Real-Time Driving Monitoring System on a Single-Board Computer Utilizing Deep Neural Networks Integrated with MiDAS

by

SAZEED KAISAR
18301181
REDUAN MASUD JIM
18301240
RIAN ISLAM
18301244

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science

Department of Computer Science and Engineering
Brac University
September 2023

Declaration

It is hereby declared that

1. The thesis submitted is my/our own original work while completing degree at Brac University.

2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.

3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.

4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

Kaisar

SAZEED KAISAR
18301181

Reduan

REDUAN MASUD JIM
18301240

Rian

RIAN ISLAM
18301244

Approval

The thesis/project titled “Real-Time Driving Monitoring System on a Single-Board Computer Utilizing Deep Neural Networks Integrated with MiDAS” submitted by

1. SAZEED KAISAR (18301181)
2. REDUAN MASUD JIM (18301240)
3. RIAN ISLAM (18301244)

Of SUMMER , 2023 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on 22 September,2023.

Examining Committee:

Supervisor: (Member)

Md. Khalilur Rhaman

Dr. Md. Khalilur Rahman
Professor
School of Data and Sciences
Brac University

Program Coordinator: (Member)

Dr. Md. Golam Rabiul Alam
Professor
School of Data and Sciences
Brac University

Head of Department: (Chair)

Dr. Sadia Hamid Kazi
Associate Professor
Department of Computer Science and Engineering
Brac University

Ethics Statement

The development and implementation of the Real-Time Driving Monitoring System on a Single-Board Computer, incorporating Deep Neural Networks integrated with MiDAS, is undertaken with unwavering commitment to the highest ethical standards. This research endeavors to adhere to a set of guiding principles that not only meet legal requirements but also align with the directives established by relevant professional bodies.

We hold the privacy, safety, and rights of all individuals involved in this research in the highest regard. To this end, we have sought and obtained all necessary permissions and consents for the collection and use of any data, images, or other information gathered during the course of this research endeavor. Furthermore, we ensure that all data is handled with the utmost care, securely stored, and exclusively utilized for the explicit purposes of this study.

Our responsibility extends to minimizing any potential negative impacts of this research on drivers, passengers, and road users. The system is designed to enhance safety and awareness without compromising privacy or infringing on individual liberties. We strive to ensure that the benefits derived from this research are shared widely with the public, contributing to the advancement of road safety measures.

Additionally, we are committed to transparency and accountability in the deployment of artificial intelligence and neural network technologies. The models and algorithms utilized in this system are rigorously trained and validated to ensure accuracy and reliability. We acknowledge the importance of avoiding biases and ensuring fairness in all aspects of system operation.

In summary, this research upholds the principles of ethical research conduct, respecting the rights and well-being of all individuals involved. It is our earnest intention that this Real-Time Driving Monitoring System ultimately contributes to the broader public interest by enhancing road safety and advancing the responsible use of advanced technologies in driving environments.

Abstract

In recent years, road safety has emerged as a critical concern due to the increasing number of accidents attributed to driver negligence and fatigue. This thesis addresses this pressing issue by proposing a Real-Time Driving Monitoring System designed for deployment on a single-board computer. The system employs a combination of cutting-edge technologies to comprehensively assess driver safety during operation.

The system's core objective is to discern whether the driver is operating the vehicle in a safe manner. To achieve this, three distinct input streams from specialized cameras are utilized. The first input stream leverages YOLOPv2, a state-of-the-art object detection model, to accurately detect road lanes and determine if the vehicle remains within the designated lane. This real-time feedback is crucial for preempting potential lane departure incidents.

The second input stream employs Monocular Depth Estimation with MiDAS, a robust and efficient technique for gauging the distance of objects in close proximity to the vehicle. By aggregating depth measurements and calculating a mean depth value, the system establishes an empirical threshold. Instances where the mean depth falls below this threshold are indicative of potential collision risks, prompting the system to identify the driver as unsafe.

Furthermore, the third input stream utilizes the front-facing camera to monitor driver behavior and detect signs of drowsiness. Through a combination of facial feature analysis and eye tracking, the system can accurately determine if the driver exhibits signs of fatigue or inattentiveness. Should the driver display drowsiness for a duration surpassing the specified threshold, an alert is triggered, thereby mitigating the risks associated with driver fatigue.

In the event that any of the aforementioned conditions persist for a predetermined duration, the system activates an alert protocol. This protocol includes the illumination of LED indicators and the sounding of a buzzer, providing immediate feedback to the driver and drawing attention to the potential safety hazard.

By combining these advanced technologies in a single-board computer-based system, this thesis presents a comprehensive approach to real-time driving monitoring. The integration of YOLOPv2 and MiDAS with deep neural networks ensures accurate and timely detection of potential safety risks, thereby contributing significantly to the enhancement of road safety standards. .

Keywords: Machine Learning; Domain Generalization; Lane Detection; Decision tree ; MIDAS ; YOLOPv2 ;

Dedication (Optional)

This research is dedicated to the visionaries and innovators whose tireless efforts have propelled the field of object detection to new heights. We extend our dedication to the future generations of researchers and developers who will continue to push the boundaries of what is achievable in driving safety systems. Furthermore, this research is a tribute to the driving safety-related foundations, recognizing their pivotal role in fostering advancements that safeguard lives on the road. The progress made in object detection holds paramount importance in identifying and averting potential threats to road safety, mitigating accidents, and ultimately contributing to the preservation and protection of our highways for generations to come. Through this endeavor, we aspire to forge new frontiers and catalyze the growth of our nation.

Acknowledgement

We would like to thank our esteemed thesis adviser Dr. Md. Khalilur Rahman Sir, Sayantan Roy Arko and Shaownak Shahriar for their invaluable advice, support, and encouragement during this research attempt. We'd also want to thank the members of our thesis committee for their insightful comments and ideas. We are really thankful to The School of Data Science at Brac University for their aid and support in completing this thesis. Finally, we would like to thank our family and friends for their unwavering support and understanding.
In addition, we would like to thank our faculties from the School of Data Science for their assistance and for creating a vibrant and collaborative environment for learning and research throughout our academic journey. Finally, we would want to express our heartfelt gratitude to everyone who helped make this thesis feasible. Their support, as well as encouragement, were important in completing this task.

