UNIVERSITY OF TWENTE.

Faculty of Electrical Engineering,
Mathematics & Computer Science

Forward Collision Warning System with visual distraction detection in bikes

Visshnu Govindaraj (s2276682)
MSc Embedded Systems

December 2021

Supervisors:
Prof. Dr. Paul Havinga
Dr. ir. Yanqiu Huang
Dr. ir. Andre Kokkeler

Faculty of Electrical Engineering,
Mathematics and Computer Science
University of Twente
P.O. Box 217
7500 AE Enschede
The Netherlands

Abstract

The recent development in urbanization has led to a persistent increase in road traffic volume and traffic accidents. Bicycles and Electric Bikes (e-bikes) are highly recommended due to the pollution generated by high-end automobiles, and they are also cost-effective, environmentally friendly, and have health advantages. The increase in traffic volume and accidents has fueled the research on Forward Collision Warning (FCW) systems in automobiles for proper safety of the driver. Forward collision occurs more frequently which leads to fatality on both sides. However, adapting the FCW systems in high-end automobiles to bicycles has limitations due to complexity, cost, and size.

In this master thesis, a multi-sensor integrated forward collision warning(FCW) system with visual distraction detection is developed. The system uses Millimeter Wave (mmWave) radar module for detecting the vehicle in the front and a camera module to estimate the head pose of the rider. The Time-To-Collision (TTC) and Minimum Safe Distance (MSD) metrics are then used to assess the threat and alert the rider. The Threshold of the TTC is varied based on the alertness of the rider. This method shows good performance with an accuracy of 96.7% in shared road scenarios.

Acknowledgement

I am delighted to have completed my masters thesis, under smart bikes project group of the pervasive system's chair. The master thesis journey was exciting and rewarding in aspect of life and academics.

I would like to sincerely express my gratitude to my supervisor dr.ing.Yanqiu Huang,for supporting me throughout my thesis. Her feedback, technical inputs and guidance helped in completing my research and also in acquiring social skills which would help me in my future. I would also like to thank Deepak Yeleshetty, Akhil Pallamreddy, and Khalil Ben Fredj, for helping me in the weekly meeting with helpful thoughts and inputs on my project. I would like to express my sincerely gratitude to Prof.dr.ing Paul Havinga and dr.ing.Andre Kokkeler for accepting to be part of my graduation committee.

Finally, I would like to thank my family and friends for supporting me and motivating me in this master program journey. I would also like to express my gratitude to my friends Kaushik, Jegan and Nithin for supporting and being with me throughout my journey.

I would like to thank the University of Twente for giving me the opportunity to do my Master program. Studying outside my home country has enabled me to gain new experiences and valuable lessons in terms of academics and life.

Contents

Abstract **2**

Acknowledgement **3**

List of Acronyms **6**

1 Introduction **11**
1.1 Motivation 12
1.2 Objective 14
1.3 Contribution 15

2 System Requirements **16**
2.1 Overview 16
2.2 Forward collision warning system with driver monitoring system 16
2.2.1 Criteria to evaluate the proposed system 17
2.2.2 Data Acquisition 17
2.2.3 Evaluation 19
2.2.4 Response 20

3 State of art and critical analysis **21**
3.1 Overview 21
3.2 Driver Monitoring System 21
3.2.1 Sensor Technologies and Evaluation Techniques in Driver monitoring system 22
3.2.2 Existing Solutions on Driver Monitoring systems 23
3.2.3 Comparison of Driver monitoring system 25
3.3 Sensor Technologies and Evaluation Techniques in Forward Collision Warning Systems 26
3.3.1 Existing Solutions for Forward Collision Warning Systems 29
3.3.2 Comparison of Forward Collision Warning Systems 33
3.4 Threat Assessment 35
3.4.1 Comparison of Threat Assessment Techniques 35

4 Theory and Background **37**
4.1 Radar Fundamentals 37
4.1.1 Range Estimation 38
4.1.2 Velocity Estimation 39
4.1.3 Angle Estimation 40
4.2 Head Pose estimation Fundamentals 40

4.2.1 Face Detection Algorithms 40
4.2.2 Histogram of Oriented Gradients Face Detection 42
4.2.3 Landmark Prediction 44
4.2.4 Perspective -N- Point (PNP) method 45

5 Proposed Forward Collision Warning System with visual distraction detection in bikes 46
5.1 System Model 46
5.2 Data Acquisition 47
5.2.1 Radar Data Acquisition 47
5.2.2 Camera Data Acquisition 48
5.3 Data Processing 48
5.3.1 Camera Data Processing for head pose estimation 48
5.3.2 Radar Data Processing 50
5.4 Threat Assessment 52

6 Testing and Results 53
6.1 Camera Sensor Performance 53
6.2 mmWave Radar Sensor Performance 54
6.3 Total System Performance 55
6.3.1 Testing Scenario 56
6.3.2 Real-road Scenario with different lighting conditions 59
6.3.3 Overall Performance 59
6.4 Limitation 63

7 Conclusion and future work 64
7.1 Conclusion 64
7.2 Future Work 65

List of Acronyms

ADAS	Advanced Driver Assistance Systems
ADC	Analog-to-Digital Converter
CAN	Controller Area Network
CASO	Cell Averaging Smallest Of
CCD	Charge Coupled Device
CFAR	Constant False Alarm Rate
C-HOG	Circular-Histogram of Oriented Gradients
CMOS	Complementary Metal-Oxide Semiconductor
CNN	Convolutional Neural Network
DBSCAN	Density Based Spatial Clustering of Applications with Noise
DLT	Direct Linear Transform
DNN	Deep Neural Network
DSP	Digital Signal Processor
DSRC	Dedicated Short-Range Communications
DGPS	Differential Global Positioning System
e-bikes	Electric Bikes
ECG	Electrocardiogram
EMG	Electromyogram
EEG	Electroencephalogram
EURO NCAP	European New Car Assessment Programme
FCW	Forward Collision Warning
FFT	Fast Fourier Transform

FMCW	Frequency Modulated Continuous Wave
FN	False Negative
FP	False Positive
GPS	Global Positioning System
GSR	Galvanic skin response
GUI	Graphical User Interface
HF	High Frequency
HOG	Histogram of Oriented Gradients
HV	Host Vehicle
I2C	Inter-Integrated Circuit
ICT	Information and Communication Technology
IHR	Intermittent Heart Rate
IF	Intermediate Frequency
IMU	Inertial Measurement Unit
IOU	Intersection over union
ISO	International Organization for Standardization
KCC	Kinematic Constraints Criterion
LCD	Liquid Crystal Display
LIDAR	Light Detection and Ranging
LF	Low Frequency
LV	Leading Vehicle
MCU	Micro Controller Unit
MIPI	Mobile Industry Processor Interface
mmWave	Millimeter Wave
MSD	Minimum Safe Distance
MTCNN	Multi-task Cascade Convolutional Networks
NHTSA	National Highway Traffic Safety Administration
NMS	Non-Maximum Suppression
PNP	Perspective -N- Point

PLL	Phase-Locked Loop
RADAR	Radio Detection and Ranging
RSSI	Received Signal Strength Indicator
RCS	Radar Cross Section
RX	Receiver
R-HOG	Rectangular-Histogram of Oriented Gradients
SAE	Society of Automotive Engineers
SNR	Signal to Noise Ratio
SSD	Single Shot Detector
SVM	Support Vector Machine
THW	Time-Headway
TN	True Negative
TP	True Positive
TTC	Time-To-Collision
TX	Transmitter
UART	Universal Asynchronous Receiver Transmitter
VGG	Visual Geometry Group
V2P	Vehicle to Pedestrian
V2I	Vehicle-to-Road Infrastructure
V2V	Vehicle-to-Vehicle

List of Figures

1.1 Levels of Advanced Driver Assistance Systems [1] 12

2.1 Forward Collision Warning System with driver monitoring system 16

3.1 Types of Distractions 21
3.2 Vision-based Monitoring [2] 22
3.3 Physiological sensing [3] 23
3.4 Sensor Technology study-1 26
3.5 Ultrasonic Sensor Working [4] 27
3.6 Radar Sensor Working [5] 28
3.7 Light Detection and Ranging (LIDAR) Sensor Working [6] 29
3.8 Vehicular Communication [7] 30
3.9 Sensor Technology study-2 34
3.10 Deterministic Approach Metrics Comparison 36

4.1 Radar System Overview 37
4.2 Radar Frame [8] 39
4.3 Objects in equal distance [8] 39
4.4 Forward Collision Warning system 40
4.5 Face Detection Flow 42
4.6 Histogram of Oriented Gradients (HOG) feature example [9] 44

5.1 System Overview 46
5.2 Camera data processing 49
5.3 68 Facial Landmarks [10] 49
5.4 Anthropometric 3D rigid model 50
5.5 Radar processing chain 51

6.1 System 53
6.2 Road Scenarios 55
6.3 Scenarios 57
6.4 Testing Area Map 60
6.5 Histogram of Time to Collision Data 61
6.6 Histogram of Range Data 61
6.7 Analysis of Accuracy of the system 62
6.8 Distribution of Accuracy 62

List of Tables

1.1 Traffic Scenarios 13
1.2 Descriptive data on cycling distractions 14

4.1 Derivative masks 43

5.1 Front-end configuration 48

6.1 Comparison of Face Detection Algorithm 53
6.2 Face Detection Algorithm Performance 54
6.3 Maximum Distance for each Time to collision(TTC) 54
6.4 Performance metric description 56
6.5 Testing results 58
6.6 Performance of the system in different lighting conditions 59
6.7 Performance Metrics 60

Chapter 1

Introduction

The rapid growth in urbanization has led to more and more people moving to cities, around 1.3 million people move into cities every day. With the help of Information and Communication Technology (ICT), cities have evolved to smart cities to tackle the problems that occur due to urbanization. The funding for the smart city initiative has been forecasted to increase from 80billion to 189.5billion by 2023 [11]. The transformation of urban cities to smart cities considers cost optimization rather than sustainability goals. Transportation has played a major role in urban sustainability. According to [12], every day 4.5 million trips are made by bus, tram, and metro in the Netherlands, one million are made by train and no fewer than 14.5 million by bicycle.

Urbanization has also led to an increase in desk-bound works such as obesity, diabetes, and hypertension. Hence physical activity is encouraged to overcome these metabolic diseases. Road traffic contributes 17.2%of greenhouse gas emissions in the Netherlands in 2019 [13]. Electric Vehicles play a key role in the reduction of air pollution in urban areas. To date, 17 countries have announced 100% zero-emission vehicle targets or the phase-out of internal combustion engine vehicles through 2050 [13].

Bicycles and e-bikes play an integral part in the transport system globally. It is considered to reduce traffic congestion in big cities, avoid dependencies on the metros and other transport systems during peak hours. Cycling is one of the most preferred physical activities as it aids both transportation and health benefits. With the recent trend of renting bikes, bicycles have gained interest as it enhances mobility and is economical. Bicycles are the Eco-friendly transportation that will be the smartest way to commute in smart cities. Sales of e-bikes are high across Europe indicating that this will slowly become the future at global level as it eases riding through automatized pedaling [14].

An increase in road transportation can increase the number of road accidents. Approximately 1.35 million people experience road accidents every year. In 2019, an estimated number of 21,400 road users were seriously injured in Dutch traffic [15]. This has led to the consideration of safety measures inbuilt in vehicles to reduce human error. The implementation of advanced driver warning systems in smart vehicles is done to counteract human error. Currently, bicycle accidents are responsible for over 7% of all severely injured traffic participants in the Netherlands [16]. Severe injury as a result of bicycle accidents has increased by 35% in the last 10 years [17].

To reduce the accidents, Advanced Driver Assistance Systems (ADAS) are introduced which helps in a mitigating collision or in rare cases reduce bodily injuries. The increase in ADAS implementation in vehicles in the recent decades shows the acceptability of the system by the people. According to [18], ADAS vehicles showed a 27% reduction in bodily injury claim frequency and a 19% reduction in property damage frequency. The ADAS systems that are implemented in vehicles are categorized into 6 levels of automation by the Society of Automotive Engineers (SAE). The levels of automation that are implemented is shown in Fig. 1.1 The levels of ADAS are introduced in high-end vehicles to reduce

Figure 1.1: Levels of Advanced Driver Assistance Systems [1]

the number of accidents and provide automatic driving. The recent implementation of ADAS systems in high-end vehicles includes automatic cruise control, collision awareness and mitigation, and driver monitoring system. Combining ADAS system and bicycles will lead to safe and Eco-friendly transportation in the near future. Smart bikes seem to be an easy solution as it aids a physical activity and a safer ride.

1.1 Motivation

The frontal collision accounts for 56% of all the fatal accidents that occurred in 2019 [19]. In this forward collision, distraction contributes to 4% where the driver is distracted from driving by phone, talking, eating, gazing outside events. The driver's attentiveness in the road plays a crucial role in avoiding collision. The recent implementation of ADAS systems in vehicles are based on collision awareness and mitigation but does not account for the driver distraction and whether or not the driver is aware of the object in the front. The risk assessment of the forward-collision warning system doesn't take into account the driver's distraction which leads to more false alarms when the driver is aware of the object but warnings are generated.

The Table 1.1 explains the scenarios considered in a forward-collision warning system in bicycles. The recent implementation of Forward Collision Warning Systems for Bicycles takes into account only scenario 1 which works without consideration of cyclist's awareness. The deviation of attention from primary tasks such as driving, pedaling and being aware of the surroundings to texting, prolonged gazing on objects in the road and cognitive thinking would lead to road accidents. The distraction of a cyclist in an urban environment not only constitutes an everlastingly latent factor for road users, but they also have a proven association with the probability of being involved in a traffic crash [20]. Furthermore, given cyclists' high physical vulnerability in traffic crashes, distraction significantly increases the risk of severe injuries or death. Approximately 71% of traffic crashes are preceded by a mistake made by one of the parties involved, and distractions have been shown to cause about 17% of driving errors in experimental tests. The crucial part of distraction is the types of distractions that are being faced by a cyclist on road. In the study by Useche [20], data was collected from 1064 cyclists to determine the type

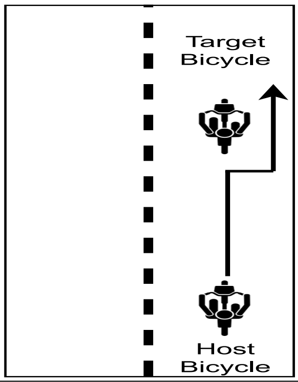
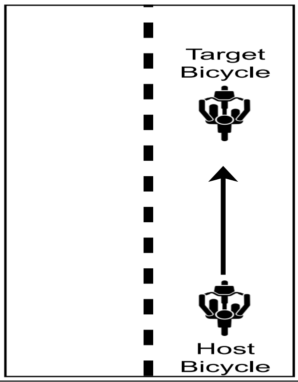
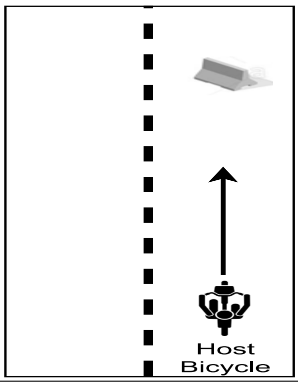
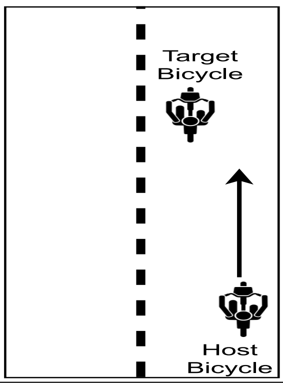
Scenario	Description	Diagram
Scenario 1	Host Bicycle trying to overtake the target bicycle.	
Scenario 2	Host Bicycle traveling along the same direction as the target bicycle. A scenario where the cyclist is distracted.	
Scenario 3	Host bicycle traveling towards a still object and the cyclist is distracted.	
Scenario 4	Host bicycle traveling along the same direction as target bicycle with TTC>5 seconds	

Table 1.1: Traffic Scenarios

of distractions that they faced, It can be seen from Table 1.2, that visual distraction ac-

Types	Percentage Reported
Text messages or chats	46.4
Phone calls	64.9
Billboards	34.7
My own thoughts or concerns	55.1
Weather conditions	68.5
The behavior of other users of the road	83.6
The obstacles in the way	83.5

Table 1.2: Descriptive data on cycling distractions

counts for around 34.7% and 83.6% of all the distractions experienced by a cyclist. The driver monitoring systems implemented for high-end automobiles detects the eye gazing, head movement, drowsiness, usage of phones, and heart rate of the driver which in turn helps in alerting the driver. Such systems are not yet implemented in bicycles to improve safety. Adapting such systems done for high-end automobiles in bikes needs attention on the position of the sensor, accuracy that can be achieved, and types of sensors needed.

Hence implementing a visual distraction detection system with forward collision warning in bikes is needed for the safety of the rider and reducing traffic accidents.

1.2 Objective

The implementation of ADAS in bicycles needs to be energy-efficient, cost-effective, and lightweight which doesn't affect the functioning of the cyclist. The system should be less complicated in both hardware and software without sacrificing accuracy. Although the forward-collision warning system with head pose estimation is present in high-end automobiles, they are not feasible for bikes. Adapting the Forward collision Warning Systems implemented in High-end Vehicles to bicycles has limitations. The primary goal of this research is the implementation of a Forward Collision Warning System that detects the forward collision threads and visual distractions seen in cyclists through the use of sensor technology which helps in increasing the safety of the rider. Thus, the question that will be addressed in this research is **"How to provide a forward-collision warning system with visual distraction detection on bikes?"**.
The key sub-questions that will be addressed in this research are:

- What are the suitable sensors for visual distraction detection and forward-collision warning system?
- How to combine forward object detection and head pose estimation to make a reliable warning system suitable for bikes?
- What is the performance of the designed system in real scenarios?

1.3 Contribution

In this section, the contributions done by this research are explained. The questions listed in Section 1.2 are addressed in chapters 2 to 5. The conclusion and future work of the research are addressed in Section 7.1 and Section 7.2. The contributions are summarized as follows:

- A complete multi-sensor integrated forward-collision warning system with visual distraction detection in bikes using mmWave radar and camera sensor is designed. The data processing and threat assessment are done in the raspberry pi single board computer.
- The radar sensor was carefully configured to achieve maximum range and velocity with low power and low computational complexity. A suitable face detection technique with reduced computational load was used for head pose estimation.
- A combined threat assessment was developed using TTC and MSD for assessing the threat in front of the bicycle.
- A warning algorithm was designed based on the head pose and forward object detection. This functionality is expected to improve safety and increase acceptability by warning the rider about potential threats.
- The proposed system was tested in different scenarios to evaluate the performance. The proposed system shows an accuracy of 96.7% in a real road scenario. The early warning for different thresholds ranges between 0.5 seconds and 4.5 seconds. This exhibits the performance and effectiveness of the system.

