeepSORT,” vol. 14, no. 8, pp. 2015–2018, 2022, [Online]. Available: http://arxiv.org/abs/2207.12202

[11] E. P. Ijjina, D. Chand, S. Gupta, and K. Goutham, “Computer Vision-based Accident Detection in Traffic Surveillance,” *2019 10th Int. Conf. Comput. Commun. Netw. Technol. ICCCNT 2019*, pp. 1–6, 2019, doi: 10.1109/ICCCNT45670.2019.8944469.

[12] Z. I. Abdelhalim and H. Aljassmi, “Educational institutions,” *Screen Print.*, vol. 91, no. 8, p. 93, 2001, doi: 10.4324/9780429044854-5.

[13] K. Liu, H. Tang, S. He, Q. Yu, Y. Xiong, and N. Wang, “Performance Validation of Yolo Variants for Object Detection,” *Proc. 2021 Int. Conf. Bioinforma. Intell. Comput. BIC 2021*, pp. 239–243, 2021, doi: 10.1145/3448748.3448786.

[14] Shrestha, S. (2020). Vehicle Tracking Using Video Surveillance. Intelligent System and Computing. doi:10.5772/intechopen.89405.

[15] Z. Q. Zhao, P. Zheng, S. T. Xu, and X. Wu, “Object Detection with Deep Learning: A Review,” *IEEE Trans. Neural Networks Learn. Syst.*, vol. 30, no. 11, pp. 3212–3232, 2019, doi: 10.1109/TNNLS.2018.2876865.

[16] B. Leibe, J. Matas, N. Sebe, and M. Welling, “Preface,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9906 LNCS, pp. VII–IX, 2016, doi: 10.1007/978-3-319-46493-0.

[17] Liu, H., Sun, F., Gu, J., & Deng, L. (2022). SF-YOLOv5: A Lightweight Small Object Detection Algorithm Based on Improved Feature Fusion Mode. Sensors, 22(15), 1–14. https://doi.org/10.3390/s22155817

[18] My vehicle dataset sample, available at <<https://universe.roboflow.com/objectdetection-gsgdp/buss-fitness-detection/dataset/3>>

Real-time Vehicle fitness Detection Using YOLOv5

ORIGINALITY REPORT

6%	4%	3%	2%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Daffodil International University Student Paper	2%
2	Vibhanshu Singh Sindhu. "Vehicle Identification from Traffic Video Surveillance Using YOLOv4", 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), 2021 Publication	1%
3	Submitted to Liverpool John Moores University Student Paper	<1%
4	www.mdpi.com Internet Source	<1%
5	dspace.daffodilvarsity.edu.bd:8080 Internet Source	<1%
6	BSH Shahyeez Ahamed, R Usha, G Sreenivasulu. "A Deep Learning-based Methodology for Predicting Monkey Pox from Skin Sores", 2022 IEEE 2nd Mysore Sub	<1%