Table of Contents

List of Figures

Nomenclature

The next list describes several symbols & abbreviation that will be later used within the body of the document

	X
ϵ	Epsilon
μ	Mu
ν	Nu
υ	Upsilon
ϖ	VarPi
CNN	Convolutional Neural Networks
FN	False Negatives
FP	False Positives
MSE	mean squared error
PSNR	Peak Signal-to-Noise Ratio
RoI	Region of Interest
SSIM	structural similarity index
TN	True Negatives
TP	True Positives

Chapter 1

Introduction

1.1 Motivation

In recent years, the escalating concern over road safety has prompted a paradigm shift in automotive technology. The surge in vehicular accidents attributed to driver negligence, distraction, and fatigue has necessitated the development of innovative monitoring systems capable of providing real-time feedback to drivers. This shift towards proactive safety measures is particularly pertinent given the rapid advancement of autonomous driving technology, which demands a seamless transition from manual to automated control.

Figure 1.1: Driver Monitor System research nowadays

The World Health Organization (WHO) estimates that over 1.3 million lives are lost annually due to road traffic accidents, making it one of the leading causes of mortality worldwide [3]. Furthermore, countless individuals suffer life-altering injuries, exerting an immense economic burden on healthcare systems and societies at large. These grim statistics underscore the critical need for comprehensive driver monitoring systems that can significantly reduce the incidence of accidents.
Traditionally, driver safety has been addressed through legislation, education, and the incorporation of passive safety features within vehicles. While these measures have undoubtedly contributed to mitigating risks, they do not sufficiently address the issue of active, real-time monitoring of driver behavior. Consequently, there exists a substantial gap in the existing safety paradigm.

1.2 Research problem

Conventional driver monitoring systems primarily rely on rudimentary sensors and heuristic-based algorithms. These systems are often incapable of discerning nuanced indicators of driver fatigue or distraction, leading to false positives or, worse, overlooking genuine instances of unsafe driving behavior. Moreover, the efficacy of these systems is contingent on ideal lighting conditions, making them less reliable during adverse weather or nighttime driving [6].

To bridge this gap, this research endeavors to integrate state-of-the-art technologies into a unified Real-Time Driving Monitoring System. By harnessing the power of Deep Neural Networks (DNNs) in conjunction with MiDAS (Monocular Depth Estimation), this system aims to revolutionize the landscape of driver safety.

Deep Neural Networks (DNNs) have demonstrated unprecedented capabilities in various computer vision tasks, including object detection, facial recognition, and image segmentation [4]. Their ability to discern intricate patterns in visual data makes them an ideal candidate for driver monitoring applications. By leveraging DNNs, the system can perform real-time analysis of visual input from multiple cameras, enabling rapid and accurate identification of potential safety risks.
Monocular Depth Estimation with MiDAS augments the system's perceptual capabilities by providing a reliable means of gauging distances to objects in close proximity to the vehicle [7]. This spatial awareness is pivotal in assessing potential collision risks, especially in scenarios where conventional sensors may fall short. By calculating the mean depth from multiple measurements, the system establishes an empirical threshold, offering a quantifiable metric for identifying unsafe driving conditions.

The genesis of this research lies in the urgent need for a sophisticated, real-time driving monitoring system that can transcend the limitations of existing approaches. By amalgamating cutting-edge technologies, we seek to create a comprehensive solution that not only addresses the challenges posed by driver negligence and fatigue but also paves the way for a safer and more secure future of transportation.

Through this amalgamation of DNNs, MiDAS, and a single-board computer, we aim to harness the full potential of modern computational capabilities to proactively enhance driver safety. This research aspires to contribute significantly to the ongoing discourse on road safety, with the ultimate goal of saving lives and reducing the societal impact of road traffic accidents.

1.3 Research Objectives

In this research endeavor, our primary aim is to construct, implement, and appraise a Real-Time Driving Monitoring System, seamlessly integrated into a single-board computer. This system will harness the power of Deep Neural Networks (DNNs) and Monocular Depth Estimation with MiDAS, representing a holistic approach to driver safety. Our research objectives can be elucidated as follows:

1. Development of a Comprehensive Driver Safety Assessment System:
a. We will incorporate YOLOPv2, a state-of-the-art object detection algorithm, for real-time road lane detection. This will ensure that the vehicle remains within the designated lanes, even under varying environmental conditions.
b.Additionally, we'll integrate Monocular Depth Estimation with MiDAS to ascertain distances of objects proximate to the vehicle. By calculating mean depth values, we aim to establish a threshold for identifying potential collision risks.

2.Integration of Deep Neural Networks for Real-Time Analysis: a.Deep Neural Networks will be deployed to process input from multiple cameras concurrently. This will enable real-time analysis of visual data streams, encompassing lane deviation detection, depth assessment, and driver behavior monitoring.
b.Object detection capabilities will be leveraged to identify road lanes from the YOLOPv2 camera feed, facilitating immediate corrective feedback to ensure the vehicle remains on course.
c.The utilization of DNNs for facial feature analysis and eye tracking will allow for real-time detection of driver drowsiness, promptly triggering alerts when necessary.

3. Real-Time Alerting Mechanism:
a.An alerting system, comprising LED indicators and a buzzer, will be integrated to provide immediate feedback to the driver. This system will be designed to activate alerts in response to identified instances of unsafe driving behavior.
b.We will develop a threshold-based logic to determine the duration and severity of unsafe driving behavior, thereby fine-tuning the alerting mechanism to respond appropriately.

Chapter 2

Related Work

The pursuit of driver safety has spurred extensive research in the development of advanced monitoring systems. Among the pivotal advancements in this domain are the application of Convolutional Neural Networks (CNNs) for real-time object detection and classification. Redmon et al. (2018) introduced the groundbreaking You Only Look Once (YOLO) models, revolutionizing object recognition by enabling rapid, one-pass processing. YOLO's efficiency in resource utilization and speed has made it instrumental in real-time applications, including the detection of road lanes. YOLOPv2, an incremental improvement introduced by the same authors, refines the balance between accuracy and speed, rendering it a pivotal tool for detecting road lanes promptly and accurately [10] [12].