Chapter 2

System Requirements

2.1 Overview

In this chapter, the forward-collision warning system with a driver monitoring system is introduced and the requirements of each component of the system are explained.

2.2 Forward collision warning system with driver monitoring system

The forward-collision warning system with driver monitoring consists of three components. The data acquisition block consists of two components,namely driver monitoring and object detection. There are various methods in detecting the object in front which helps in analyzing the situation. The driver monitoring system helps in constantly monitoring the driver and taking action in uncertain situations. The second block is responsible for processing the sensor data and generating threats based on the data. The third part is the response part where a warning is given to the driver through audio, visual, or haptic response. The framework of the whole system is shown in Fig. 2.1

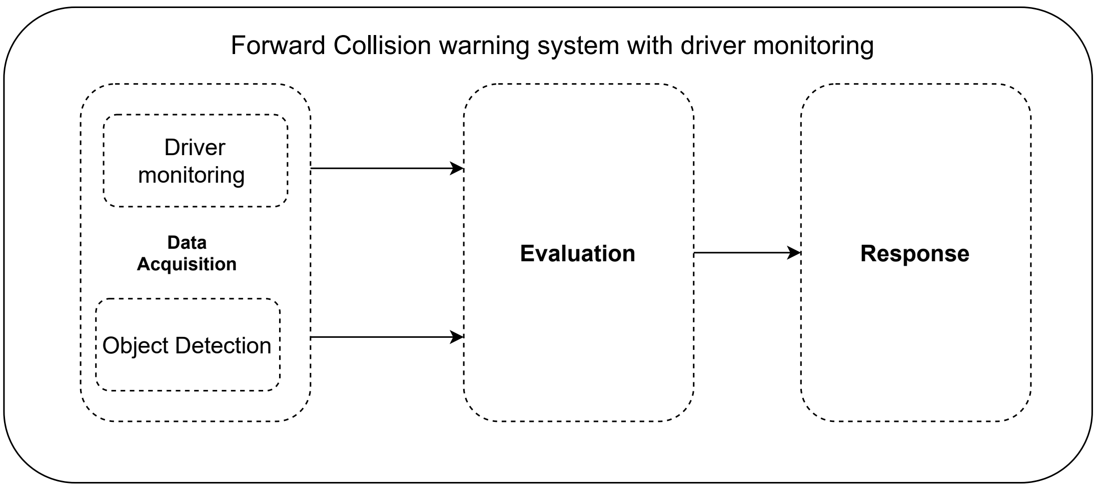


Figure 2.1: Forward Collision Warning System with driver monitoring system

2.2.1 Criteria to evaluate the proposed system

The performance of Forward collision warning system are usually measured using metrics derived from the confusion matrix. Same will be used for evaluating the performance of the proposed system which includes:

- **Accuracy:** Accuracy is defined as the fraction of correct predictions by the system. This values determines the overall correctness of the system. It is anticipated that the accuracy of the proposed system should be high.

- **Precision:** Precision is defined as the quality of the classification done by the system. This metrics quantifies the number of true positive predictions made. It is also known as detection rate of the system. The precision is expected to be high.

- **Sensitivity :** The Sensitivity is to indicate how good the system is at recognizing the positive occurrences. It is also known as True positive rate. The sensitivity of the system should be moderate to high so that the system is not too susceptible to errors.

2.2.2 Data Acquisition

The data acquisition block for the designed forward-collision warning system with driver monitoring consists of two components as Driver monitoring and forward object detection. This block plays a crucial role in acquiring the environmental information related to the bicycle and the rider. This block determines the overall performance of the system. The information extracted using this block is used in assessing the potential threat.

Requirements for Driver monitoring system

The criteria such as accuracy, sensitivity to external conditions, cost, and complexity needed for this block as explained as follow:

- **Accuracy:** Accuracy is defined as the sensor's ability to detect the distraction experienced by the rider. Since there are many distractions that a rider can experience, the accuracy is determined for each distraction and compared. A high-accuracy sensor helps in better detection of distraction and assessment. Needless higher the accuracy can lead to more complexity in implementation and computing. For this system, moderate accuracy sensor technology is needed to detect distractions in riders.

- **Cost:** Cost is defined as the price of the sensor that is needed for the system. The cost of the sensor must be addressed because this technology is aimed at bicycles. This method requires a low-cost sensor.

- **Complexity :** The sensor's implementation complexity is critical for this system. Component constraints, energy consumption, and processing difficulty all rise as the system's complexity grows. This project necessitates a sensor that is easier to deploy.

- **External Sensitivity:** External Sensitivity to environmental conditions plays a crucial role in sensor performance. External sensitivity is defined as the sensor's ability to detect distraction in the rider without being affected by environmental conditions. The proposed system needs a moderate sensitivity that can work in all conditions. This criteria also depends on the type of distraction that needs to be detected.

Requirements for Forward object detection

The death fatality due to forward collision has fueled the implementation of a forward-collision warning system to detect and mitigate from objects in the path. This block contains the sensor technology that is utilized to detect the object in the rider's path. The vulnerability of a rider to detect the object in the path when distracted increases the severity of the accident. The object in the path of the rider is detected by various sensor technology. This detected object is assessed in the algorithm to determine whether the object is a threat or not. The criteria for selecting the suitable sensor technology are discussed below:

1. **Range:** The range is the maximum distance up to which the sensor can detect an object. This maximum range is calculated based on the threat assessment technique and maximum speed of the host vehicle. Range lesser than this can result in low reaction time for an object in front of the rider. The sensor should have long range to generate warnings ahead of collision.

2. **Field of View:** The Field of View of a sensor determines the coverage area. Wider the coverage area, too many objects in the detected frame and the processing of each object increases the complexity. Hence a considerable field of view is needed by this research.

3. **Accuracy:** The accuracy of the system plays a crucial role in detecting the potential threats in the environment. Accuracy is the measure of correctness in detecting the object in the path. Having a high accuracy can help in reducing the false detection done by the sensor. Hence higher the accuracy, the better the detection.

4. **Cost:** Since this system is targeted at a low-end vehicle, the cost of the sensor needs to be considered. A low-cost sensor is needed for this system.

5. **Complexity :** Complexity plays a vital role in implementation. As the complexity of the system increases, the component constrains, energy usage and computation load increases. This project requires a sensor with less complexity in implementation.

6. **Sensitivity to external conditions:** The sensitivity to external conditions plays an important role in the proper working of the system in all conditions. Susceptibility to weather can affect the functioning of the sensor. The sensor for detecting the object in the front needs to be less susceptible to weather conditions.

7. **Estimation Capability:** Estimation capability is the sensor's ability to extract characteristic features such as distance, speed, and position of the target. These features help in the better judgement of the object and judge the severity of the threat.

2.2.3 Evaluation

The Evaluation block consists of Driver Distraction Evaluation and Threat Assessment block. The data gathered from the data acquisition block is processed and the severity of the threat is assessed in this block.

Driver Distraction Evaluation

This block is responsible for the evaluation of the data gathered from the sensors to determine the distraction level of the rider. Some of these most used methods and algorithms are analyzed to find the challenges in implementing for bicycles in the next chapter. In this component, the requirements of the algorithm for driver monitoring systemn are explored. The algorithm's accuracy, complexity, resource usage are explored.

1. **Accuracy:** The accuracy of the algorithm plays a crucial role in driver monitoring system. The accuracy of the algorithm should be high to have more accurate results but the higher the accuracy, the higher the resource usage and complexity. For this system moderate to high-level accuracy is needed to balance the trade-off between accuracy and complexity.

2. **Complexity:** The complexity of the algorithm is important to the implementation. With high complexity, the computing power, resource usage will increase. For this system, a low complexity algorithm is needed to balance the complexity of the forward obstacle detection and driver monitoring system.

3. **Resource usage:** The resource usage by the algorithm is directly linked to the accuracy and complexity. For a low-end device, the resources are constrained, hence the algorithm needed should be able to work in low-resource constrain devices.

Threat Assessment Evaluation

This block is in charge of assessing the threat by evaluating the data collected from the sensors. The requirements of the threat assessment used for assessing the object in the path of the rider are explored. The requirements of the threat assessment to be used are explained.

1. **Accuracy:** The accuracy of the threat assessment, is the proper assessment of a potential threat. The accuracy of the algorithm for assessing the threat should be moderate to high.

2. **Complexity:**The algorithm's complexity has an impact on its implementation. As both the driver monitoring and forward object detection are working simultaneously ,the complexity of the threat assessment should be low to balance the complexity by the driver monitoring system.

3. **Acceptability:** The Acceptability of the threat assessment algorithm is important. Choosing an algorithm that warns with a very low time for reaction can affect the system's functionality. The system needs to have good acceptability since the warnings are based on this algorithm.

2.2.4 Response

This block is responsible for conveying the warnings to the rider. This block needs to be user-friendly, not overly distractive, low cost, and less complex. This block can have two types of implementation based on the type of the assistance system. The two types of systems are as follows:

1. **Feedback:** These systems warn the driver of any potential danger via the optical, acoustic, or tactile sensory channels of human beings. These warning systems are primarily concerned with delivering supporting information to the driver so that he or she may make better decisions for safer driving. The driver is still responsible for making good use of the information to increase overall safety.

2. **Automatic Response:** These systems take over vehicle control from the driver during a potential threat. The system gathers information from the environment, assesses the threat, and makes essential vehicle actions such as applying brakes or changing the vehicle's trajectory.

Due to time constraints and complexity in choosing an optimal method, this block was not explored in the proposed system. The warnings generated by the system were stored in a csv file and used to compare with ground truth.

Chapter 3

State of art and critical analysis

3.1 Overview

The types of sensors and evaluation techniques used in forward collision warning system and driver monitoring system are explained and critically analysed.

3.2 Driver Monitoring System

There has been significant growth in the ADAS systems in automobiles in sensing the environment and taking necessary measures when the driver is distracted. The sensor technology needed for driver monitoring plays a crucial role in the performance of the system. The Distraction that a rider experiences on road are categorized into three types such as Fig. 3.1.

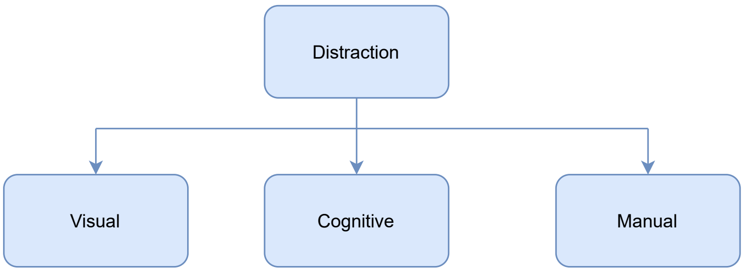


Figure 3.1: Types of Distractions

- **Visual Distraction:** Distraction due to change of focus on the road to secondary tasks.
- **Manual Distraction:**Distraction due to lack of hands on the steering wheel.
- **Cognitive Distraction:** Distraction caused by Prolonged thinking, mind-wandering, Stress, and Cognitive load.

All these types of Distractions lead to less response time to an obstacle on road. Depending on the type of Distraction, the sensing mechanism responsible for it changes and evaluation changes.

3.2.1 Sensor Technologies and Evaluation Techniques in Driver monitoring system

As mentioned in the above section, the sensing technology depends on the type of distraction that is needed to be evaluated. Each technology has its pros and cons associated with it. Thus, it is necessary to analyze different technologies to understand the trade-offs and select the best-suited sensor. In this section, a brief description of the three major categories of sensor technologies in driver monitoring systems is given to gain a better understanding of the following literature.
The three major categories of sensor technologies in Driver Monitoring systems are as follows:

- **Vision-based Monitoring:** This category includes sensors that are used to capture the facial features and evaluate the distraction level through Image processing algorithms for eye tracking, drowsiness detection, and head movement. These sensors are versatile and can provide much useful information about the driver. Such vision-based systems are prone to weather conditions, lighting conditions, and the position of the sensors. The operation of a vision-based monitoring system is shown in Fig. 3.2.

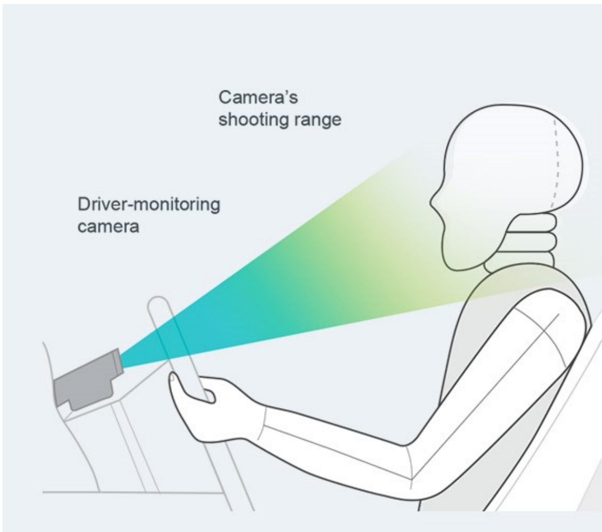


Figure 3.2: Vision-based Monitoring [2]

- **Physiological Sensors:** This category includes sensors that capture the vital signs of the driver to deliver the distraction level. These types of systems are gaining increasing popularity due to the discovery of miniaturized sensors in the biomedical domain. Such systems can accurately determine the distraction of a driver through the use of signal processing algorithms. The physiological sensor-based monitoring system is shown in Fig. 3.3. Some of the common psychological sensing technologies to detect Distraction are:
  - **Electrocardiogram(ECG):** The Electrocardiogram (ECG) signals are used to determine the Heart rate variability of the driver. Heart rate variability is the variation of time intervals between heartbeats. According to the heart

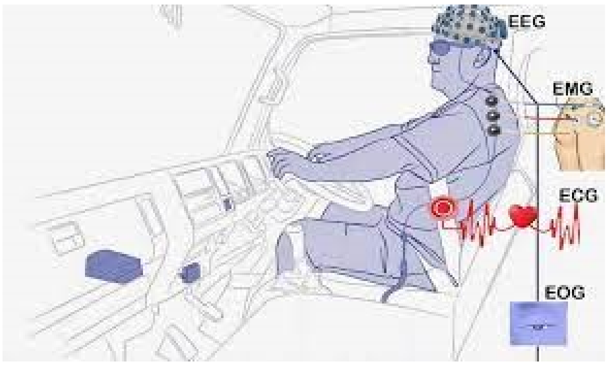


Figure 3.3: Physiological sensing [3]

rate variability metrics calculated through frequency domain such as Low Frequency (LF),High Frequency (HF),LF/HF power density values, the cognitive load, stress, fatigue, and Drowsiness of the driver can be determined.

  – **Electroencephalogram (EEG):** The Electroencephalogram (EEG) signals are electrical signals detected due to the electrical activity of the brain. This signal when processed in the frequency domain, three major rhythms are seen which can be used to denote the cognitive distractions in humans. The three rhythms used are alpha(8-13Hz), beta(13-30Hz), and delta(¡3Hz).

  – **Electromyography (EMG) :** Electromyogram (EMG) is a diagnostic procedure that evaluates the health condition of muscles and the nerve cells that control them. These nerve cells are known as motor neurons. They transmit electrical signals that cause muscles to contract and relax. These signals are used to detect muscle activity near the eyes which helps in determining whether the person is sleeping or not.

- **Manual Distraction :** The manual distraction of a driver depends on the control of the steering wheel. This type of distraction is detected through the use of Inertial Measurement Unit (IMU) sensors, pressure and vibration sensors on the vehicle and steering wheel to detect the sudden change of driver's control on the vehicle.

3.2.2 Existing Solutions on Driver Monitoring systems

In this sub-section, the existing driver monitoring systems that are related to the research are studied and criticized.

Physiological Sensing

According to the article[21], the author presents an Enhanced Random Forest algorithm to detect the attention level in drivers using ECG data from physiological sensors. ECG, EMG, timestamp, foot Galvanic skin response (GSR), hand GSR, Intermittent Heart Rate Intermittent Heart Rate (IHR), respiration, and marker data were collected from the driver who was monitored in different environments(rest, City, Highway) and the stress levels in the drivers were calculated. In article[22], the data from the previous research was used with galvanic skin resistance data to detect stress levels in drivers. The R-R intervals were calculated for the Heart rate Variability to categories the driver

into low, medium, high levels of stress. William J. Horrey et al proposes a system[23] in cars to detect the level of engagement and distraction of the drivers. The cerebral oxygenated hemoglobin, heart rate, and eye movement of the driver are detected during driving simulation. The drivers are subjected to distraction through auditory materials. Three different classifications are done such as interesting, boring, and baseline.
To summarize, physiological-based driver monitoring systems as a wide range of applications. The system has high accuracy in detecting different distractions seen by the driver but there are some shortcomings. The system is highly sensitive to external factors such as weather and illumination. As the physiological sensors are used in vehicles where the rider is still and no physical activities are done, adapting this system to bikes will affect the performance of the system. Thus, it can be concluded that physiological sensing is not suitable for this research.

Vision Based Solutions

The system presented in [24] depends on the movement of eyelids to detect the drowsiness of the driver. A camera is used to detect the real-time video processing, the camera is placed in front of the dashboard through which the video is captured, this data is processed in raspberry pi and if the eyelids are closed, an alarm is given through the speaker. This device is designed specifically for cars. Ashish Kumar et al proposes a system for the detection of drowsiness and distraction in car drivers [25]. A camera sensor is used to detect the face in the frame and a Histogram of Oriented Gradients is used. An Support Vector Machine (SVM) classifier is used to classify faces depending on HOG features. Landmark detection is done on the face detected frames which help to detect the eye aspect ratio, mouth opening ratio, and head bending. With these data, the system is trained for a threshold, and drowsiness and distraction are detected. The system is tested on a laptop for real-time acquisition of camera data from the webcam. The complexity of implementing the system in an embedded device is not discussed as the environment variables play a crucial role when implementing it in cars. The performance of the system can degrade when implemented in an embedded device due to the complexity of the system. Vaibhav Rathod et al proposes a system for detection of distraction on drivers using a camera [26]. The camera is mounted on the front of the driver and the movement of the driver is tracked and a corresponding alert is given when the driver is distracted from driving. Haar cascade classifier is used to detect and track the motion of the driver and an alert is given using an Liquid Crystal Display (LCD). An alcohol sensor is also provided to detect whether the driver has consumed alcohol. The environmental variables such as placement of the camera, camera calibration, and lighting are not considered in this system. The classifier used has less accuracy compared to other classifiers used in other systems .
To summarize, vision-based sensing technology are seen to be straight forward approach for a driver monitoring system. The complexity of the system depends on the image processing algorithm. The system is flexible, low cost, and highly accurate compared to other monitoring systems. Due to the wide range of algorithms present for driver monitoring systems with high accuracy and less complexity. The system can be adapted to low-end automobiles. Hence, a camera-based driver monitoring system is suitable for this research.

Non-Vision Based Solutions

Jaehoon Jung et al proposes a driver monitoring system using mmWave radar [27]. The frequency spectrum of the Intermediate Frequency (IF) signal is monitored for different head movements of the driver. A 2D rectangular window function is applied for the spectrum and slid at an interval of 0.5s to capture the movement of the head. The data collected are stored and used to train a Convolutional Neural Network (CNN) for different movements of the head. The CNN architecture was used to classify the four types of driver's motions. The classification results showed that the trained CNN model can classify four types of driver's motions with over 80% of accuracy. The system can detect the movement of the driver but doesn't estimate the yaw, pitch, roll angle of the head. The use of CNN for classifying the driver's head movement increases computing complexity. Rachel Chae et al proposes a radar-based driver monitoring system for head pose estimation [28]. A coherent Frequency Modulated Continuous Wave (FMCW) radar was used to detect the movement of the driver such as dorsal flexion, dorsal hyperextension, lateral bending, and rotation. A Range-Doppler evolution was created from the range profile and used to detect the movement of the head. The system was not able to distinguish all the movements of the head. The system was not tested in real-time.
To summarize, non-vision-based driver monitoring systems are an upcoming research domain due to their less sensitivity to the environment. The system shows moderate accuracy but lacks the classification of different head movements. The classification of the head pose can be done by high computational algorithms. The system as such increases in cost and complexity when used for the classification of head pose. Thus, it can be concluded that non-vision-based sensing technology is not suitable for this research.

3.2.3 Comparison of Driver monitoring system

Physiological Sensing is gaining increasing popularity in a wide range of applications. These systems have high accuracy in monitoring the driver but suffer from high complexity and adverse weather conditions. These sensors need contact with the skin which increases complexity in using them in low-end automobiles. The algorithm used for monitoring the driver suffers from the cyclist functioning and needs more validation techniques for the data acquired. The sensing technology is affected by adverse weather conditions and the algorithm has high computing complexity when used in low-end automobiles. This method of sensing technology is not ideal for this research. Non-Vision sensors such as radar for driver monitoring systems are newly emerging sensor technology in ADAS. These systems have less accuracy and high computing complexity. The sensing technology needs high computation to classify the distraction detected. The existing algorithms have issues in differentiating the action of the driver. The accuracy of such a system is low compared to other sensing technologies. This makes this sensor technology not suitable for this research.
Camera-based driver monitoring systems are widely used in high-end automobiles for the past decade. The complexity of the system can be altered based on the image processing algorithm. The accuracy, performance, and cost of the system depend on the algorithm and can be configured. It can be concluded that Visual sensor technology is more suitable for this research. The suitable algorithm for this research has to be planned based on the accuracy, computational complexity, and performance needed. The comparison between the different sensors used for driver monitoring is shown in Fig. 3.4. Though radar has good working in external conditions, it suffers due to the complexity of the

Sensors	Accuracy	Cost	Complexity	External Sensitivity
Driver Monitoring System				
Physiological Sensors	High	Low	High	High
Camera	High	Moderate	Moderate	Moderate
Mm-Wave radar	Moderate	High	High	Low

Figure 3.4: Sensor Technology study-1

system. Combining the radar system in front and the driver monitoring system needs to have low computing complexity and cost.

3.3 Sensor Technologies and Evaluation Techniques in Forward Collision Warning Systems

The sensing technology for forward collision warning systems depends majorly on the range, accuracy, and target's characteristics needed. Thus, it is necessary to analyze different technologies to understand the trade-offs and select the best-suited sensor. In this section, a brief description of the three major categories of sensor technologies in Forward Collision Detection is given to gain a better understanding of the following literature.
The three major categories of sensor technologies in Driver Monitoring systems are as follows:

- **Visual Sensors:** In this category, the obstacle in the front of the vehicle is tracked using a camera. This system depends on the lighting of the environment, weather conditions. The accuracy of the system to calculate the velocity of the obstacle is based on approximation. A Camera sensor made up of Charge Coupled Device (CCD) or Complementary Metal-Oxide Semiconductor (CMOS) technology helps in recognizing vehicles, obstacles, and lane markers. The range of the sensor purely depends on the camera capability and the image processing technique used which is generally up to 50m. Types of Visual Sensors used for Forward Collision Warning systems are
  - **Stereo Camera:** The stereo camera consists of two camera sensors that are separated horizontally by a predefined distance. The range of the object in the frame is measured depending on the position of the object in both the

cameras. Image processing is used to detect the type of object in the image. The accuracy of the system depends on the resolution of the camera.

- **Monocular Camera:** A Monocular Camera uses a single camera sensor to detect obstacles in the front. The range of the obstacle from the host vehicle is estimated based on the size and position of the object in the frame using pinhole geometry. The speed of the vehicle is also calculated by the change of pixel and/or the change in the size of the object in the frame. The problem with this type of sensing is its estimation of range and speed reduces the accuracy and increases the load on the system.

- **Non-Visual Sensors:** This category includes sensors that capture the obstacle in the front through wave propagation technologies. These types of systems are gaining increasing popularity due to their lightweight, low energy consumption, and robustness. Such systems can accurately determine the range, velocity, and angle of the obstacle in the front. Some of the common wave propagation technologies used in forward-collision warning system are as follow

  - **Ultrasonic Sensor:** Ultrasonic sensors are the most cost-effective technology used for a long time in ADAS systems. The system generates a bust of ultrasound which gets reflected by the object. This echo signal received time is used to calculate the distance between the host vehicle and the target object. The distance of the object from the host vehicle is calculated from the time taken for the signal to receive and the speed of sound in the air. The range of the ultrasonic sensor is short-ranged around 2 meters and this is affected by environmental variables such as humidity, temperature, and wind speed. The operation of an Ultrasonic Sensor Technology is shown in Fig. 3.5.

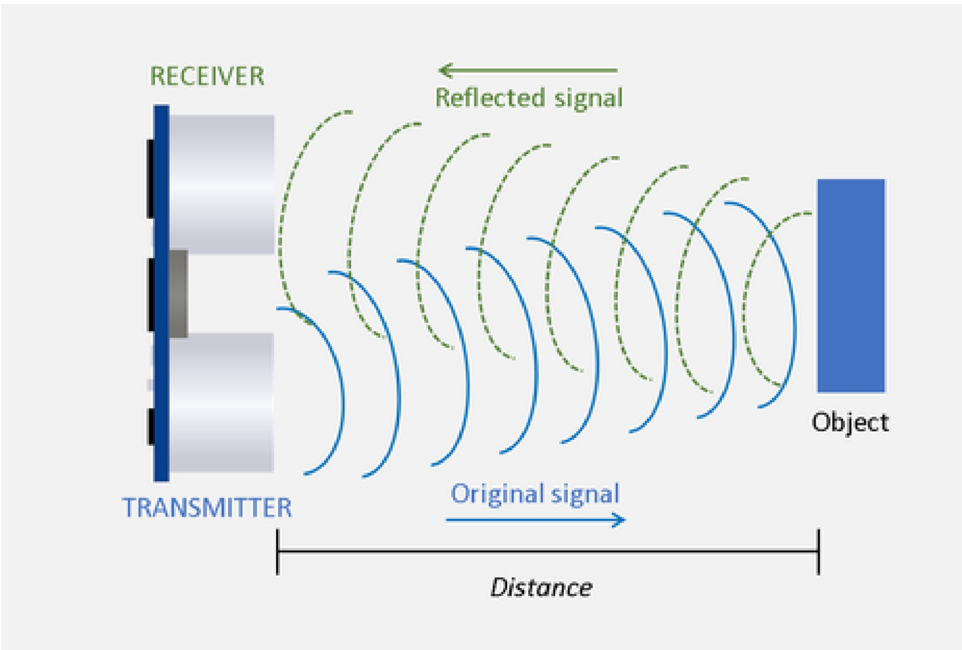


Figure 3.5: Ultrasonic Sensor Working [4]

– **Radio Detection and Ranging (RADAR):** RADAR sensors are electromagnetic sensors that transform microwave echo signals into electrical signals. Radar can help in detecting the motion, position, shape, and motion trajectory of the object. It has a higher range compared to ultrasonic sensors and is safer for humans and animals. Through the use of Doppler effect, the motion and velocity of the object can be detected using radar. The object's relative distance and relative velocity can be detected using radar sensors. A FMCW RADAR uses carrier frequency modulation between a bandwidth to detect the velocity and range of the object through frequency comparison. The factors such as physical size, transmitted power, antenna beamwidth, and atmospheric attenuation vary with frequency. As the frequency is higher, the shorter the wavelength hence the Physical size reduces but transmit power increases. The

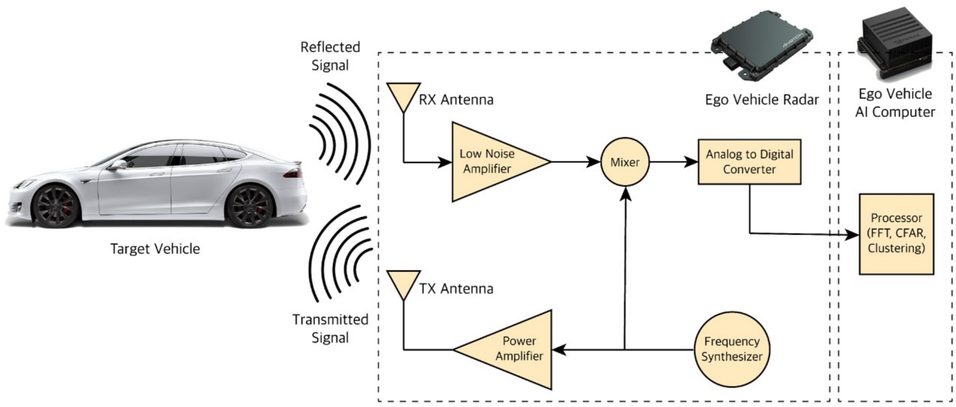


Figure 3.6: Radar Sensor Working [5]

operation of a Radar sensor Technology is shown in Fig. 3.6. The new FMCW radar uses a millimeter-wave band. The millimeter-wave band has a shorter wavelength due to which the system is small and has high detection accuracy. The mmWave RADAR has a Lower radar cross-section (greater stealthiness), Long target-detection range, All-weather capabilities, and greater reliability.

– **LIDAR :** LIDAR is a remote sensing method that uses light in the form of a pulsed laser to measure the distance between the host vehicle and the object. It forms a pixel-type view of the world known as a point cloud. The LIDAR generates high-speed pulses of laser around 150,000 pulses per second. LIDAR uses eye-safe laser beams to create a 3D representation of the surveyed environment. The time it took for each pulse to return is used to calculate the distance it has traveled. A LIDAR instrument principally consists of a laser and a scanner. The level of accuracy can vary from sensor to sensor depending on the power, configuration, and resolution of the unit. The operation of a LIDAR Sensor is shown in Fig. 3.7. LIDAR technology works well in rain and snow but suffers during adverse weather conditions. The disadvantage of the system is its propagation through materials. If the unit is obscured by an object in close range then there will be a loss of data. The cost of an LIDAR technology is high compared to other technologies used in ADAS systems.

- **Vehicular Communication :** A vehicular Communication system is the most popular new generation technology that is being implemented in automotive ve-

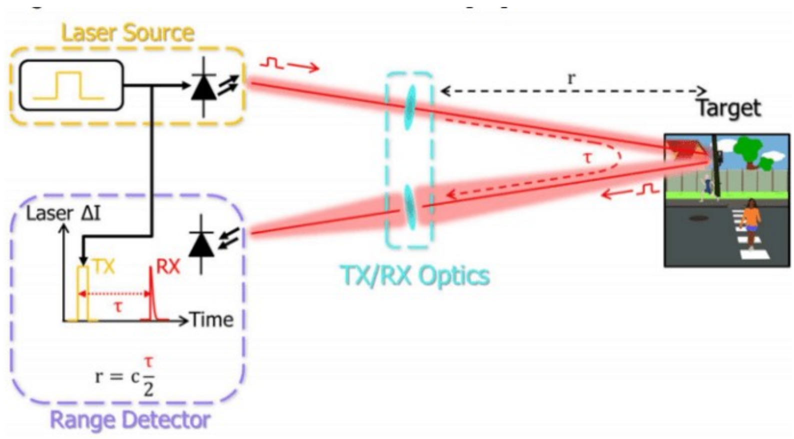


Figure 3.7: LIDAR Sensor Working [6]

hicles for communicating the ongoing traffic and safety warnings. The vehicular Communication system used in accident prevention is further categorized into the following

- **Vehicle-to-Vehicle (V2V):** V2V communication system enables communication between vehicles in real-time. The vital information such as steering angle, braking, speed, ongoing traffic, driver's behavior can be communicated to the nearby vehicles. This information can help the driver in judging the situation and taking action.
- **Vehicle-to-Road Infrastructure (V2I):** V2I communication system enables the communication between vehicles and road infrastructure. This is a bidirectional data exchange where the vehicle communicates the vital information of the vehicle while the infrastructure communicates the traffic lights, lane markers speed limits to the vehicle.
- **Vehicle to Pedestrian (V2P):** V2P communication system enables the communication between vehicles and pedestrians. The information such as walking pedestrians, children on the road are communicated to the vehicle to avoid accidents.

The operation of a Vehicular Communication Technology is shown in Fig. 3.8. The major disadvantage of this system is that all the vehicle in the environment needs to implement the same communication system to avoid accidents. The data traffic due to frequent communication can lead to data packet loss and communication errors.

3.3.1 Existing Solutions for Forward Collision Warning Systems

In this section, the Existing Forward Collision Warning system solutions related to the research are examined and criticized.

Vision Based Solutions

Shivam Kumar et al proposes a system for forward collision warning system [29]. The system detects the car in front of the host vehicle with the use of a camera. For identifying

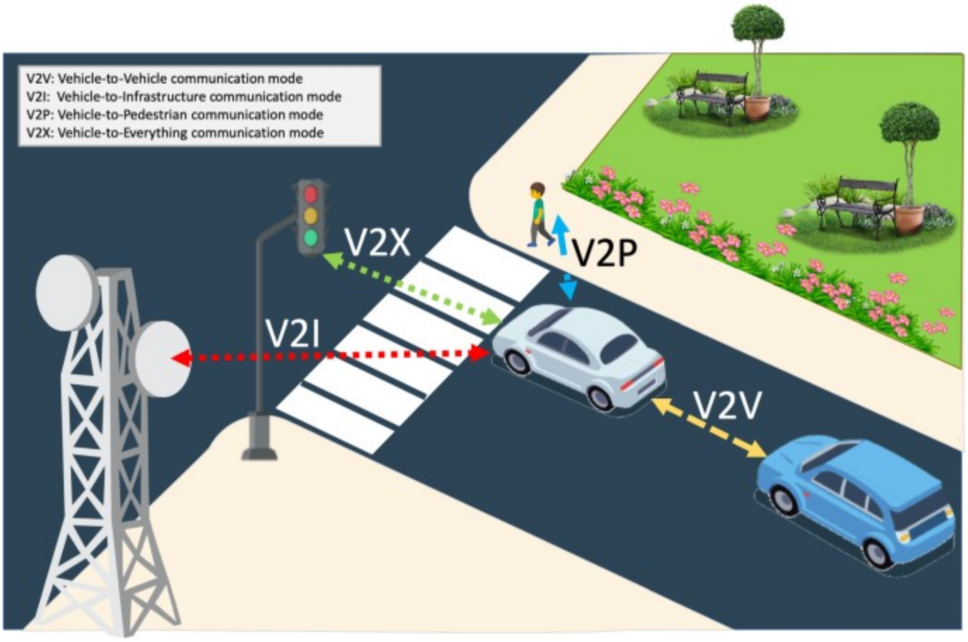


Figure 3.8: Vehicular Communication [7]

the distance between the cars, the system employs a CNN. The spacing between the cars is fixed at 1 meter, 1.5 meters, 2 meters, and 2.5 meters. When the gap between the cars is less than 1 meter, a warning is provided. This system was trained using two-car images and the threshold value changes depending upon the car model. The system was evaluated on a laptop, and no optimizations were made to make it suitable for use in an embedded device.

Huieun Kim et al proposes a system for vehicle and lane detection for forward collision warning [30]. The car is recognized and tracked in the second stage using the Hough forest algorithm and Kalman filter. This algorithm has a greater lane detection and vehicle detection accuracy with fewer false positives. When utilized in embedded devices, the system has a reduced frame rate, and the accuracy of the system is entirely dependent on camera calibration and illumination. The illumination in the surroundings affects vehicle and lane detection. Wai Chun Phoon et al proposes a system for detecting Forward collision [31]. Using a camera sensor, the system recognizes the vehicle in front of it and determines a safe distance between the host and target vehicles. The alert is sent using a buzzer. The system was tested in different lighting scenarios to increase the accuracy. The technique was incorporated into an embedded device and tested in real-time. The system suffers from issues such as the lighting conditions and the environmental noise which greatly influences the system's performance. The warning methodology given can greatly influence the driver's attention and led to distraction. Qun Lim et al, proposes a system using a monocular camera for Forward Collision warning [32]. The distance between the host car and the leading vehicle is measured using the camera sensor. The width of the vehicle is assumed as 1.8m and the width of the vehicle in the pixel is detected through the camera The camera data is double filtered using the Kalman filter to better estimate the distance and velocity between the target vehicle and the host. The accuracy of the system is affected by lighting, camera calibration, and focal length of the lens. Yuwei Lu et al proposes a system for forward collision warning using camera [33].

The device detects automobiles using a camera sensor. The camera is calibrated first, and then the car is detected. The camera's calibration ignores extension distortion, which leads to increased aspect ratio inaccuracies. In addition, the system was created on a laptop and not tested on an embedded device. Yeong-Kang Lai et al proposes a forward-collision warning system using camera [34]. The system uses a camera sensor for forward vehicle detection. HAAR algorithm is used to detect the features of the camera data and this is given as input to the cascade classifier. The target vehicle's rectangle bounds are produced, and the Euclidean distance between the target car and the host vehicle is determined. This Euclidean distance is used to determine the time to contact, which is used to issue a warning. Because the system relies on a camera, it is subject to camera calibration, weather, and illumination challenges, all of which impair the accuracy of vehicle detection. Sunghoon Hong et al proposes a forward collision warning system [35]. The system detects and tracks the vehicle in front using a camera sensor. Haar feature-based cascade classifier is used to extract the car in the sensor data. SMID processing is used to increase the processing speed and the multi-thread method is done to reduce the complexity of the algorithm. The system, complexity is high and the warnings method is not clear. The accuracy of the system is less compared to other FCW algorithms using a camera.