Monocular Depth Estimation, a critical facet of our system, has emerged as a central theme in recent research endeavors. Zhou et al. (2017) introduced Monocular Depth Estimation with Pre-trained Scene-specific Networks (MiDAS), a state-of-the-art method for inferring depth from monocular images. The approach leverages pre-trained networks on scene-specific data, resulting in impressive accuracy in estimating distances to objects in proximity to the vehicle. MiDAS plays a fundamental role in our system by providing essential spatial awareness, which is crucial for evaluating potential collision risks.[5] The depth estimation equation utilized in MiDAS can be represented as follows:

$$D = f.\frac{B}{D} \tag{2.1}$$

where: D = Estimated Depth,
f = Focal Length of the Camera,
B = Baseline (distance between the two cameras),
d = Disparity.

Driver drowsiness detection has been an enduring area of interest, with a particular emphasis on facial cues and eye movement patterns as indicative markers. Xu et al. (2016) proposed a comprehensive framework employing a combination of facial landmark detection, eye state analysis, and head pose estimation. Their approach, akin to ours, endeavors to swiftly discern drowsiness indicators, ensuring timely alerts to the driver. By utilizing multiple facial features, their system demonstrates commendable accuracy in detecting signs of driver fatigue, a pivotal aspect of our integrated monitoring system [9].

While the individual components of our system have been explored in isolation, the amalgamation of YOLOPv2, MiDAS, and Deep Neural Networks (DNNs) on a single-board computer signifies a novel advancement in this domain. This integration harnesses the complementary strengths of each technology, resulting in a comprehensive real-time monitoring system capable of accurately evaluating driver safety across multiple dimensions.

2.1 Comparative Analysis of Approaches

In a comparative analysis, it becomes evident that our integrated system represents a convergence of cutting-edge technologies, leveraging the strengths of YOLOPv2, MiDAS, and DNNs. This comprehensive approach stands as a notable departure from previous systems that have predominantly relied on conventional sensors and heuristic algorithms. Previous efforts, while commendable, often face limitations in accuracy, real-time processing capabilities, and adaptability to diverse environmental conditions. For instance, traditional driver monitoring systems relying on rudimentary sensors can struggle with nuanced indicators of driver distraction or drowsiness, leading to either false positives or overlooking genuine instances of unsafe driving behavior. In contrast, the incorporation of YOLOPv2 ensures real-time and precise lane detection, providing immediate feedback to the driver in scenarios of potential lane deviation. Moreover, the integration of MiDAS offers a significant advancement in spatial awareness, enabling the system to gauge distances to objects near the vehicle with remarkable accuracy. This capability is crucial in assessing potential collision risks, particularly in situations where conventional sensors may fall short. The utilization of DNNs for real-time analysis of multi-camera inputs represents a paradigm shift in driver monitoring. This approach empowers the system to perform intricate pattern recognition and analysis, facilitating the detection of road lanes and assessment of driver behavior. By leveraging the strengths of DNNs, our system is adept at processing complex visual data streams from various angles, resulting in a more comprehensive evaluation of driver safety.

2.2 Future Implications

The integration of YOLOPv2, MiDAS, and DNNs in our real-time driving monitoring system holds profound implications for the future of road safety. This integrated approach not only addresses the limitations of existing systems but also sets a new standard for the industry. By combining the capabilities of advanced neural networks and depth estimation techniques, we open the door to a new era of driver safety. Furthermore, the adaptability and scalability of our system position it as a potential cornerstone in the development of autonomous vehicles. The technology and methodologies employed in our research can serve as a foundation for future innovations in automated driving, ensuring that safety remains a paramount consideration. In conclusion, our research represents a significant step forward in the pursuit of enhanced driver safety. By integrating state-of-the-art technologies, we aim to contribute to a future where road accidents and their associated human and economic costs are drastically reduced. The successful implementation of our integrated monitoring system showcases the potential for advanced technologies to make a tangible difference in road safety.

2.3 Literature Review

As we know that the world population is currently increasing, so are motorized vehicles. If we think only about countries where owning a private car isn't much affordable, there are more than 4.4 million motorized registered vehicles up to 2021. In between these, more than 370k are private passenger cars. If we look back in 2003, there were only 303k motorized vehicles. Up to 2022 there are about 5403590 vehicles registered in Bangladesh Road Transport Authority (BRTA) [9]. Here is a chart that shows how alarming the situation is-

Type of Vehicles	2010	2011	2012	2013	March 2014
Bus	27,778	29,539	30,978	32,085	32,391
Jeep	32,286	34,420	35,989	37,303	37,658
Microbus	66,379	70,430	73,474	76,011	76,957
Minibus	25,644	25,920	26,169	26,317	26,357
Taxicab	44,380	44,455	44,627	44,678	44,685
Truck	82,871	90,198	94,533	99,662	102,144
Auto-rickshaw	126,763	147,186	170,731	186,428	191,376

Figure 2.1: NUMBER OF REGISTERED MOTOR VEHICLES IN BANGLADESH (YEARWISE)[8]

When in a lower middle economic country where GDP per capita $2503 facing an increment of car in this rate, then how alarming the situation is in countries like the USA where GDP per capita is $69287.54 **10**. As vehicles are increasing, the accidents are also increasing. To reduce accidents there are many researches and proposals that are working on driving monitoring. So mostly there are two categories - i) Drivers behavior monitoring ii) Driving or car monitoring According to [11] we need to monitor 3 major things that may be the reason for 90% of car accidents. Those are- • Distraction: Distraction from inside or outside of cars. 2. Fatigue: Unwell feeling of drivers (Drowsiness) 3. Aggressive Driving: Over speeding, hard-breaking traffic laws.

According to [1] we can predict drivers behavior with different parameters like facial expression, steering wheel controlling etc. We can predict major and minor accidents. Using visual sensors we can get some data of the driver like eye blinking frequently, percentage of eye closure, reflen towards particular actions. Also there can be some non visual measurements like EEG, ECG, EOG, PPG monitoring. Though these features seem complicated, for steering wheel observation we can mention the steering wheel standard deviation. For visual purposes we can use NIR illumination. Also according to PFL in Switzerland, we can detect distraction [13]. Another research from D'Orazio with some classical algorithms for detecting eyes can generate facial recognition images with a camera. Also for using IR sensors we had some problems in the past. Like the rays from IR creates problems when there are glasses and red eye problems. Using IR illumination, we can produce different lighting and solve the problem. For producing a 3d image of the driver's face we can place two cameras like in the picture.