To summarize, vision-based forward-collision warning systems are upcoming driver assistance systems in automobiles. The vision-based system designed for high-end automobiles can help in obstacle classification and trajectory estimation. The accuracy of the system directly depends on the computing power and image processing techniques. The vision-based systems are affected by poor lighting and adverse weather conditions. The approximation of speed and distance can affect the performance of the system. For good performance of the system, high computational algorithms are needed to calculate the distance and speed. This research requires sensors that can provide considerably good accuracy, relatively lower computational complexity, and work in all adverse conditions. Hence, it can be concluded that this type of sensor technology is not suitable for this research.

Non-Vision Based Solutions

L.H. Eriksson et al proposes a forward-collision warning system using mmWave RADAR [36]. The radar-based system was used to gather, analyses, and tracks the object in front of it. The threat is assessed using the relative velocity and range between the host vehicle and the target vehicle. The system was tested in various weather conditions to test the performance of system. The system had low degradation in performance in adverse weather conditions. The lack of a risk assessment technique and restriction of the field of view leads to an increase in false alarms. P. Ganci et al proposes a forward-collision warning system using 77 GHz RADAR [37]. The mmWave RADAR detects and tracks the vehicle in front. The vehicle's relative velocity and range are determined. Before deciding whether a vehicle is in the same lane as the host car or not, the lateral distance to the vehicle ahead is determined using the Curve Algorithm, which is independent of road curvature. When compared to other forward collision warning systems, the system has a high level of accuracy, range, and cost. The technology was put to the test in both controlled and high-traffic environments. The system lacks a risk assessment technique to distinguish threat and non-threat vehicles which leads to a higher number of false

alerts. Thomas Fitzgerald Stevens proposes a system for forward collision warning using LIDAR [38]. This system uses the LIDAR sensor to obtain point clouds of data that represent the surrounding environment. These point clouds are segmented and obstacles are parsed out. The obstacle properties are searched in order to differentiate between hazardous obstacles and cone obstacles that mark the road boundaries. The system doesn't consider moving vehicles. The proposed system is expensive and has high computing complexity. Henrik Clasen et al propose a system for Forward and rear collision warning in e-bikes [39]. The system uses mmWave radar for forward and rear vehicle detection and tracking. Depending on the Time to Collision statistic, the rider receives a visual and auditory warning. For a 360° vision, two radars are used at the forward and back ends, and the radars are adjusted to detect up to 93m. The system expresses a lower number of tracking compared to other approaches due to the complexity of the algorithm utilized. The alerting mechanisms might be a source of additional distraction for both the host and the pursuing vehicle while driving.
To summarize, non-vision-based forward-collision warning systems are existing driver assistance systems in automobiles. These systems are tested in all adverse weather conditions and delivers a relatively good performance. The sensors in this category have their own set of advantages in terms of cost, detection range, accuracy, resolution, and susceptibility to bad weather. The computing complexity of the system is low compared to other sensor technologies. The speed and distance of the target object are given without the use of high computational algorithms. Hence, it can be concluded that non-vision-based sensors are suitable for this research. The sensors in non-vision technology are further compared in terms of range, accuracy, cost, and susceptibility to external factors are done in order to find the suitable sensor for this research.

Sensor Fusion Systems

Yimin Wei et al[40] proposes a sensor fusion system for frontal collision warning. The proposed system uses radar and LIDAR for estimating the longitudinal and lateral position of the target in cars. In an ideal situation, the data from both sensors are used for tracking the target's position. The gyroscope, accelerometer, speedometer data from the car are used to detect the motion of the target which is used for neglecting the unnecessary warning during turning or accelerating. Brake, throttle sensor data are used for validating the driver's behavior before and after the warning is given. The warnings are given when the Time to Collision is lesser than 2 seconds. This system has higher accuracy compared to other single sensor systems and avoids unnecessary warning during turning and acceleration. This system lacks the prediction of the driver's intention and whether or not the obstacle is seen by the driver. Moreover, the system warnings are given through a Human interface which might lead to more distraction of the driver.

Shigetaka Suzuki et al proposed a system based on camera and radar for pedestrian collision warning system [41]. The camera sensor is used to detect the crosswalk present in the field of view and combined with the radar sensor data, the position of the pedestrian and crossing speed of the pedestrian are calculated. The warnings are given if the Time to Collision is lesser than 4 seconds. This system detects and tracks the movement of a pedestrian in a crosswalk only. The warning is given irrespective of the driver's awareness of the object.
To summarize, sensor fusion technology used in forward-collision warning system show high accuracy compared to other systems. Sensor fusion technology has a wider range of

applications in ADAS-enabled automobiles. The existing sensor fusion technologies are mainly focused on increasing accuracy and performance which increases the computing complexity and cost of the system. Hence, the existing sensor fusion technologies are not suitable for this research.

V2V Communication

Xiang Yang et al proposes a system for forward-collision warning system using Zigbee technology [42]. The system communicates with the target car using two techniques and estimates the distance between them based on which warning is delivered. This technique was also employed during overtaking maneuvers when communication between the host and target vehicles assisted in avoiding collisions. All cars should be equipped with ZigBee technology as part of this system. If there is no communication between them, they may collide, resulting in an unending cycle. The host can only use the Received Signal Strength Indicator (RSSI) of a ZigBee if the target vehicle is in line of sight. The RSSI signal is packed with noise at greater ranges, which can lead to errors. Tangtao Yang et al proposed a system for forward collision warning using V2V technology [43]. For improved accuracy, the system employs Dedicated Short-Range Communications (DSRC) technology in conjunction with Differential Global Positioning System (DGPS). The DGPS calculates the car's longitudinal and latitudinal coordinates and sends them to the rear vehicle. The relative distance and Time to Collision metrics are determined using this sensor data. The TTC threshold is set at 2s, and if it drops below that, a warning is issued. The DGPS technology has a lower error rate than other technologies, and the system's precision is greater than other systems. Replacement of DGPS receivers with higher performance can also increase DGPS location accuracy. This system can only be implemented in a smart city and if and only if all the vehicles are equipped with it.

Ahmed Hosny et al proposes a forward collision avoidance algorithm using V2V communication technology [44]. The technology communicates acceleration, direction, and braking information, as well as Global Positioning System (GPS) data, to the following vehicle, which aids in the calculation of the Distance to Collision statistic. Three levels of warning are given by visual and audio signals depending on the distance between the leading and following vehicle. This system only functions in an infrastructure where all vehicles are equipped with communication radios. Because the warning signs include both a visual and an audible warning, they might be disturbing to the rider.
To summarize, Vehicular Communication is an upcoming research field in ADAS. The systems use inboard sensors to gather data and communicate to the outside environment. Vehicular Communication systems lack the basic infrastructure which increases the cost of implementation. All automobiles need to be equipped with the same vehicular communication system to enable the proper functioning of the system. Hence, it can be concluded that this type of sensor technology is not suitable for this research.

3.3.2 Comparison of Forward Collision Warning Systems

The comparison of different types of sensors that are used in forward-collision warning systems is concluded in this section. The vision-based system designed for high-end automobiles can help in obstacle classification and trajectory estimation. The accuracy of the system directly depends on the computing power and image processing techniques. The vision-based systems are affected by poor lighting and adverse weather conditions. It can be concluded that vision-based systems are not suitable for this project due to their poor performance in adverse weather conditions and low range. The approximations did for calculating the range and speed of a vehicle reduce the accuracy of the system. Sensor fusion technology has a wider range of applications in ADAS-enabled automobiles. The existing sensor fusion technologies are mainly focused on increasing accuracy and performance trading the computing complexity and cost of the system. The existing sensor fusion discussed in the literature is considered to be not suitable for this research. It can be seen that these sensor fusion system increases the computing complexity of the system and cost of implementation. Vehicular Communication systems are the new generation sensor technologies used in ADAS. Vehicular Communication systems lack the basic infrastructure which increases the cost of implementation. All automobiles need to be equipped with the same vehicular communication system to enable the proper functioning of the system. The methods discussed in vehicular Communication are considered to be not suitable for this research. These methods need pre-existing infrastructure for good performance.

Sensors	Range	Field of View	Accuracy	Cost	Complexity	External Sensitivity	Estimated Capability
Forward Collision Warning System							
24-GHZ radar	Moderate	Upto 120°	Moderate	Moderate	Moderate	Low	Distance,Speed,Angle
Lidar	High	Upto 360°	High	High	High	High	Distance,Speed,Angle
mm-Wave radar	High	Upto 120°	High	Moderate	Moderate	Low	Distance,Speed,Angle
Camera	Moderate	Upto 180°	Moderate	Low	Moderate	Moderate	Distance,Speed
Ultrasonic	Low	15°	Moderate	Low	Low	High	Distance

Figure 3.9: Sensor Technology study-2

The non-vision-based system has been used in ADAS for the past decade. The sensors in this category have their own set of advantages in terms of cost, detection range, accuracy, resolution, and susceptibility to bad weather. The computing complexity of the system is low compared to other sensor technologies. It can be concluded that Non-vision based systems are the ideal choice for this research due to their low cost, high accuracy, high range, and less complexity. The type of sensor in non-vision technology needs a bit more examination. The comparison between the sensor technologies is shown in Fig. 3.9. It can be concluded that mm-wave radar aligns with the requirements of this project. Though the LIDAR has a higher Field of View and Accuracy, the cost of the equipment and the sensitivity to external conditions makes it less suitable for this project.

3.4 Threat Assessment

The Risk Assessment system is used to determine the risk of collision from the target object. There are several techniques for risk assessment we will be addressing Deterministic approach

- **Deterministic Approach:** In Deterministic Approach, the warning is given as a binary value based on whether the collision will occur or not. It is a rule-based system that computes the output prediction by comparing a conservative estimate of prospective exposure to the threshold risk value. Most of the forward collision warning and mitigation systems implement a Deterministic Approach. The drawback in such an approach is its inability to explicitly model the uncertainties of its input data.
  - **Time-Headway (THW):** THW is calculated by dividing the range between the target vehicle and host vehicle divided by the velocity of the host vehicle. For safe driving, most Departments of Motor Vehicles suggests that drivers maintain a two-second time headway. Time-headway is a useful metric because it corresponds to the amount of time the driver of a host vehicle has to match the braking profile of the lead vehicle.
  - **Time To Collision(TTC):** Using a time-to-collision criterion for triggering forward collision warning alerts is based on the theory that humans directly perceive time-to-collision. Time-to-collision is calculated by dividing the range between the host vehicle and the target vehicle to the relative velocity between them.
  - **Kinematic Constraints Criterion (KCC):** Using the instantaneous target- and host-vehicle speeds and accelerations, these algorithms calculate the minimum collision-avoidance range that will result in the host vehicle missing the lead vehicle by a specified margin. The minimum collision-avoidance range is calculated by assuming the brake rate and brake reaction time of the vehicle.
  - **Minimum Safe Distance(MSD):** This algorithm is based on distance. It is calculated as the minimum distance that needs to be between the host vehicle and the target vehicle.
- **Probabilistic Approach:** The probabilistic approach is used to get more precise estimates of danger frequency and damages. Probabilistic assessment depends on uncertainty and randomness of the danger. Probabilistic assessment simulates the disaster that how likely it would occur. The algorithm takes into account all the possible scenarios and occurrences of an accident. The algorithm suffers from the computing complexity and training of the algorithm. Fuzzy logic, Markov processes, and Bayesian networks are some of the most common methods used in probabilistic approaches.

3.4.1 Comparison of Threat Assessment Techniques

A Threat Assessment Technique is needed to identify potential threats and alert the rider. A Threat Assessment algorithm calculates the risk of collision of the host vehicle from the target object. Existing solutions for threat assessment are categorized into probabilistic and deterministic approaches. The probabilistic approach are complex and has high computing power. They are termed to be not suitable for this research. The Deterministic Approach are less complex and has good performance.

Criteria	TTC	KCC	THW	MSD
Complexity	Low	High	Moderate	low
Acceptability	High	Low	Moderate	Moderate
Accuracy	High	High	Low	Moderate

Figure 3.10: Deterministic Approach Metrics Comparison

The Fig. 3.10 shows the comparison of the commonly available metrics. The KCC metric calculates the minimum collision avoidance range through the use of parameters such as brake rate, brake reaction time, and acceleration of the host and the target vehicle. This makes the algorithm too complex for a low-end vehicle. The THW metrics takes into consideration only the host vehicle speed which reduces the accuracy of the system. Thus, MSD and TTC have good accuracy and performance for long-range and short-range warnings. Combining these methods leads to less complex threat assessment metric which would help for long range and short range.

Chapter 4

Theory and Background

4.1 Radar Fundamentals

The IWR1642 uses Millimeter-wave radar technology to detect an object in real-time. Millimeter-wave radar uses short-wavelength electromagnetic waves for determining the range, velocity, and angle of the object. As the mmWave radar uses wavelength in the millimeter range, the size of the system and antenna are small compared to other radar systems. The operating frequency of a mmWave radar is from 76-80 GHz. The system overview is shown in Fig. 4.1. IWR1642 is an FMCW radar that transmits frequency-modulated signals continuously to measure angle, range, and velocity.

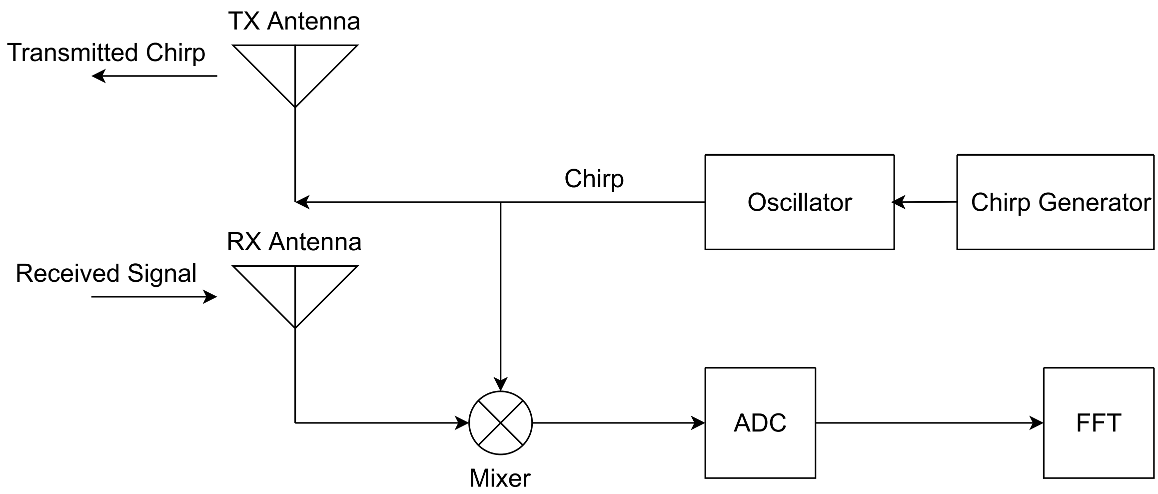


Figure 4.1: Radar System Overview

In FMCW radar, a chirp signal is transmitted to obtain the velocity, range, and angle of the object. A chirp is a signal in which the frequency increases gradually. The chirp signal is transmitted and on reflection from an object, the reflected chirp is received by the radar. The transmitted signal and received signal is combined in the mixer. The signal from the mixer which is the combination of the transmitted and received signal is called the IF signal.

$$IF = sin[(\omega_1 - \omega_2)t + (\phi_1 - \phi_2)] \tag{4.1}$$

The phase of the Intermediate Frequency(IF) is the phase difference between the transmitted and received signal. The time delay in the received signal is formulated in Eq. (4.2), where d is the distance to the object and c is the speed of light.

$$\tau = \frac{2d}{c} \tag{4.2}$$

The initial phase of the IF signal is the difference between the phase of the Transmitter (TX) chirp and the Receiver (RX) chirp at that time instant corresponding to the start of the IF signal. The initial phase of the IF signal is given by

$$\phi_0 = \frac{4\pi d}{\lambda} \tag{4.3}$$

where d is the distance from the object and λ is the wavelength of the signal. The IF signal is given as

$$\text{IF} = Asin(2\pi f_0 t + \phi_0) \tag{4.4}$$

This IF signal is processed using Fourier Transform to separate different tones. On Fourier Transforming the result signal consists of peaks of different tones which denotes each object detected. The range, velocity, angle estimation that was done using the received signal is discussed in the following section.

4.1.1 Range Estimation

Range plays a significant role in forward collision warning systems. The maximum range of a radar system depends on the Signal to Noise Ratio (SNR) and the maximum intermediate Frequency(IF) supported by the system. The beat frequency is the absolute difference between the transmitted and received chirp. The beat frequency is directly proportional to the distance of the system. The formula to determine the beat frequency is given is Eq. (4.5).

$$f_b = \frac{S2d}{c} \tag{4.5}$$

Therefore the maximum range Rmax can be derived from Eq. (4.6) depending on its maximum IF bandwidth by

$$R_{max} = \frac{IF_{max}c}{2S} \tag{4.6}$$

Where S is the slope of the transmitted chirp and c is the speed of light. The maximum range that can be detected also depends on the SNR of the signal. The Rmax is given by the formula Eq. (4.7)

$$R_{max} = \sqrt[4]{\frac{P_{tx}G_{tx}G_{rx}\lambda^2 N T_c \sigma}{(4\pi)^3 SNR_{min} kTF}} \tag{4.7}$$

Where Ptx is the transmission power, Gtx and Grx are the transmitter and receiver antenna gain,σ is the Radar Cross Section (RCS) of the object to be detected, k is Boltzman Constant, T is antenna temperature, λ is the wavelength ,T_c is Chirp Time, N is number of chrips in a period. The maximum IF bandwidth of a system is limited by

$$IF_{BW} = 0.9 * ADC_{sampling} \tag{4.8}$$

4.1.2 Velocity Estimation

To detect the velocity of the target object, the FMCW radar transmits two chirps separated by a time T_c. On processing the received signal to detect the range of the object through Fast Fourier Transform (FFT), the velocity of the signal can be calculated from the difference in phase of the signals. The phase difference is formulated as

$$\Delta\phi = \frac{4\pi v T_c}{\lambda} \tag{4.9}$$

The velocity of the detected object can be calculated from the above formula such as

$$v = \frac{\Delta\phi\lambda}{4\pi T_c} \tag{4.10}$$

For higher velocity, the transmission time is reduced. In real-time to detect the velocity of multiple objects, N equally spaced chirps are transmitted. The Range-FFT results in a set of N peaks with a different phase which incorporates the velocity of the detected objects. The Range-FFT of two objects in equal distance is shown in Fig. 4.3, it shows

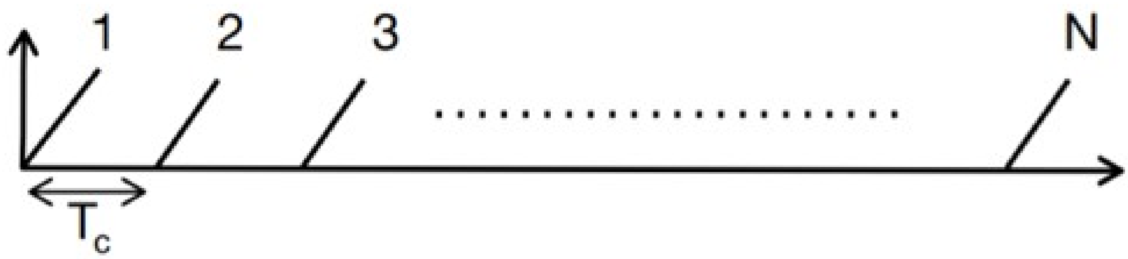


Figure 4.2: Radar Frame [8]

the phase contributions of each object. On applying a second FFT named Doppler-FFT on the Range-FFT to separate the two objects.