Figure 2.2: Two cameras at the front

Here a neural classification can be used to categorize the eye from the image. Also in this case the frequency of eye blinking is less important than the duration of eye closing. Here we can use a multivariate Gaussian mixture model characterizing the two parameters of eye blinking. After giving proper data to this model we can characterize the normal value to compare. In this case, eye detection is very convenient. It works in three The two cameras placed on the car dashboard[14]steps. Firstly iris recognition using transform. Secondly, searching the symmetrical region from the limited area of the image. Lastly using neural classification, an eye validation algorithm is used. Here the writers also put some test results to compare. A method of detection is [2].

In[15] we can see , we can recognize different facial expressions from the driver to detect their emotional status. Taking the image of the driver, the system can categorize their emotional status using a deep convolutional neural network. In another research[16] the researchers used two cameras. One is to check the driver's status and sight, another one is to check the lane condition. With these two data, they observe the front position of the car(The heading direction). Here they have calculated a coefficient to determine the relation between the driver's sight detection and the lane path. From this, a pattern can be generated [16]. In [17] recognized local binary patterns from a video using facial and eye recognition. Then using the LBP operation they recognized the status of the eye. Here the eye region is divided into small blocks. Then calculate the feature vector for each block. Then combining all we can get a feature vector for the whole eye area. Now from this using AdaBoost [18] we can determine weak and strong classified to determine the useful data from it. After collecting data with a training data set, a system was built to detect drowsy driving [17].
In[19] we can see the LSTM model from driving behavior classification. Different sensors were used to get different data. Firstly acceleration is measured along three axes. Then Yaw , roll, and pitch angle are measured. Also to determine vehicle speed GPS sensor is used . A camera is used to calculate the distance from ahed vehicle and number of vehicles in front [19]. Using the raw data from the sensor's time stamp, distance ahed vehicles in the current lane, a number of vehicles ahead, etc.

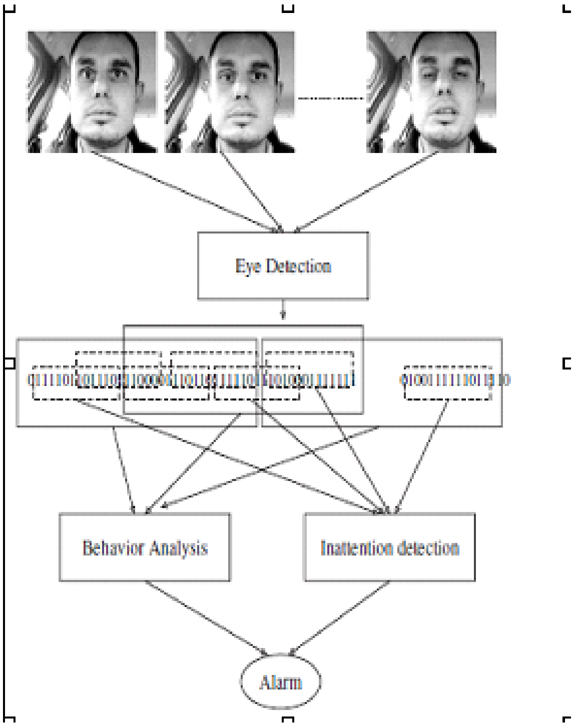

Figure 2.3: Two cameras at the front

can be kept. The most interesting part from this research is all the data is collected by the smartphone sensors. In [20] the researchers had worked on a fuzzy-based system for safe driving. They presented FDMS1 and FDMS2, these two systems. Both of these two systems recognize the driver's situational awareness. In FDMS1 they worked on outside temperature, noise, heart responses, and respiratory rate. These all things worked as input in FDMS2. In this system sometimes the output might lead in the wrong direction. Like though the situation is normal, it can detect unawareness from the driver. In[21] the researchers talked about Mag track. Here using magnetic accessories they have worked with the motion of hand and head. They have used a sensing algorithm , with that using a single manometer in a hand watch it takes data. With this data the system can verify the driving safety. With this system the accuracy is very high [21]. For driving monitoring another important data is to determine the traffic situation in the surroundings.
In[22] the system reflects laser light and that reflects from other obstacles and with that distance can be measured using the Kalman Filter. Here the researchers have calculated some equations for speed and distance. In driving monitoring, in some cases the vehicles are needed to be tracked. In this case there is a work in [23]. In this research they have used LPC2148 . It takes input from temperature, alcohol , and GPS sensors. And using that, it gives output through Buzzer and motors with an external monitor, it can show output status from the system [23]. In another research [24] the front car is recognized by the tail light. It is specially for night time car following. It puts a box and identifies the front cars in it. It uses a camera in front . Though there are some drawbacks in this system. Firstly as it follows the tail light so it doesn't work in daylight. Secondly as it uses research from about 30 years ago some new technologies can be implemented to remove those drawbacks [24].

Nowadays some well known car companies are developing advanced drive assistance system.ADAS in short. ADAS actually alerts about the upcoming danger [25]. One of the features to be implemented into ADAS is the traffic sign recognition system. In [26] we came to know about how we can use deep convolutions. Neural network for Real Time Traffic sign Recognition System. In this system lightweight algorithms can be used for faster recognition. As it will give real time results,parallel programming can be used. In this system CNN model can be used for image classification. In this image is classified into 3 parts Hue, saturation and value. As traffic signs have categorized color for each kind of sign throughout the world, the recognition won't be that much complicated. So using this feature in the driving monitoring system drivers can be warned. Also if someone doesn't abide by any sign that can be kept in record. There is another parameter for judging if the driver is safely driving or not that is lane centering. Keeping the car in the center of lane and not changing frequently is an indication to safe driving. In[27] for lane centering for various image processing, for various image processing RANSAC can be used. From the front camera the lane can't be identified exactly. For this a bird's eye image should be needed here. The RANSAC is needed to be applied in HSV on hue saturation value. The output from

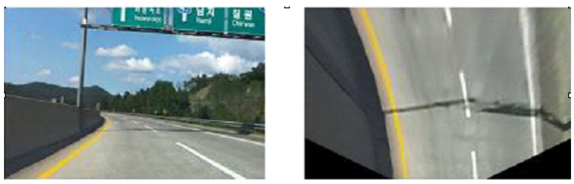

Figure 2.4: Input image(Left) and (b) bird's-eye view(Right)

this system is quite good. Using the RANSAC with Kalman filter, it gives correct markings on 86.39% and gives incorrect 12.71%. It misses 0.78% [27].

Chapter 3

Methodology

The Real-Time Driving Monitoring System is a sophisticated integration of hardware and software components designed to operate on a single-board computer. It incorporates three distinct cameras, each serving a unique purpose in assessing driver safety. The architecture, illustrated in Figure 1, encompasses a YOLOPv2-enabled lane detection camera, a Monocular Depth Estimation with MiDAS camera for object proximity assessment, and a front-facing camera for monitoring driver behavior.

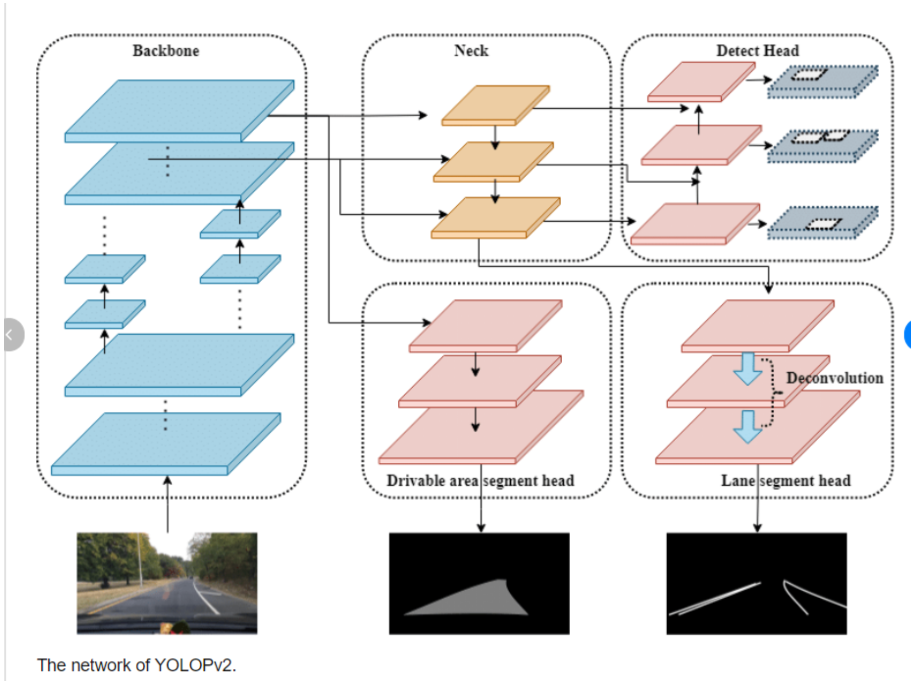

Figure 3.1: Top Level Overview of Proposed Model

3.0.1 YOLOPv2 Lane Detection

:
Central to the system's functionality is the YOLOPv2 model, a cutting-edge object detection algorithm. Trained on an extensive dataset, this model excels at real-time lane detection, capable of accurately identifying lane boundaries across diverse

road environments. Its strength lies in its adaptability to varying lighting and weather conditions, a crucial attribute for robust performance in real-world driving scenarios. Operating in real-time, YOLOPv2 processes images from the dedicated lane detection camera, swiftly generating bounding boxes around detected lanes.

3.0.2 Monocular Depth Estimation with MiDAS:

Monocular Depth Estimation is facilitated by the MiDAS framework, an advanced methodology for inferring depth from monocular images. The framework leverages a pre-trained neural network, fine-tuned on scene-specific data, to achieve remarkable accuracy in estimating distances to objects proximate to the vehicle. The depth estimation equation, represented as = D=f d B , encapsulates the critical relationship between estimated depth (D), focal length of the camera (f), baseline distance (B), and disparity (d). This depth information serves as a pivotal dimension in evaluating potential collision risks.

3.0.3 Driver Drowsiness Detection:

A front-facing camera is dedicated to monitoring driver behavior, with a specific emphasis on detecting signs of drowsiness. The system employs a hybrid deep neural network, trained on a diverse dataset of facial expressions and head orientations. This network performs comprehensive facial feature analysis, eye state detection, and head pose estimation. Through continuous monitoring, the algorithm identifies indicators of driver fatigue. If drowsiness persists for a duration surpassing the predefined threshold, an alert is promptly issued.

3.0.4 Data Fusion and Decision Logic:

To form a comprehensive safety assessment, the output from each camera undergoes fusion. The lane detection output verifies if the vehicle remains within designated lanes, offering immediate corrective feedback. Simultaneously, depth information is utilized to gauge distances of objects near the vehicle. The aggregation of multiple depth measurements enables the computation of mean depth. If this mean value falls below the predefined threshold, the system deems the driver to be in an unsafe situation, warranting an alert.

In parallel, the drowsiness detection algorithm continually evaluates the driver's level of alertness. If drowsiness persists beyond the stipulated threshold, a prompt alert is triggered, ensuring the driver is promptly informed of the identified safety concern.

3.0.5 Alerting Mechanism:

The system employs a dual-alert mechanism to provide immediate feedback to the driver. LED indicators, strategically positioned within the driver's line of sight, illuminate to convey visual alerts. Additionally, a buzzer generates an audible alert, offering an additional layer of notification. This multi-modal alerting system ensures that the driver is promptly informed of any identified safety concern.

3.0.6 Experimental Setup:

A rigorous evaluation of the Real-Time Driving Monitoring System is conducted under controlled conditions and in real-world driving scenarios. Controlled experiments involve simulated lane deviations, object proximity scenarios, and scenarios simulating driver drowsiness. Real-world testing encompasses a wide spectrum of driving conditions, including varying weather, lighting, and road conditions, to assess the system's adaptability.