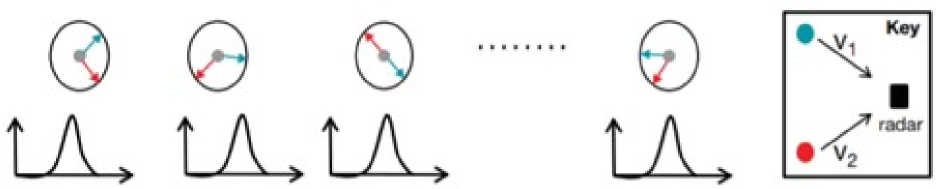


Figure 4.3: Objects in equal distance [8]

4.1.3 Angle Estimation

An FMCW RADAR can be used to estimate the angle of the reflected signal. Angular estimation is based on the observation that a small change in the distance of the object results in a phase change in the Range-FFT peak. On using at least two RX antennas, the differential distance from the object to each of the antennas results in a phase change in the peak of Range FFT. The angle of Arrival is calculated as

$$\theta = sin^{-1}(\frac{\lambda\Delta\phi}{2\pi l}) \tag{4.11}$$

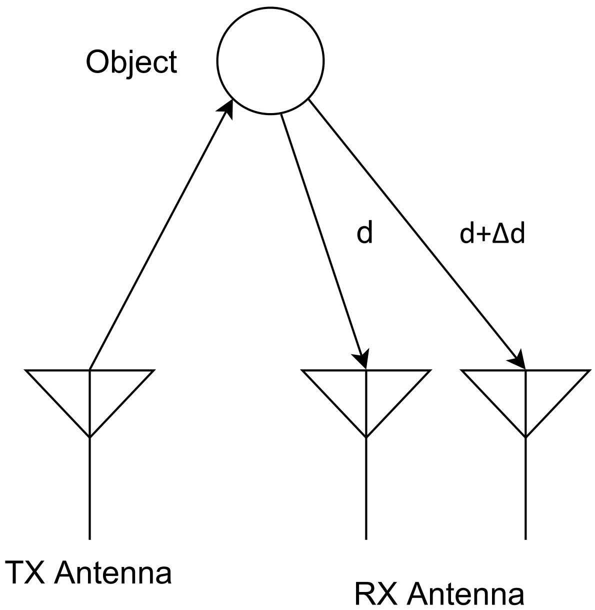


Figure 4.4: Forward Collision Warning system

4.2 Head Pose estimation Fundamentals

The Head pose of the rider is estimated by using three components such as face detector, landmark predictor, and Perspective -N- Point(PNP) method.

4.2.1 Face Detection Algorithms

This section explains the background on the different face detection algorithms present and the performance of the system for the research question at hand.

HAAR Cascade Based Face Detector

The first-ever object detector was proposed by Viola and Jones which was based on HAAR Wavelets [45]. A HAAR Wavelet is an array of square-shaped functions. This feature continuously transverses through the image to find the change in intensity of the pixels. The darker pixels are given a value of 1 and lighter pixels are given a value of 0.

Other intensities are given a value between 0 to 1. The difference between the sum of the darker pixels and the lighter pixels is used to detect edges in the image and by which the outline of the image can be estimated. An AdaBoost Technique was used to apply 180,000 features on the image to create separate Learners which helps in reducing the set of features to 6000 features. These features are separated into stages and continuously transverse on the image to detect facial features. As the feature set increases each stage it reduces the false positives.

Histograms of Oriented Gradients for Face Detection

A HOG feature descriptor-based object detector is widely used in computer vision. The HOG feature descriptor is used in object detection algorithms for high accuracy and less computing complexity. The change in intensity helps in detecting the edges of an object in images. A Gradient is calculated within an image per block. A block is a grid of pixels in an image with magnitude and direction based on the difference in intensity of the pixel inside that block. For face detection, a window size of 128 x 144 is used to match the aspect ratio of the face. The HOG descriptors are calculated for 8 x 8 dimensions of pixels. This 8 x 8 pixel is called a block. Each descriptor value in the blocks are quantized into bins which represents the directional angle and magnitude which is the summation of all values in the bin. The histogram values are normalized over a 16 x 16 size to minimize the accuracy drop due to light conditions. An SVM model is trained using a number of HOG vectors for multiple faces.

Deep Learning Methods

Deep Learning methods of face detection have a lot of advantages such as high accuracy, robustness, and precision. The existing state-of-the-art algorithms used are discussed in this section.

CNN based face detector

The Multi-task Cascade Convolutional Networks (MTCNN) is a popular algorithm that is based on [46]. The algorithm consists of three stages of convolutional networks for detecting faces. The MTCNN is a faster approach in CPU compared to other algorithms and can detect landmarks such as eyes, nose, and mouth. In the first stage, a shallow CNN is used to quickly produce the candidate windows. In the second stage, the windows without the face are refined through a more complex CNN. In the third stage, more complex and powerful CNN is used to refine the result.

In Stage 1, a fully convolutional network named Proposal Network(P-Net) is used to obtain the windows and the bounding box regression vector. A Non-Maximum Suppression (NMS) is used to merge the overlapped bounding boxes.

In Stage 2, Refine Network(R-Net) is used to reject the false positives and perform calibration with the bounding box regression and NMS merges.

In Stage 3, more refining is done to the candidate window to find the landmarks in the face.

The filters used in the algorithm are reduced to 3x3 filters to reduce the computing complexity while increasing the depth for better performance. Initially, the image is resized for the first stage. The stages in the model are not connected instead, the outputs of the previous stage are fed to the next stage. The Accuracy of the system is high compared to Viola and Jones method and the HOG+SVM method. But in real-time the frames per system are low.

Deep Neural Network (DNN) based face detector

Deep Neural Network-based face detectors are of high accuracy and robust compared to other existing solutions. The most commonly used face detector using Deep Neural Network is based on the Caffe model. A Caffe model is based on Single Shot Detector (SSD) which is the top layer. The backbone of the model uses ResNet-10 or Visual Geometry Group (VGG)-16.

To the base VGG network, additional convolutional layers are added at the end of the base network for detection. The Convolutional layers decrease in size gradually till the end which helps in the detection of objects at multiple scales. The feature map size decreases progressively which helps in finding objects more accurately. The base network performs the feature extraction and creates multi boxes on the image. During training, the boxes with the highest overlap are chosen where Intersection over union (IOU) between the predicted boxes and ground truth should be more than 0.5. To the predicted boxes, Non-Max suppression is done to reduce the number of detection. Initially, the image is divided into grids, and the location of an object is done in these grids. SSD-based DNN face detectors are of high accuracy and robustness but introduce latency in real-time when used in embedded devices.

4.2.2 Histogram of Oriented Gradients Face Detection

Various Face detection algorithms are used in computer vision for an accurate estimate of head pose estimation of the rider. The algorithm used are explained in Section 4.2.1. Histogram of Gradient Orientation is used in this algorithm since it has moderate accuracy and good frames per second compared to other algorithms. The Histogram of Oriented Gradient method was suggested by Dalal and Triggs [47]. The face detection system is shown in Fig. 4.5

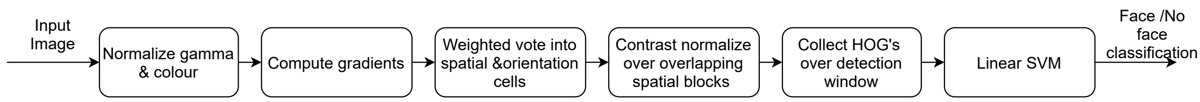


Figure 4.5: Face Detection Flow

Gamma/Colour Normalization

The Gamma/Colour normalization is used to improve the performance of the algorithm. A Gamma/Colour normalization block reduces the impact of light on the image. The input image should be in RGB or LAB color as this method has good results in these metrics compared to grayscale images.

Gradient Computation

Gradient computation helps in extracting useful information from the image such as silhouettes, contour, and textures. This information is used to determine the object to

uncentered	[-1,1]
centered	[-1,0,1]
cubic-corrected	[1,-8,0,8,-1]

Table 4.1: Derivative masks

be detected. The gradient computation is done using Gaussian smoothing followed by a derivative mask. Many derivative masks are used, some are listed in Section 4.2.2. The most common derivative used is the 1-D $\begin{bmatrix}-1, 0, 1\end{bmatrix}$ mask. However, it was shown in [2] that the most effective technique for gradient computation is the application of a 1-D, centered, point discrete derivative mask to filter the color or intensity of the image. The gradient magnitude and direction are calculated in each pixel using Eq. (4.12).

$$\begin{aligned} Gradient\ Magnitude &= \sqrt{g_x^2 + g_y^2} \\ Gradient\ Angle &= arctan(\frac{g_y}{g_x}) \end{aligned} \tag{4.12}$$

The g_x and g_y are the respective horizontal and vertical components of the change in pixel intensity.

Spatial/Orientation Binning

The magnitude and angle computed in the gradient computation part are used in this block to generate the histogram of oriented gradient. The picture is divided into small spatial sections called cells and a local 1-D histogram of gradient orientation is accumulated in a 1D histogram with a predefined number of bins. According to its orientation, each pixel votes for one of the histogram bins, and the magnitude is increased. These orientation bins are uniformly spaced from 0 to 180. The Histogram of Oriented Gradients with SVM method consists of 9 bins for spatial/orientation binning.

Normalization and Descriptor Blocks

Gradient strengths vary over a wide range due to the light variations, hence normalization is used to reduce these issues and increase performance. The cells are grouped into blocks. Two arrangements are used for normalization, Rectangular-Histogram of Oriented Gradients (R-HOG) and Circular-Histogram of Oriented Gradients (C-HOG) The R-HOG are calculated using $n \times n$ pixel cells and histogram bins in $m \times m$ grids (m being the number of cells in each block). The 2×2 cells of 8×8 pixels with histogram bins are used idly in this algorithm.

Detector Window and Context

A Detection window of 64x128 pixels is used for detecting the object. Decreasing this window affects the performance of the system Decreasing it from 16 to 8 pixels reduces the performance by 6%.

Classifier and Non-Maximum Separation Algorithm

A soft linear with SVM Light is trained using the input for high performance. A Gaussian kernel SVM is used when performance is to be increased. The SVM algorithm creates a hyperplane that separates the data into classes. Support vectors are the data points closer to the hyperplane. These points help in defining the separating line better by calculating margins. The raw data produced by the Histogram of Oriented Gradients are given to the SVM to classify the points as a face or not. The non-Maximum Separation Algorithm is used to reduce the number of bounding boxes generated for an object. The predicted boxes are compared and IOU values are compared. If the IOU is greater than a threshold then that prediction is selected, The threshold set is 0.5%. Hence, the confidence values greater than this threshold are selected. An example of an image after applying the HOG filter is shown in Fig. 4.6

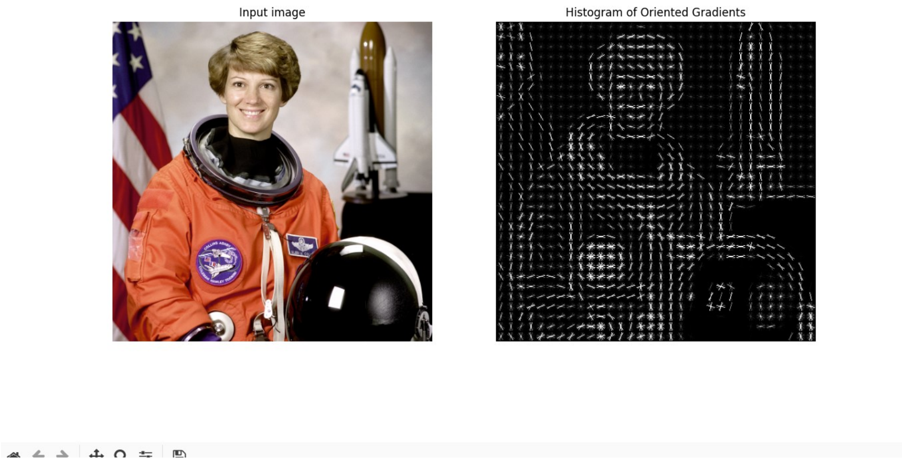


Figure 4.6: HOG feature example [9]

4.2.3 Landmark Prediction

The landmark prediction is an algorithm that is a part of the head pose estimation process. The most commonly used landmarks are the 68 landmarks present in the face. The vector of all landmarks is denoted by S=$\left(x_1^T, x_2^T, ..., x_p^T\right)^T \epsilon \mathbb{R}^{2p}$ where x_i denotes the x, y coordinates of the landmark. The $\hat{S}^{(t)}$ is used to denote the current estimate of S. A cascade of regressor tree is used to predict the landmark. Each regressor predicts an update vector and is added to $\hat{S}^{(t)}$ to improve the estimate. The regressor predictions are based on pixel intensity. The initial landmark vector is chosen by the mean shape of the training data-centered and scaled according to the bounding box output of the generic face detector., The regressor is trained using a gradient tree boosting algorithm with the sum of square error loss. The training data set is made into triplets of face images to generate the regressor function. This triplet of face images is updated for each regressor until a cascade of regressors are learned which when combined gives high accuracy.

Each regression function rt is built around the tree-based regressors that were fitted to the residual targets during the gradient boosting procedure. In the regression tree, at each split node, a decision is made based on thresholding the intensities difference of two pixels. The input image is warped to the mean shape depending on the current shape estimate. The sum of square error is computed between the mean-shape facial landmark points and the warped shape. A random set of candidate split is used to train the regression tree where the candidate with minimum mean square error is selected. The gradient boosting algorithm is modified to account for weight factors due to missing labels. This is done by initializing the model with a weighted average of targets and fitting regression trees to the weighted residuals.

4.2.4 PNP method

Perspective-n-Point(PNP) is the most common method is estimating the pose of an object. On using landmark predictor, the 2d projections of the landmarks are given. This 2d projection and the 3d points of those projects in the real world are used to estimate the pose of the object. The camera has to be calibrated for this method to work. 6 degrees of freedom are given based on this method such as the rotation vector and translation vector. The rotation and translation vector with respect to the world is given by

$$s\,p_c = K\,[\,R\,|\,T\,]\,p_w. \tag{4.13}$$

where $p_w = \begin{bmatrix} x & y & z & 1 \end{bmatrix}^T$ is the 3d coordinates of the object,$p_c = \begin{bmatrix} u & v & 1 \end{bmatrix}^T$ is the corresponding image point, K is the calibrated camera parameters, s is the scalar factor for the image point and R, T are the rotation and translation vector of the 3d coordinates in real time. The Eq. (4.13) can be transformed into:

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & \gamma & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \tag{4.14}$$

where f_x and f_y are the focal length in x and y direction and v_0, v_0 are the optical center of the camera. This equation is solved by using Direct Linear Transform (DLT). Direct Linear Transform(DLT) is used in solving the above equation and finding the unknowns when ≥ 6 points are given. The observed image coordinates is given by

$$x_i = PX_i \tag{4.15}$$

where on expanding the matrix we get

$$x_i = P_{3x4}X_i = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} X_i \tag{4.16}$$

Combining the rows as vectors such as A, B, C and substituting in Eq. (4.16), the equation becomes, $x_i = \begin{bmatrix} A^T \\ B^T \\ C^T \end{bmatrix} X_i$ combining the elements of P into a vector we get $p = \begin{bmatrix} A \\ B \\ C \end{bmatrix}$ This helps in simplifying the equation into $a_{x_i}^T p = 0$ and $a_{y_i}^T p = 0$ Each point generates these two equations. This equation is solved using Singular Vector Decompositions which is a standard method. On using the singular vector decomposition approach, the singular values of the vector p are found. As P is a combination of the Rotation vector, intrinsic parameters, and Translation vector. On solving, p, rotation vector, and translation vector are found. Since DLT is not very accurate, Levenburg -Marquardt Optimization method is used to reduce reprojection errors. The Levenburg -Marquardt Optimization randomly changes the pose and the reprojection error is checked. If the reprojection error is decreased, then this estimate is accepted as a new estimate of the pose. If not, the pose is changed again until the reprojection error reduces.

Chapter 5

Proposed Forward Collision Warning System with visual distraction detection in bikes

In this chapter, a brief description of the system model, challenges in implementation, and the design trade-off considered in order to maximize the performance of the system are discussed.

5.1 System Model

The Forward Collision Warning System algorithm considers the RADAR data and Camera data to estimate the level of warning that is needed in the situation. The proposed system is made up of three parts such as Data Acquisition, Data Processing, and Thread Assessment. The overview of the System model is shown in Fig. 5.1. The Data Acquisition part consists of the RADAR module for detecting the object in front and a camera module is used to detect the head pose of the rider. Data Processing consist of two parts, namely RADAR data processing, and camera data processing.