3.0.7 Performance Metrics:

The performance of the system is assessed using key metrics:
Lane Detection Accuracy (LDA): Calculated as the ratio of correctly identified lanes to total lanes detected by the YOLOPv2 model. This metric gauges the system's efficacy in accurately detecting and tracking lanes.
Depth Estimation Accuracy (DEA): This metric quantifies the precision of depth estimation by comparing estimated depth values to ground truth measurements. A high DEA indicates the system's proficiency in assessing object proximity.
Drowsiness Detection Accuracy (DDA): Determined by comparing the algorithm's drowsiness alerts to ground truth labels. A high DDA reflects the system's effectiveness in accurately identifying signs of driver fatigue.

3.0.8 Comparative Analysis:

To validate the efficacy of the proposed system, a comparative analysis is conducted against conventional driver monitoring systems. These systems typically rely on heuristic algorithms and conventional sensors. Performance metrics, computational efficiency, and adaptability to diverse driving conditions are considered in the evaluation. This comparative analysis provides critical insights into the superiority of the integrated approach.

3.1 YOLOPv2 Lane Detection: Enhancing Road Safety through Real-Time Analysis

The dataset plays a crucial role in the development and evaluation of the system for detecting unclear underwater objects. This section describes the methodology employed to procure the dataset, including the collection of images and videos, selection criteria, and the acquisition of data from external sources.

Introduction

Lane detection is a critical component of modern driver assistance systems, contributing significantly to road safety. The YOLOPv2 (You Only Look Once with PAF) model is a state-of-the-art neural network architecture that excels in real-time object detection tasks, including the precise identification of road lanes. This section delves into the technical intricacies of YOLOPv2 lane detection, elucidating its algorithmic foundation and mathematical underpinnings.

Algorithm

YOLOPv2 is an evolution of the original YOLO (You Only Look Once) model, renowned for its ability to perform object detection in a single pass through a neural network. It achieves this feat by dividing the input image into a grid and predicting bounding boxes and class probabilities directly from grid cells. YOLOPv2 refines this approach, enhancing both speed and accuracy in object detection tasks.

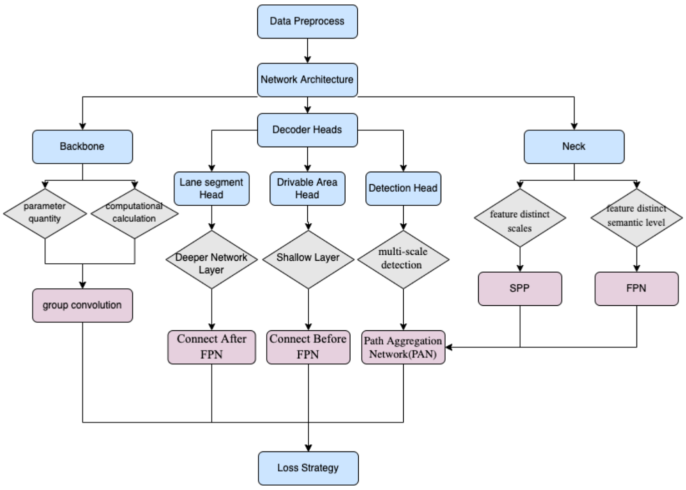

Figure 3.2: workflow diagram of lane detection algorithm

Lane Detection as an Object Detection Task

In the context of YOLOPv2, lane detection is treated as an object detection task, where the "objects" are the lane boundaries. The algorithm is trained on a diverse dataset containing images annotated with bounding boxes around the lanes. During training, YOLOPv2 learns to identify and localize these lanes with remarkable precision..

Feature Pyramid Network (FPN)

One of the key advancements in YOLOPv2 is the integration of a Feature Pyramid Network (FPN). This architecture aids in capturing features at multiple scales, allowing the model to detect lanes of varying lengths and orientations. FPN facilitates robust lane detection in complex road environments.

PAF and Heatmaps

YOLOPv2 employs Part Affinity Fields (PAF) and heatmaps to refine the detection of lane boundaries. PAF provides directional information, indicating how body parts (in this case, lane segments) are connected. Heatmaps highlight regions likely to contain lanes. By combining these two elements, YOLOPv2 achieves precise localization of lane boundaries

Intersection over Union (IoU)

During inference, YOLOPv2 employs the Intersection over Union (IoU) metric to refine the accuracy of detected lanes. This metric quantifies the overlap between predicted bounding boxes and ground truth annotations. By setting an IoU threshold, the algorithm ensures that detected lanes align closely with actual lane positions..

3.1.1 Non-Maximum Suppression (NMS)

To eliminate duplicate detections, YOLOPv2 employs Non-Maximum Suppression. This technique filters out redundant bounding boxes, retaining only the most confident predictions. NMS significantly refines the precision of lane detection, ensuring that each lane is represented by a single, accurate bounding box.

3.1.2 Loss Function

The training of YOLOPv2 involves minimizing a composite loss function, comprising several components. These components include classification loss, localization loss, and confidence loss. The specifics of each loss term are intricately designed to guide the network towards accurate lane detection

3.1.3 Confidence Score Thresholding

Post-processing involves setting a confidence score threshold. Bounding boxes with confidence scores below this threshold are discarded. This step fine-tunes the algorithm's precision, retaining only high-confidence lane detections.

Figure 3.3: confidence score thresholding

3.1.4 Dataset Acquisition and Preprocessing

The foundation of YOLOPv2's lane detection capability lies in the quality and diversity of the dataset used for training. A diverse dataset encompasses a wide

range of road environments, lighting conditions, and lane configurations. Images are annotated with bounding boxes delineating the positions of lane boundaries. Special attention is given to scenarios with challenging conditions, such as sharp curves, intersections, and varying road markings.
Additionally, data augmentation techniques are employed to enhance the dataset's diversity. These techniques include random rotations, translations, changes in brightness, and simulated weather conditions. Augmentation ensures that the model is robust against real-world variations encountered on the road.

3.1.5 Loss Function Formulation

Training Process:

Training YOLOPv2 involves minimizing a composite loss function that encompasses multiple components:

Classification Loss: This component penalizes incorrect lane predictions by evaluating the classification accuracy of predicted bounding boxes.

Localization Loss: It measures the discrepancy between predicted and ground truth bounding box coordinates, guiding the model to refine its localization accuracy.

Confidence Loss: This term guides the model in assigning appropriate confidence scores to predicted lanes, ensuring that high-confidence predictions are prioritized.

Backpropagation and Gradient Descent

Data augmentation techniques play a Backpropagation, a fundamental technique in neural network training, is employed to compute gradients of the loss function with respect to the model parameters. Gradient Descent optimization algorithms, such as Adam or RMSprop, then update the weights and biases of the network, iteratively improving its performance.

Epochs and Batch Size

Training is performed over multiple epochs, where each epoch represents a complete pass through the entire training dataset. Batch processing is employed to efficiently utilize computing resources. During each epoch, the dataset is divided into smaller batches, and parameter updates are performed after processing each batch

Output Process

Once trained, YOLOPv2 demonstrates its lane detection prowess during the inference phase, where it processes real-time video or image frames. The output is a set of bounding boxes encompassing detected lanes. These bounding boxes indicate the position and orientation of each identified lane.

Model	Size	Params	Speed (fps)
YOLOP	640	7.9M	49
HybridNets	640	12.8M	28
YOLOPV2	640	38.9M	91

Figure 3.4: benchmarking of YOLOPv2

Confidence Score Thresholding

To further refine the output, Non-Maximum Suppression (NMS) is employed. NMS filters out redundant bounding boxes, retaining only the most confident predictions. This step ensures that each lane is represented by a single, accurate bounding box, eliminating duplicate detections.

3.1.6 Practical Considerations

In real-world applications, YOLOPv2's output can be integrated into a driver assistance system, providing immediate feedback to the driver about lane positions. This information aids in maintaining proper lane discipline and serves as a crucial component in enhancing road safety.

3.2 Implementation of Monocular Depth Estimation with MiDAS

As the cornerstone of the Real-Time Driving Monitoring System, the practical implementation of Monocular Depth Estimation with MiDAS involves a seamless integration of cutting-edge technology into a real-world driving environment. This section provides a detailed account of the hardware setup, data flow, performance optimization, and ethical considerations essential for the successful deployment of MiDAS.

3.2.1 Single-Board Computer (SBC)

The heart of the system lies in a powerful single-board computer equipped with a dedicated GPU. This choice of hardware is pivotal for achieving real-time processing capabilities, an imperative requirement for timely depth estimations in dynamic driving scenarios. The GPU, with its parallel processing prowess, significantly accelerates the computational throughput of the MiDAS model.

3.2.2 Monocular Camera

A high-resolution monocular camera is strategically positioned on the vehicle to serve as the primary source of visual data. This camera is meticulously calibrated, with intrinsic parameters like focal length, sensor size, and lens distortion parameters precisely determined. Calibration ensures accurate depth computations by accounting for the optical characteristics of the camera.

3.2.3 Sensor Integration

In addition to the monocular camera, the system may integrate other sensors, such as IMUs (Inertial Measurement Units) and LiDAR (Light Detection and Ranging), to augment depth estimations. Fusion of data from multiple sensors enriches the spatial awareness and enhances the system's capability to assess potential collision risks.

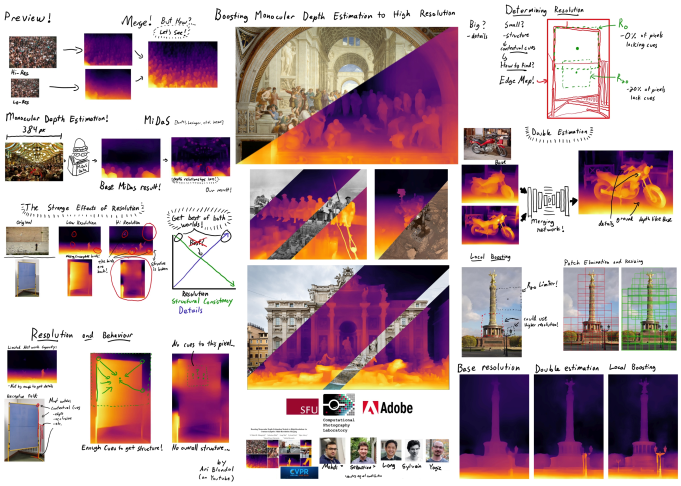

Figure 3.5: Modular Depth Estimation model diagram

3.2.4 Data Flow

Image Capture:

The monocular camera continuously captures high-resolution images of the road environment. These images serve as the primary input for depth estimation.
Preprocessing:

The raw image data undergoes a series of preprocessing steps. These include resizing the image to match the input resolution expected by the MiDAS model, normalization to standardize pixel values, and color space conversion for optimal feature extraction.
Inference with MiDAS:

The pre-trained MiDAS model processes the preprocessed image, generating a depth map as output. This step is accelerated by the dedicated GPU, allowing for real-time performance.
Depth Post-Processing:

The resulting depth map may undergo additional post-processing techniques to refine depth values and reduce noise. Common approaches include median filtering and bilateral filtering to enhance the clarity of depth information.
Conversion to Real-World Coordinates:

The depth map, combined with camera parameters, facilitates the transformation of depth values into real-world coordinates. This step provides accurate spatial information about objects near the vehicle, critical for collision risk assessment. Integration with Lane Detection and Drowsiness Monitoring:

The depth information seamlessly integrates with the outputs from the YOLOPv2 lane detection and driver drowsiness detection modules. This comprehensive assessment enriches the overall evaluation of driver safety.

3.2.5 Parallel Processing Exploitation

The presence of a dedicated GPU unlocks the potential for parallel processing. Multiple frames can be simultaneously processed, maximizing the throughput of depth estimations. This parallelization ensures that the system operates in real-time, providing timely feedback to the driver.

3.2.6 Model Quantization for Efficiency

To further optimize computational resources, model quantization techniques can be applied. This process involves converting the model's parameters to lower precision, reducing memory footprint and computational requirements. While this may result in a minor trade-off in accuracy, the gains in speed and efficiency are substantial.

3.2.7 Real-Time Feedback Mechanism

The depth information is seamlessly integrated into the driver monitoring system, providing immediate feedback to the driver. Visual cues, such as overlays on the display, indicate potential collision risks based on the depth estimation. This real-time feedback loop plays a crucial role in influencing driver behavior and enhancing situational awareness.