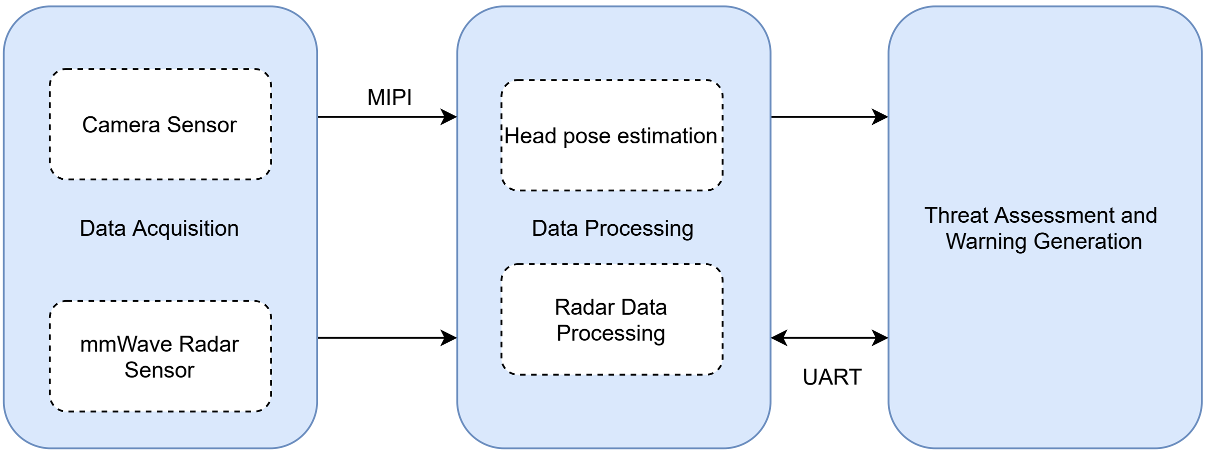


Figure 5.1: System Overview

The radar data processing consists of processing the frame acquired from the radar front-end and the Time to Collision(TTC) calculation. The camera data processing consists of a face detection module, landmark detection, and head pose estimation. The data acquired from the radar-backend and camera are processed using a raspberry pi Single Board Computer. The following section explains in-depth the system's components and the implementation.

5.2 Data Acquisition

The sensing technology used in detecting the obstacle in the front and head pose of the rider is done in this component of the system. The camera data acquired is transmitted through Mobile Industry Processor Interface (MIPI) protocol and for mmWave radar Universal Asynchronous Receiver Transmitter (UART) protocol is used. The radar and camera sensor data acquisition is explained in the following section.

5.2.1 Radar Data Acquisition

The IWR1642 mmWave radar sensor from Texas Instrument is used in sensing the environment. IWR1642 booster pack consists of a single chip mmWave radar sensor with Digital Signal Processor (DSP) module, ARM processor. The IWR1642 uses Frequency Modulated Continuous Wave radar technology. FMCW radar has higher accuracy, robustness, and rapid sensing. The range calculation for a Continuous Wave radar lacks timing marks which reduce the accuracy compared to FMCW radars. The Ti RADAR module which uses RFCMOS mmWave sensor can store 512 chirps before the frame starts compared to SiGe-based solutions which are limited in the storage of chirps. The Ti radar can be configured with multiple configurations settings which helps in extracting more data from the environment. The radar system consists of a front end and a back end. The IWR1642 is a self-contained, single-chip solution that facilitates the deployment of mmWave sensors in the 76 to 81 GHz frequency range. A monolithic implementation of a 2TX, 4RX system with built-in Phase-Locked Loop (PLL) and A2D converters are included in the IWR1642. The radar signal processing is handled by TI's high-performance C674x DSP. Front-end configuration, control, and calibration are handled by an R4F-based CPU subsystem. The Micro Controller Unit (MCU) and DSP peripherals can be programmed with integrated peripherals(UART, Controller Area Network (CAN), Inter-Integrated Circuit (I2C)). The IWR1642 sensor evaluation board is flexible and easy to use which eases the complexity in prototyping.

The IWR1642 sensor consists of two components such as a front end and a back end. The front end defines the sensing capabilities of the system. The tuning of the RF radio for the transmission and reception of the radar frames is done in the front end of the system.

As discussed in Chapter 4, the maximum velocity and range that can be detected are computed. The configuration for the system is shown in Table 5.1. Using this configuration results in a 5000ksps sampling rate in ADC. The maximum range that can be achieved is theoretically computed as 120m. Increasing the range of the system reduces the resolution of the system. Which leads to reduced resolution in the near field. Increasing the Resolution decreases the maximum distance that can be detected. The received signal is processed and the environment information is extracted in the back end of the system. The configuration of the front end and the data transfer is done on the integrated ARM R4F microcontroller.

Parameter	Value
Start Frequency (GHz)	76.01
Analog-to-Digital Converter (ADC) start time (s)	4.8
Ramp end time (s)	56
Number of ADC samples	256
Frequency slope (MHz/s)	5.6
ADC sampling frequency (ksps)	5000
Bandwidth (MHz)	409

Table 5.1: Front-end configuration

5.2.2 Camera Data Acquisition

A raspberry pi camera is used in collecting the head pose of the rider. Raspberry pi camera consists of Sony IMX219 8-megapixel sensor. The Sony IMX219 sensor has high image quality, color fidelity, and low light performance and supports 1080p resolution at 30 frames per second and 720 resolution at 60 frames per second. IMX219 sensor consists of CMOS active pixel sensor. The sensor has low power consumption, high sensitivity. The preexisting method for acquiring the camera data is through the OpenCV library method. This method has a low frame rate issue due to its blocking operation. Creating a separate thread to access the latest frame from the camera helps in reducing the latency between frames. This method of accessing the camera was used in this model to reduce latency in the camera processing.

5.3 Data Processing

This block is responsible for analyzing the data acquired from the sensors. It consists of a Camera and radar processing components. The camera data processing component consists of a face detector, face landmark predictor, and head pose estimator. The radar processing component consists of the radar back-end responsible for estimating the range, velocity, and position of the detected object.

5.3.1 Camera Data Processing for head pose estimation

In this component, the frames acquired from the camera are processed using a face detection and prediction algorithm. The face detection is done using the Histogram of Oriented Gradient method combined with SVM. The face detection algorithm was custom trained. The custom-trained algorithm was combined with a preexisting algorithm used by dlib library for better performance.

Face Detector

This section explains the method in finding the suitable face detector algorithm for our system. There are various methods in detecting the face in a frame in real-time. The background on these methods are discussed in Section 4.2.1. The Face detection is done using Histograms of Oriented Gradients(HOG) with Support vector machines(SVM). The

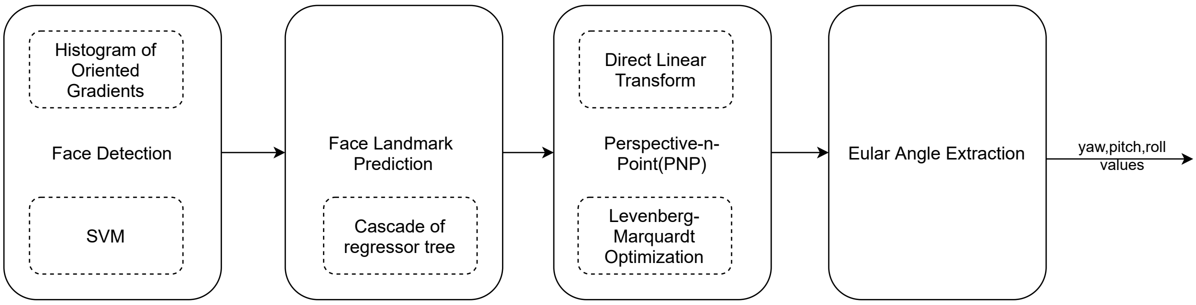


Figure 5.2: Camera data processing

dlib method which uses the HOG+SVM algorithm is implemented in this system due to its accuracy and high frames per second in embedded devices.

The face detector was trained using dlib library training dataset and custom dataset. A custom dataset was created from a video captured during cycling for training the algorithm. The detection rate of the existing face detector and custom dataset trained face detector was observed. The custom face detector shows a high detection rate and score compared to the existing dlib face detector.

Face Landmark Predictor

Facial Landmarks are shown in Fig. 5.3, these landmarks are used in detecting the head pose of the person. The face detector returns a rectangle with x, y, w, h parameters

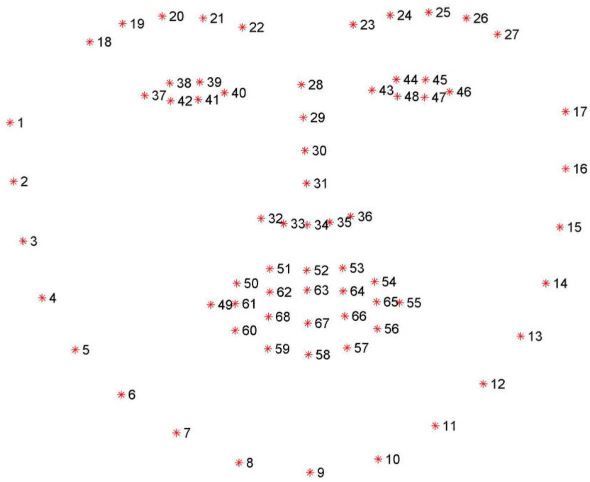


Figure 5.3: 68 Facial Landmarks [10]

which are used in this algorithm. The x, y, w, h coordinates of the face in the frame is parsed to the predictor algorithm to specify the region of interest where the face is present. The algorithm is trained using a labeled facial landmarks dataset that contains the x, y coordinates of each landmark in a pts format. This dataset is used to train an ensemble of regression trees to estimate the facial landmark position. Each regressor in the cascade is used to predict an updated value to the current shape estimate in order to improve the estimate. The regressor predictions of the features in the image are based on the pixel intensity value. This algorithm was pre-trained using 300W Dataset[48] and its mirror images for high accuracy and precision.

Head Pose Estimation

The Head pose estimation is done using the Perspective-n-Point method. The Perspective-n-point problem helps in generating the Rotation and Translation Vector of the object. A Rotation Vector is the rotation of the camera with respect to a still object and vice versa. The Translation Vector is the movement of the camera from its previous coordinates to a new location. The Rotation and Translation Vector has three degrees of freedom in X, Y, and Z coordinates. The Perspective-n-point problem is done assuming that the camera is calibrated. To compute the head pose of the person in the frame, the 3d coordinate and the 2d coordinate of the facial landmarks are needed. The 2d coordinates of the facial landmarks in the image are generated by the face landmark predictor. The 3D coordinates of the facial landmarks are derived from an anthropometric 3D rigid model of the human head [49]. The anthropometric 3D rigid model as 58 3D coordinates which is shown in Fig. 5.4 The estimation done using the 2d and 3d coordinates is used to

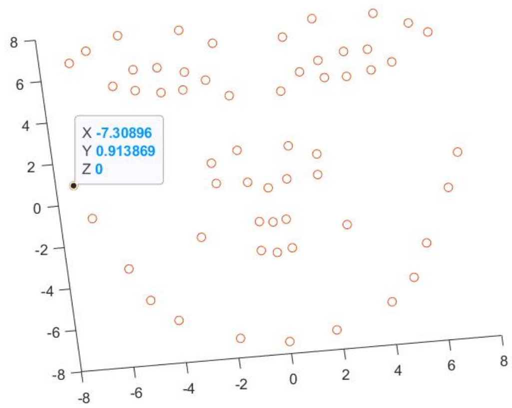


Figure 5.4: Anthropometric 3D rigid model

track the head movement. The rotation vector gives the rotation in 9 axes which are converted to 3axis such as pitch, yaw, and roll. This angle is used to denote the head movement of the rider. On testing the systems depending on the movement of the head, a threshold was set for left and right movement. This change in head movement is taken into consideration during the thread assessment.

5.3.2 Radar Data Processing

The processing of the RX signal is done by the integrated C674x DSP. The Back-end is responsible for the processing of the frames from the front end. The data received from the front-end is processed through Range FFT and Doppler FFT to estimate the Range and Velocity of the object. This data is pruned using the Constant False Alarm Rate(Constant False Alarm Rate (CFAR)) algorithm to detect and group the range and velocity estimates. CFAR is an adaptive algorithm used to detect target data against noise, clutter, and interference. The process chain with the front end is shown in Fig. 5.5.

For range direction cell-average smallest of- CFAR is used. The cell-average-smallest-of(Cell Averaging Smallest Of (CASO))-CFAR takes into consideration the side of the reference window which holds the smaller averaging value which minimizes the over-filtration of reflections from objects. Then Cell-Averaging-(CFAR) is done for estimating the velocity direction. The determined value is multiplied by the threshold scaling factor for determining the detection threshold. The detectability of the radar is determined by this threshold scaling factor. This threshold depends on the Radar Cross Section(RCS) of the target.

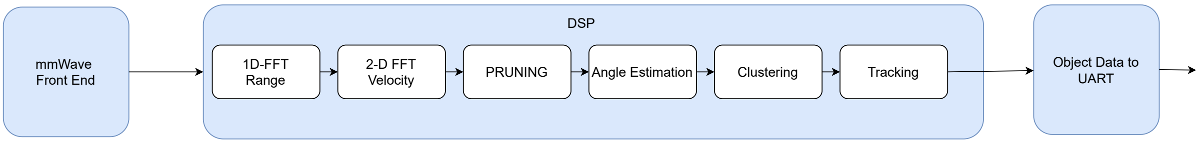


Figure 5.5: Radar processing chain

The Radar Cross Section is the ability of an object to reflect the radar signal towards the radar. The RCS value of the target as to be determined. Depending on the target, the RCS value changes where car and truck have a high RCS of value 5 sq.m and 100 sq.m. The RCS value of bicycles is assumed to be 1.5 sq.m which is higher than pedestrians and lower than cars. Pruning is done on the data to remove clutter from the target. The angle of Arrival of the points are estimated and the position of the points are mapped in X and Y locations in meters. Clustering is done on the point data to reduce the number of reflected points from one target. The clustering is done using the Density Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm. The algorithm groups together point based on a minimal distance and speed between them. On using this algorithm, the mean location of the target object and the X and Y locations are determined.

The tracking of the object is done using the Extended Kalman Filter which has four states as X-position, Y-position, relative velocity along the respective direction. Further Pruning is done to this data to reduce false detection by the radar. The velocity of the object in the front is detected as negative sing which is due to the anti-clockwise movement of the IF signal phasor. The object that has no movement are also detected as the relative velocity between the still object and the bicycle is the velocity of the bicycle. The x-axis has a constraint of 20cm to the right and left which is the width of the bicycle. This helps in reducing the warnings for objects parked on the side and vehicles in other lanes. Time to Collision is calculated for the target object based on the relative velocity and longitudinal range of the object from the host vehicle. This process is done in the mmWave radar DSP chipset and the x location, y location, range, speed, and TTC calculated for each of the target objects are sent to the raspberry pi through UART protocol.

5.4 Threat Assessment

In this section, the threat assessment of a potential thread based on the radar data and the camera data is done. The warning algorithm for the system is shown in Algorithm 1. The

Algorithm 1 Forward Collision Warning System with Adaptive Threshold

```
At time t
Get TTC(t)
Get Head Pose status(t)
Initialize TTC 1 = 5
Initialize TTC 2 = 3.5
Initialize Warning 1 = False
Initialize Warning 2 = False
if TTC(t) <= 5seconds then
    if Head Pose status = Distracted then
        Warning 1 = True
        if TTC(t) < 3.5seconds then
            Warning 2 = True
        end if
    else if Head Pose status = Forward then
        Warning 1 = False
        if TTC(t) < 3.5seconds then
            Warning 2 = True
        end if
    else if Head Pose status = Unknown then
        Warning 1 = True
        if TTC(t) < 3.5seconds then
            Warning 2 = True
        end if
    end if
else
    Warning 1 = False
    Warning 2 = False
end if
```

Warning Algorithm uses the Time to Collision(TTC) metric and the head pose to assess the threat. The minimum reaction time of a rider on a bicycle is 2 seconds. On testing the system, the time to collision (TTC) for the bicycle is set as 5 seconds for distracted riders and 3.5 seconds for alert riders. For an inattentive rider, the system alerts the rider for an object with TTC less than 5 seconds which helps to increase the system's safety. For an alert rider, the system warns only for TTC less than 3.5 seconds, which helps to increase the acceptability of the system. In this system, a Fail-safe algorithm is implemented in case the face detection fails due to external circumstances. This Fail-Safe method consists of a counter that counts down until the face detection work is recovered and then raises an exception. During that time, the threat is evaluated and a warning is issued based only on radar data. The head pose task returns a 0 when the camera sensor fails to capture the head pose of the rider.

Chapter 6

Testing and Results

In this chapter, the testing results of the individual components of the system and system as a whole is briefly explained. The mounting of the system on the bicycle are shown in Fig. 6.1.

(a) Radar sensor mounting

(b) Camera sensor mounting

(c) Camera sensor mounting(angle changed)

Figure 6.1: System

6.1 Camera Sensor Performance

An algorithm for face detection is required to estimate the rider's head pose during cycling in order to assess the situation and deliver the warning. We first compare the frames per second and the complexity of the algorithm. It is shown in Section 6.1. Complexity is the measure of implementation difficulty, training time, and resources needed by the algorithm. It can be concluded that the deep learning methods have high accuracy

Algorithm	Frames per second		Complexity
	(Intel core i7)	Raspberry pi	
HAAR Face Detector	29	15	Low
HOG+SVM Face Detector	24	12	Moderate
CNN Face Detector	13	3	Very High
DNN Face Detector	15	2	Very High

Table 6.1: Comparison of Face Detection Algorithm

in face detection,but the computing complexity of the model are high which makes it less adaptable to embedded devices. Dalal and Triggs method of face detection using HOG and SVM has moderate accuracy and less complexity which suits this research project. The algorithm was implemented in raspberry pi and tested for its accuracy and capability of detection. A dataset captured during testing was used to train the face detector. A new dataset was made to validate the custom-trained face detector and existing dlib face detector. It was inferred that the custom-trained algorithm shows a good detection rate and confidence score compared to the existing face detector. The angle of the camera was changed for the maximum detection rate. The results of testing the face detection algorithm in bicycles are shown in Table 6.2. The detection rate of the custom-trained algorithm is higher compared to the pre-existing dlib face detector algorithm. The custom-trained face detection algorithm was tested in different lighting conditions and it was inferred that the face detection algorithm suffered during poor lighting conditions. Though the dlib pre-existing algorithm has a low detection rate, the sensitivity to change in lighting conditions was low. Hence to increase the performance of the system, both the custom-trained algorithm and dlib face detection algorithm were used to detect the face of the rider.

Algorithm	Detection Rate(%)	Confidence Score
Custom Trained HOG+SVM	99.6	1.0
dlib face detector	48.32	0.9584

Table 6.2: Face Detection Algorithm Performance

6.2 mmWave Radar Sensor Performance

The mmWave radar was tested in different real-time scenarios for determining the performance of the system. The RADAR was mounted in the front of a dutch bicycle and connected to a laptop to visualize the radar data received during testing through MATLAB Graphical User Interface (GUI). A mobile phone was mounted in front to record the objects in front of the cycle for ground truth. The maximum distance up to which the radar as to detect for each $TTC_{Threshold}$ is calculated and shown in Table 6.3. It is assumed that the maximum speed achievable and the average speed of a bicycle as 7m/s and 3m/s. The maximum range for the radar to detect a bicycle is 20m and 14m in the case where the bicycle in front is traveling at average speed and the host bicycle is traveling at maximum speed. The detection range of the system was tested in different

Parameter	TTC=5s	TTC=3.5s
Maximum Speed(m/s)	7	7
Average Speed (m/s)	3	3
Maximum Range (m)	20	14

Table 6.3: Maximum Distance for each Time to collision(TTC)

real-road scenarios for targets such as cars, bicycles, scooters, and humans. The three lane scenarios tested are shown in Fig. 6.2.

(a) Shared Road with bicycle lane

(b) Separate Bicycle lane

(c) Shared Road without bicycle lane

Figure 6.2: Road Scenarios

1. **Shared Road with bicycle lane scenario:** This scenario is shown in Fig. 6.2a where there is a bicycle lane but it is shared with other vehicles. In this scenario, 5 objects were detected and tracked. The average range of detection was 35.1m. Cars, buses, and bicycles were the target of interest in this scenario.

2. **Separate Bicycle Lane scenario:** This scenario is shown in Fig. 6.2b in which there is a separate bicycle lane outside of other vehicle lanes. In this lane, the bicycles are the object of interest. In this scenario,7-objects were detected at an average range of 46.0m.

3. **Shared Road without Bicycle Lane :** This scenario is shown in Fig. 6.2c where there are no separate bicycle lane markings. This scenario is common in inner parts of cities. In this scenario, all types of vehicles are to be considered. The average detection range was 39.6m.

6.3 Total System Performance

The system was tested in different scenarios to evaluate the performance metrics. A confusion matrix was drawn based on the data acquired. This confusion matrix helps in calculating the performance of the system. The warning generated was stored in a csv file and validated. Two cameras were used to validate the warnings generated by the system. One camera was installed in front of the bicycle to record the test ride and the other camera was installed facing the rider to validate the head pose estimated done by the system. The terms used in the confusion matrix are explained below:

1. **True Positive(True Positive (TP)):** True Positive is the total number of times

the warning system warned the rider depending on the obstacle in front and the head pose of the rider.

2. **True Negative(True Negative (TN)):** True Negative is the total number of datapoints where the warnings were not generated when there were no object in the path of the rider The head pose of the driver is also taken into account during this calculation.

3. **False Positive(False Positive (FP)):** The total number of times a warning was provided without an object in the rider's path is termed as False Positive. This might happen as a result of the clutter in the environment.

4. **False Negative(False Negative (FN)):** The number of times the rider was not given a warning despite the presence of an object in the rider's path is termed as False Negative. This figure also includes the total number of times the face was not identified in such scenarios.

The performance metrics used for evaluating the performance are shown in Table 6.4.

Performance Metrics	Description	Formula
Accuracy	Accuracy is defined as the number of correctly categorized data instances divided by the total number of data instances.	$\frac{TP+TN}{Total}$
Precision	Precision is a statistic that measures how many correct positive predictions have been made.	$\frac{TP}{TP+FP}$
Sensitivity	Sensitivity is a metric that indicates how good a model is at recognizing positive occurrences. It is also known as True Positive Rate	$\frac{TP}{TP+FN}$
Specificity	Specificity is defined as the proportion of real negative situations that were predicted as negative by our model. It is also known as True Negative Rate.	$\frac{TN}{FP+TN}$
Misclassification rate	Misclassification rate is defined as the amount of incorrectness in the system. It is the fraction of predictions where the actual value is incorrect.	$\frac{FP+FN}{Total}$

Table 6.4: Performance metric description

6.3.1 Testing Scenario

The Forward collision warning system with distraction detection was tested by the following scenarios. Since there are no standard test procedures for this type of system in International Organization for Standardization (ISO),National Highway Traffic Safety Administration (NHTSA), European New Car Assessment Programme (EURO NCAP). The following scenarios were tested which were adapted from NHTSA and EURO NCAP [50]. The scenarios tested for forward-collision warning system are

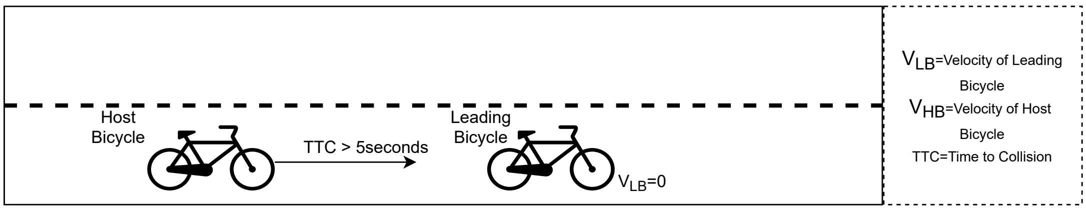
1. **Scenario 1:** The Host Vehicle (HV) and Leading Vehicle (LV) travel in the same direction with TTC >5 seconds. The leading vehicle stops suddenly and the system is checked whether warnings are generated.

2. **Scenario 2:** In this scenario, the host vehicle and leading vehicle travel at a constant speed. The leading vehicle slows down gradually and the system is checked whether warnings are generated or not.

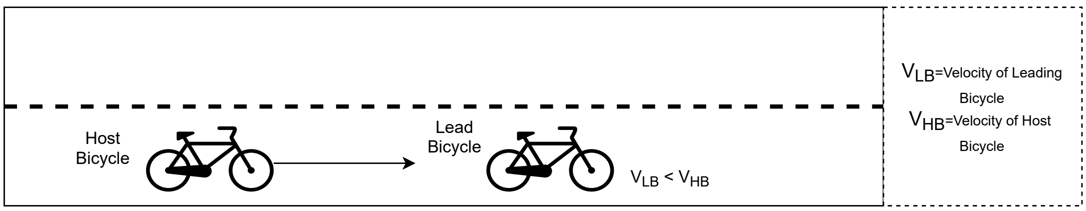
3. **Scenario 3:** The HV and LV travel in the same direction. The leading vehicle travels at a constant speed and the host vehicle accelerates and the systm is checked for warnings.

4. **Scenario 4:** In this scenario, the host vehicle is traveling in one direction and the target vehicle is traveling in the opposite direction. The system is tested to neglect such tracking and detect the object in the path.

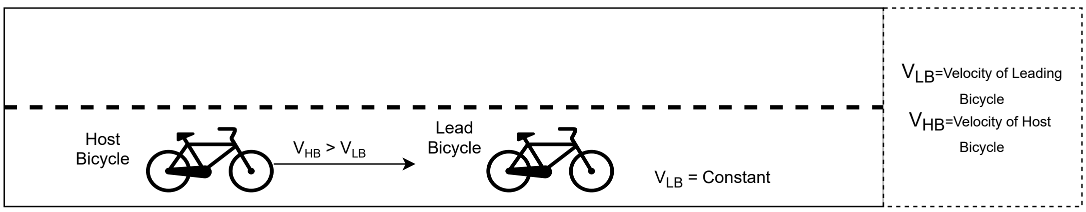
Since the system uses head pose estimation data to refine the warnings, the system was tested in the same scenario including the head-pose data. The scenarios are shown in Fig. 6.3



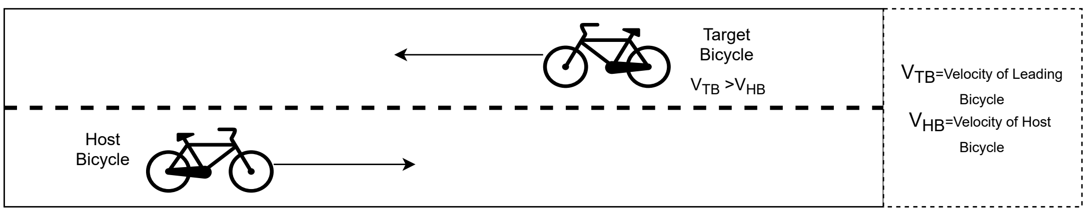
(a) Scenario 1



(b) Scenario 2



(c) Scenario 3



(d) Scenario 4

Figure 6.3: Scenarios

The Table 6.5 shows the detection distance and the time of collision of the object in the testing scenarios. Scenario 4 was tested in real-time and it was observed that the vehicles traveling towards the host vehicle in the adjacent lane doesn't trigger warnings.

The radar configuration to omit detection beyond 20cm in the x-axis helps in neglecting the objects detected in adjacent lanes which doesn't affect the rider. The potential threats are detected and the warnings are delivered to the rider with time to mitigate the collision. These results demonstrate the effectiveness of the proposed system in forward collision situations.

Relative Distance(m)	Time to Collision(s)
8.1	3.6
1.1	1
3.7	1.2

(a) Scenario 1

Relative Distance(m)	Time to Collision(s)
8.1	3.6
3.4	3.4
9.3	3.7

(b) Scenario 2

Relative Distance(m)	Time to Collision(s)
12.2	4.9
9.2	4.3
3	3
1.8	1.2

(c) Scenario 3

Table 6.5: Testing results

6.3.2 Real-road Scenario with different lighting conditions

The system was tested in different light conditions to evaluate the performance of the system. The testing scenario consist of shared road with bicycle lane and without bicycle lane road infratructure.

1. **Day Light condition(12:00pm-3:00pm):** In this condition, the system shows good performance. More extensive testing was done in this condition to evaluate the head pose estimation component of the system. The precision of the system is comparatively high in this condition. The face detection rate in this scenario is moderate. This degradation of face detection rate is due to the change of light conditions while riding the bicycle. The system performance is shown in Table 6.6

2. **Twilight condition(4:30 pm-5:30 pm):** Twilight light condition is the time of the day between daylight and darkness. There was a good performance shown by the system in this condition. The precision and sensitivity of the system degrades in this condition. The face detection rate in this condition is better compared to night time due to the normal lighting conditions without sunlight influence. The system performance is shown in Table 6.6

3. **Night condition(6:00 pm onwards):** The camera suffers during the poor light condition. The system throws more exceptions due to the problem in face detection. Fail-safe algorithm implemented to compensate this limitation helps in generating warnings in this scenario. The sensitivity of the system is low in this condition due to the low face detection rate. The system performance is shown in Table 6.6

Performance Metrics	Day-light Condition	Twilight Condition	Night Condition
Accuracy	0.97	0.96	0.97
Precision	0.93	0.88	0.89
Sensitivity	0.67	0.64	0.57
Specificity	0.995	0.99	0.99
Misclassification rate	0.03	0.04	0.03

Table 6.6: Performance of the system in different lighting conditions

The accuracy of the system is maximum at the day-light condition and this is due to the steady camera's performance.

6.3.3 Overall Performance

The overall performance of the system was tested and validated. The system was tested at Schiedam, the Netherlands. The testing procedure was conducted on November,2021. The testing process was 4 minutes and the testing area is shown in Fig. 6.4. This process was repeated for 9-times. Each testing process covered 1.34km. This area consists of both the shared road with a bicycle lane and without bicycle lane scenarios. The data obtained for every 100m was considered as a dataset for this analysis. A total of 113 datasets was obtained from this experiment. The mean values of the performance metrics are shown in Table 6.7.

Figure 6.4: Testing Area Map

Performance Metrics	Mean	S.D	1-sigma LCL	1-sigma UCL
Accuracy	0.967	0.0321	0.9349	0.9991
Precision	0.8945	0.277	0.6175	1
Sensitivity	0.627	0.433	0.194	1
Specificity	0.9943	0.0139	0.9804	1

Table 6.7: Performance Metrics

Fig. 6.5 shows the distribution of Time to Collision(TTC) for object assessed as a threat. It can be inferred that the system provides 77.54% of warnings before 2 seconds which helps the rider to mitigate the collision. 44.89% of the warnings are given 3.5-5seconds earlier before the collision. These warnings are initial warnings that are generated when the rider is distracted. The second warning where the time to collision is less than 3.5seconds accounts for 55.1% of the total warnings generated. These warnings are crucial regardless of the rider's attention level.

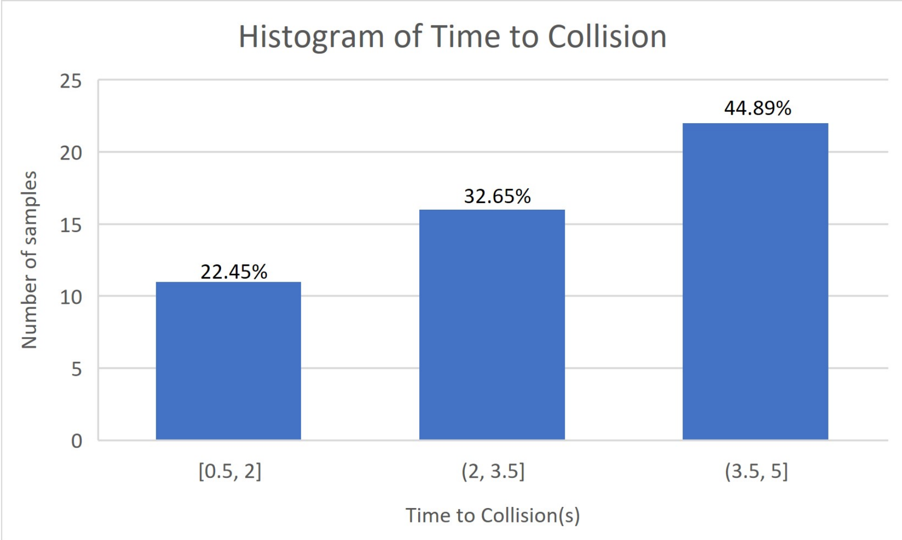


Figure 6.5: Histogram of Time to Collision Data

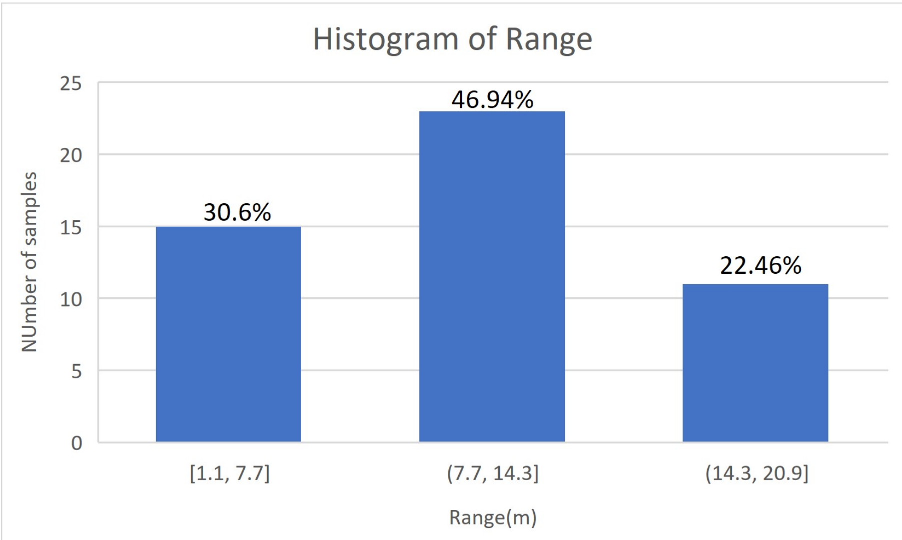
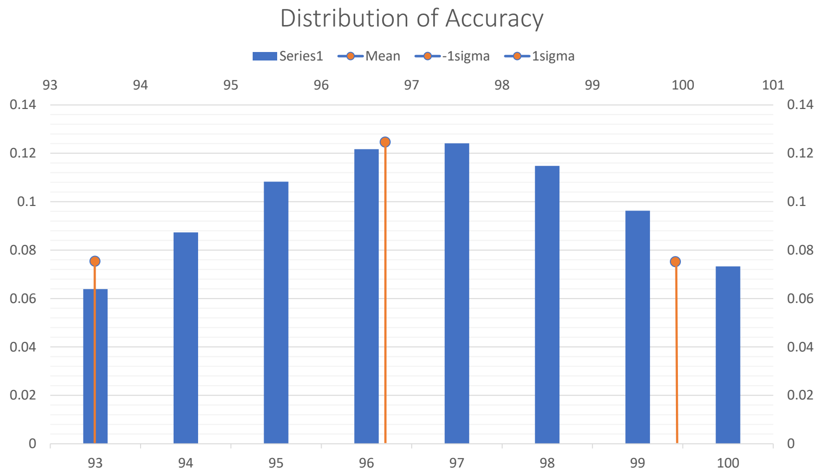


Figure 6.6: Histogram of Range Data

The distribution of range data for each object termed as a threat is shown in Fig. 6.6. It can be inferred that 69.4% of the detection are above 7.7m. 46.94% of the object assessed as a threat are at a range of 7.7-14.3m. About 30.6% of data fall in the minimum range and the minimum range of the object detected as a threat is 1.1m.
The normal distribution of Accuracy is shown in Fig. 6.7. The average accuracy of the system is 0.967(96.7%). One sigma standard deviation was done to the accuracy of the system and 68% of the data falls in this control limits. The accuracy is distributed with a standard deviation of 0.0322(3.2%). The system has high accuracy in the tested area, however the sensitivity of the system degrades due to performance of the camera, configuration of the radar, and the sampling rate. The low lighting conditions and angle of the camera led to low face detection in this experiment. The configuration of radar to detect an object in long-range leads to less resolution in short-range scenarios. Hence



Parameters	Values
Standard Deviation	3.21
-1sigma	93.49
Mean	96.7
1sigma	99.91

68%

Figure 6.7: Analysis of Accuracy of the system

the warnings generated for a vehicle reduces when the vehicle is in this short range. The relative velocity in the shared road scenario is low compared to the dedicated lane scenario. Hence the warnings generated in this scenario is low.

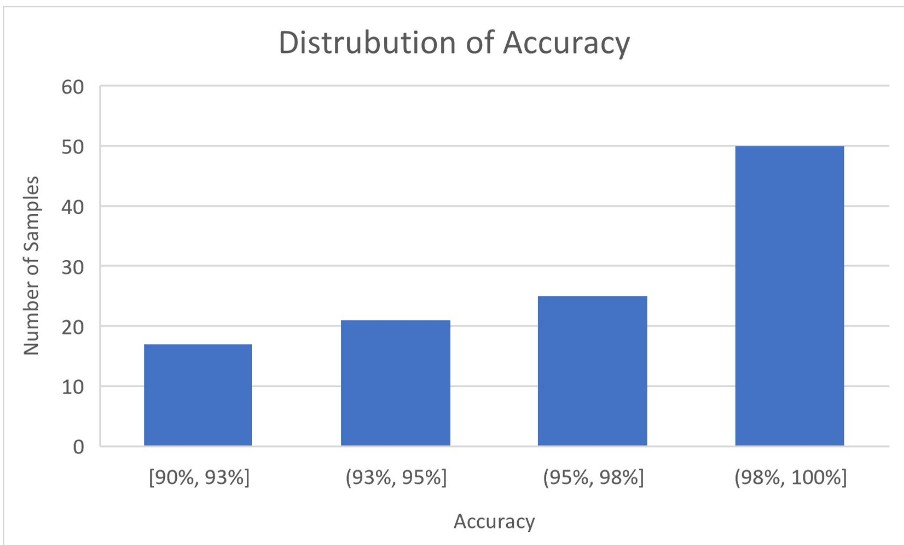


Figure 6.8: Distribution of Accuracy

6.4 Limitation

The system shows good performance in the real-road scenario where the warning are given according to the object in front and head pose of the rider. However, the system has limitations that reduce the sensitivity of the system. The face detection done using a camera has a high dependency on the illumination of the environment. The warnings generated when the face is not detected lead to False Negatives. This restricts the system from being used at all times. This can be solved by using a robust dataset collected during a low light situation and training the model. Moreover, the angle of the camera plays a major role in the detection of the head pose of the rider. The placement of the camera was changed until face detection had maximum true positives. Since this plays a crucial role in the development of this system, a structure can be used to mount the camera to increase the sensitivity of the system. More accurate face detection algorithms such as CNN and DNN can be trained and pruned to work in low-end embedded devices which would help in reducing the false negatives.

The mmWave radar for forward collision detection, has a longer range, high accuracy, and precision but lacks resolution in the near-field. As mentioned in Section 5.2.1, to acquire a high range, the resolution of the detection in near-field reduces. Though the object is detected from a long distance and tracked, the objects in the near field are not detected by the radar and leads to more false negatives. This is due to the configuration of the radar. This limitation can be solved by changing the configuration of the radar which will reduce the maximum range that can be detected. The sampling rate of the system can be altered to overcome the range resolution issue but this can lead to a delayed warning.

Chapter 7

Conclusion and future work

7.1 Conclusion

In this research, a multi-sensor forward-collision warning system was implemented and tested in real-road scenarios. The system uses a sensor fusion of camera and mmWave radar. The mmWave radar is used to detect the object in path of the bicycle. The camera is used to detect the head pose of the rider to detect whether the rider is aware of the obstacle in front. The mmWave radar detects an object with higher accuracy and precision which helps in warning assessment. The TTC and MSD metrics are used to assess the object in front. The mmWave radar helps in providing the range, speed, position of the object in path of the rider, which helps in calculating the Time To Collision(TTC). The camera was used to detect the yaw, pitch, and roll angle of the rider which helps in detecting whether the rider was aware of the object or not. The system shows a mean accuracy of 96.7% in shared road scenarios. Implementation of ADAS system in bicycles improves the safety of the rider. Nearly 70% of the warnings generated by the system was at a range of greater than 7m with Time To Collision(TTC) more than 2.5seconds. The limitation in the system are explained in Section 6.4. The fail-safe algorithm used in the system helps in overcoming the limitations due to the camera in low-light scenarios. The research outcomes motivate more research that has to be done in rider monitoring systems for distraction detection. The threat assessment threshold used help in maintaining the balance between safety and acceptability. More surveys should be done in the threshold values to make the system more acceptable for the rider.

7.2 Future Work

Due to the limitations in the system discussed in Section 6.4. There are a lot of improvements that can be done to the system to increase its performance. A method to detect the type of lane can help in altering the range of the radar which increases accuracy and precision. The angle of the camera affects the detection of the head pose estimation. Determining the correct angle of the camera that provides a relatively higher face detection rate can help in improving the performance of the whole system. A method to detect the illumination in the environment and increase the visibility of the camera during low light situations can help in increasing the accuracy of the system and reducing the situations when face detection fails. Estimating the gaze angle of the rider and the angle of the vehicle can help in reducing the warnings and widen the scenarios in which this system can be used. Due to the sensitivity of the camera sensor on the environment, training the head pose estimation algorithm with dataset from different lighting conditions can help in increasing the face detection rate in low light conditions. The deep learning algorithms for head pose estimation can be implemented on this system after pruning for increasing the accuracy of the camera component. Housing for the radar and the camera should be incorporated to reduce the effect of environmental conditions on the system.

Bibliography

[1] "How ai can help cities to better manage transport." https://www.itu.int/en/myitu/News/2020/11/02/16/11/How-AI-can-help-cities-to-better-manage-transport-Cities-Today. [Online; accessed 15-August-2021].

[2] https://www.hyundai.com/content/dam/hyundai/kr/ko/data/ir-schedule/2019/02/21/hmc-ir-pt-update-19-02-21.pdf. [Online; accessed 25-October-2021].

[3] M. Doudou, A. Bouabdallah, and V. Berge-Cherfaoui, "Driver drowsiness measurement technologies: Current research, market solutions, and challenges," *International Journal of Intelligent Transportation Systems Research*, vol. 18, no. 2, pp. 297–319, 2020.

[4] "how to create distance measuring system." https://www.c-sharpcorner.com/article/how-to-create-distance-measuring-system-using-arduino-uno-r3/. [Online; accessed 15-August-2021].

[5] B. Shahian Jahromi, T. Tulabandhula, and S. Cetin, "Real-time hybrid multi-sensor fusion framework for perception in autonomous vehicles," *Sensors*, vol. 19, no. 20, 2019.

[6] P. A. M. Sandborn, *FMCW Lidar: Scaling to the Chip-Level and Improving Phase-Noise-Limited Performance.* University of California, Berkeley, 2017.

[7] S. Zeadally, J. Guerrero, and J. Contreras, "A tutorial survey on vehicle-to-vehicle communications," *Telecommunication Systems*, vol. 73, no. 3, pp. 469–489, 2020.

[8] "Fundamentals of mmwave radar." https://www.ti.com/lit/wp/spyy005a/spyy005a.pdf?ts=1639925019407&ref_url=https%253A%252F%252Fwww.google.com%252F. [Online; accessed 25-October-2021].

[9] S. Van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, and T. Yu, "scikit-image: image processing in python," *PeerJ*, vol. 2, p. e453, 2014.

[10] G. Amato, F. Falchi, C. Gennaro, and C. Vairo, "A comparison of face verification with facial landmarks and deep features," 04 2018.

[11] "Artificial intelligence and urban development." https://www.europarl.europa.eu/RegData/etudes/STUD/2021/690882/IPOL_STU(2021)690882_EN.pdf. [Online; accessed 16-August-2021].

[12] "Public transport in the netherlands." https://www.emta.com/IMG/pdf/brochure.pdf. [Online; accessed 16-August-2021].

[13] "greenhouse gas emissions in the netherlands 1990- rivm." https://www.rivm.nl/bibliotheek/rapporten/2021-0007.pdf. [Online; accessed 15-August-2021].

[14] "Serious road injuries in the netherlands." https://www.swov.nl/en/facts-figures/factsheet/serious-road-injuries-netherlands. [Online; accessed 16-August-2021].

[15] "Serious road injuries in the netherlands." https://www.swov.nl/en/facts-figures/factsheet/serious-road-injuries-netherlands. [Online; accessed 16-August-2021].

[16] https://ec.europa.eu/transport/road_safety/sites/default/files/ersosynthesis2015-pedestrianscyclists25_en.pdf. [Online; accessed 20-September-2021].

[17] S. T. N. H. V. Crispijn L. van den Brand, Lennard B. Karger and K. Jellema.Neurotrauma, "Bicycle helmets and bicycle-related traumatic brain injury in the netherlands," pp. 201–206, 01 2020.

[18] https://risk.lexisnexis.com/about-us/press-room/press-release/20200618-vehicle-build. [Online; accessed 16-August-2021].

[19] K. Digges and G. Bahouth, "Frequency of injuries in multiple impact crashes," in *Annual Proceedings/Association for the Advancement of Automotive Medicine*, vol. 47, p. 417, Association for the Advancement of Automotive Medicine, 2003.

[20] S. A. Useche, F. Alonso, L. Montoro, and C. Esteban, "Distraction of cyclists: how does it influence their risky behaviors and traffic crashes?," *PeerJ*, vol. 6, 2018.

[21] S. Smaldone, C. Tonde, V. Ananthanarayanan, A. Elgammal, and L. Iftode, "Improving bicycle safety through automated real-time vehicle detection," 09 2010.

[22] S. Raviteja and R. Shanmughasundaram, "Advanced driver assitance system (adas)," in *2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)*, pp. 737–740, 2018.

[23] SamYong Kim, Se-Young Oh, JeongKwan Kang, YoungWoo Ryu, Kwangsoo Kim, Sang-Cheol Park, and KyongHa Park, "Front and rear vehicle detection and tracking in the day and night times using vision and sonar sensor fusion," in *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2173–2178, 2005.

[24] Sen Ma, S. Chen, C. Yang, S. Chu, and J. Pan, "Vision based front and rear vehicle collision warning system," in *2015 Third International Conference on Robot, Vision and Signal Processing (RVSP)*, pp. 22–26, 2015.

[25] A. Kumar and R. Patra, "Driver drowsiness monitoring system using visual behaviour and machine learning," in *2018 IEEE Symposium on Computer Applications Industrial Electronics (ISCAIE)*, pp. 339–344, 2018.

[26] V. Rathod and R. Agrawal, "Camera based driver distraction system using image processing," in *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, pp. 1–6, 2018.

[27] J. Jung, S. Lim, B.-K. Kim, and S. Lee, "Cnn-based driver monitoring using millimeter-wave radar sensor," *IEEE Sensors Letters*, vol. 5, no. 3, pp. 1–4, 2021.

[28] R. Chae, A. Wang, and C. Li, "Fmcw radar driver head motion monitoring based on doppler spectrogram and range-doppler evolution," in *2019 IEEE Topical Conference on Wireless Sensors and Sensor Networks (WiSNet)*, pp. 1–4, 2019.

[29] S. Kumar, V. Shaw, J. Maitra, and R. Karmakar, "Fcw: A forward collision warning system using convolutional neural network," in *2020 International Conference on Electrical and Electronics Engineering (ICE3)*, pp. 1–5, 2020.

[30] H. Kim, Y. Lee, T. Woo, and H. Kim, "Integration of vehicle and lane detection for forward collision warning system," in *2016 IEEE 6th International Conference on Consumer Electronics - Berlin (ICCE-Berlin)*, pp. 5–8, 2016.

[31] W. C. Phoon and P. Y. Lau, "Real-time forward collision alert system using raspberry pi," in *2019 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*, pp. 1–2, 2019.

[32] Q. Lim, Y. He, and U.-X. Tan, "Real-time forward collision warning system using nested kalman filter for monocular camera," in *2018 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pp. 868–873, 2018.

[33] Y. Lu, Y. Yuan, and Q. Wang, "Forward vehicle collision warning based on quick camera calibration," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2586–2590, 2018.

[34] Y.-K. Lai, Y.-H. Chou, and T. Schumann, "Vehicle detection for forward collision warning system based on a cascade classifier using adaboost algorithm," in *2017 IEEE 7th International Conference on Consumer Electronics - Berlin (ICCE-Berlin)*, pp. 47–48, 2017.

[35] S. Hong and D. Park, "Lightweight collaboration of detecting and tracking algorithm in low-power embedded systems for forward collision warning," in *2021 Twelfth International Conference on Ubiquitous and Future Networks (ICUFN)*, pp. 159–162, 2021.

[36] L. Eriksson and B. As, "A high performance automotive radar for automatic aicc," in *Proceedings International Radar Conference*, pp. 380–385, 1995.

[37] P. Ganci, S. Potts, and F. Okurowski, "A forward looking automotive radar sensor," in *Proceedings of the Intelligent Vehicles '95. Symposium*, pp. 321–325, 1995.

[38] T. F. Stevens, "A lidar based semi-autonomous collision avoidance system and the development of a hardware-in-the-loop simulator to aid in algorithm development and human studies," 2015.

[39] C. Englund, H. Clasen, T. Bui, D. Lindström, and J. Andersson, "Radar system for bicycle -a new measure for safety," 10 2019.

[40] Y. Wei, H. Meng, H. Zhang, and X. Wang, "Vehicle frontal collision warning system based on improved target tracking and threat assessment," pp. 167 – 172, 09 2007.

[41] S. Suzuki, P. Raksincharoensak, I. Shimizu, M. Nagai, and R. Adomat, "Sensor fusion-based pedestrian collision warning system with crosswalk detection," in *2010 IEEE Intelligent Vehicles Symposium*, pp. 355–360, 2010.

[42] X. Yang, C. Lu, and W. Pan, "A zigbee-based highway vehicles prevent collision warning system research," in *2011 Fourth International Symposium on Computational Intelligence and Design*, vol. 1, pp. 270–273, 2011.

[43] T. Yang, Y. Zhang, J. Tan, and T. Z. Qiu, "Research on forward collision warning system based on connected vehicle v2v communication," in *2019 5th International Conference on Transportation Information and Safety (ICTIS)*, pp. 1174–1181, 2019.

[44] A. Hosny, M. Yousef, W. Gamil, M. Adel, H. Mostafa, and M. S. Darweesh, "Demonstration of forward collision avoidance algorithm based on v2v communication," in *2019 8th International Conference on Modern Circuits and Systems Technologies (MOCAST)*, pp. 1–4, 2019.

[45] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," vol. 1, pp. I–I, 2001.

[46] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," vol. 23, pp. 1499–1503, 2016.

[47] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 1, pp. 886–893 vol. 1, 2005.

[48] `https://ibug.doc.ic.ac.uk/resources/300-W/`. [Online; accessed 15-November-2021].

[49] `https://aifi.isr.uc.pt/HeadPoseEstimation.html`. [Online; accessed 15-November-2021].

[50] "Autonomous emergency braking collision mitigation testing." `https://www.zenmicrosystems.co.in/autonomous-emergency-braking-collision-mitigation-testing/`.

Appendix

Real-time testing in different lighting conditions

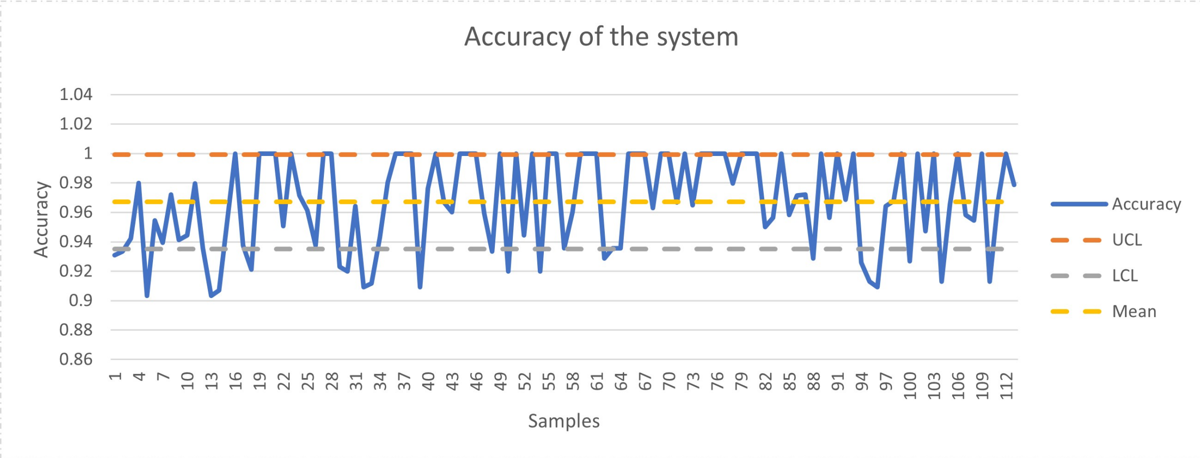
The different lighting conditions in the overall performance testing are shown in Section 7.2. It is clear that the light condition is better in twilight and daylight which helps in increasing the face detection rate.

(a) Daylight (b) Twilight (c) Night

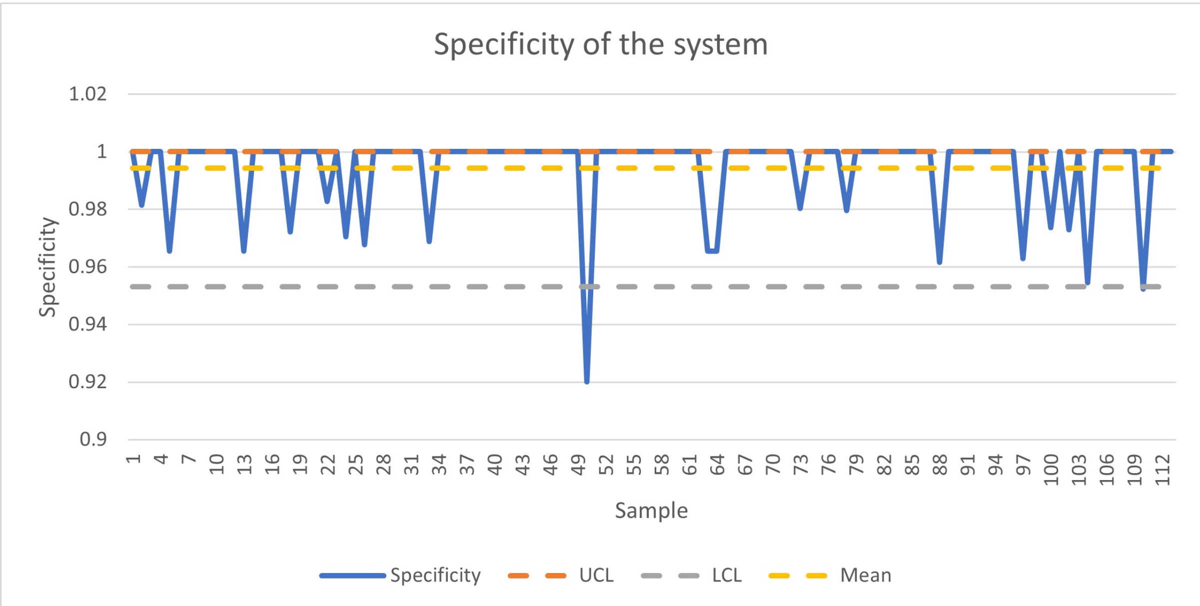
Real-time testing on different lighting conditions

Total System Performance

The accuracy of the system was analysed statistically ,the control chart of the accuracy calculated for the system is shown in Section 7.2. The specificity parameter gathered during the tested scenario is statistically analyzed. The average specificity of the system is 0.994 with standard deviation of 0.0139.



Statistical Analysis of Accuracy of the system



Statistical Analysis of Specificity of the system