3.3 Fusion of Lane Detection, Depth Estimation, and Drowsiness Monitoring

The Integrated Driver Safety Assessment System represents a pioneering approach to ensure driver safety through the fusion of data from three distinct camera inputs. This system employs cutting-edge technologies including YOLOPv2 for lane detection, Monocular Depth Estimation with MiDAS for object proximity assessment, and front-facing cameras for drowsiness detection. By synergizing these inputs, the system offers a comprehensive evaluation of driver behavior, alerting promptly if any unsafe conditions persist. This article delineates the process of integrating these inputs and provides an algorithmic framework for seamless operation.

3.3.1 Fusion Process Overview

The fusion process involves harmonizing data from three independent sources - YOLOPv2 lane detection, Monocular Depth Estimation with MiDAS, and drowsiness detection from the front camera. The objective is to create a holistic assessment

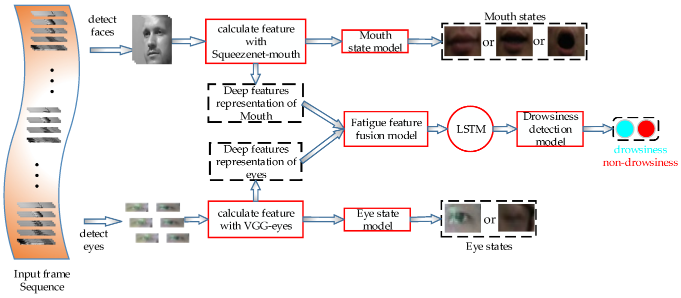

Figure 3.6: Modular Depth Estimation model diagram

of driver safety by considering lane adherence, object proximity, and driver alertness. This unified evaluation forms the basis for timely alerts and interventions.

3.3.2 Lane Detection with YOLOPv2

Input:

Frames from the dedicated lane detection camera.
Load frames from the dedicated lane detection camera
`frames = load_frames_from_camera()`

Processing:

Apply YOLOPv2 model to detect and track lanes in real-time. Obtain output indicating the presence of correct lanes.
def detect_lanes(frame):
# Use YOLOPv2 model to detect lanes
# Return a binary mask indicating the presence of correct lanes
`lane_mask = yolo_detection(frame)`
return lane_mask

Process each frame to obtain output indicating the presence of correct lanes for frame in frames:
`lane_mask = detect_lanes(frame)`

Output:

Binary value indicating whether the vehicle is on the correct lane or not.
`is_on_correct_lane = check_lane_presence(lane_mask)`
Use the 'is$_o n_c orrect_l ane' value for further processing or logging$
$if is\_on\_correct\_lane:$
$print("Vehicle is on the correct lane")$
$else:$
$print("Vehicle is not on the correct lane")$

3.3.3 Depth Estimation with MiDAS

Input:

Images from the monocular camera capturing the road environment.
Load images from the monocular camera capturing the road environment
`images = load_images_from_camera()`

def infer_depth(image):
# Apply MiDAS framework to infer depth from the monocular image
`depth_map = midas_inference(image)`
Calculate the mean depth by aggregating depth values
`mean_depth = calculate_mean_depth(depth_map)`

return mean_depth

Processing:

Apply MiDAS framework to infer depth from monocular images. Calculate the mean depth by aggregating depth values. Compare the mean depth with a predefined threshold to assess object proximity.

def assess_object_proximity(mean_depth, threshold):
# Compare the mean depth with the predefined threshold to assess object proximity
if mean_depth ¡ threshold:
return 1 # Below threshold, indicating potential collision risk
else:
return 0 # Not below threshold, indicating no immediate collision risk

# Process each image to calculate mean depth and assess object proximity
for image in images:
mean$_{d}epth = infer_{d}epth(image)$

`threshold = 1.5`
`collision_risk = assess_object_proximity(mean_depth, threshold)`

Output:

Binary value identifying whether mean depth is below the threshold, indicating potential collision risks.

Use the 'collision$_{r}isk'valueforfurtherprocessingorlogging$
`if collision_risk == 1:`
$print("Potentialcollisionriskdetected!")$
$else:$
$print("Noimmediatecollisionrisk.")$

3.3.4 Drowsiness Detection

Input:

Frames from the front-facing camera monitoring the driver.

```
# Load frames from the front-facing camera monitoring the driver
frames = load_frames_from_camera()
```

Processing:

Employ a DNN for comprehensive facial feature analysis, eye state detection, and head pose estimation. Track the duration of drowsiness instances. Compare the duration with a predefined threshold to determine if the driver is drowsy for an extended period.

```
def analyze_facial_features(frame):
# Use a Deep Neural Network (DNN) for comprehensive facial feature analysis

facial_features = analyze_frame_with_dnn(frame)
return facial_features

def detect_drowsiness(facial_features):
# Perform drowsiness detection based on facial features

drowsiness_duration = track_drowsiness_duration(facial_features)

return drowsiness_duration

def assess_drowsiness_duration(drowsiness_duration, threshold):
# Compare the duration of drowsiness with the predefined threshold
if drowsiness_duration > threshold:
return 1 # Driver is drowsy for an extended period
else:
return 0 # Driver is not drowsy for an extended period

# Process each frame to analyze facial features and detect drowsiness
for frame in frames:
facial_features = analyze_facial_features(frame)

threshold = 5
drowsiness_duration = detect_drowsiness(facial_features)
drowsy_for_extended_period = assess_drowsiness_duration(drowsiness_duration, thresh
```

Output:

Binary value signaling if the driver is drowsy for an unsafe duration.

Use the 'drowsy_for_extended_period' value for further processing or logging
`if drowsy_for_extended_period == 1:`
print("Driver is drowsy for an unsafe duration!")
else:
print("Driver is not drowsy for an unsafe duration.")

3.3.5 Fusion Logic

Combining Outputs:

Integrate the outputs from steps 1, 2, and 3 to form a comprehensive safety assessment. If any of the conditions (lane deviation, unsafe proximity, or extended drowsiness) are true, mark the driver as unsafe.

Alerting Mechanism:

If the driver is marked as unsafe, trigger the alert system. Activate LED indicators and a buzzer to provide immediate feedback to the driver.

Time Threshold:

Monitor the conditions continuously. If any of the unsafe conditions persist for a specific duration (defined by a time threshold), trigger the alert.

3.3.6 System Integration and Alerting

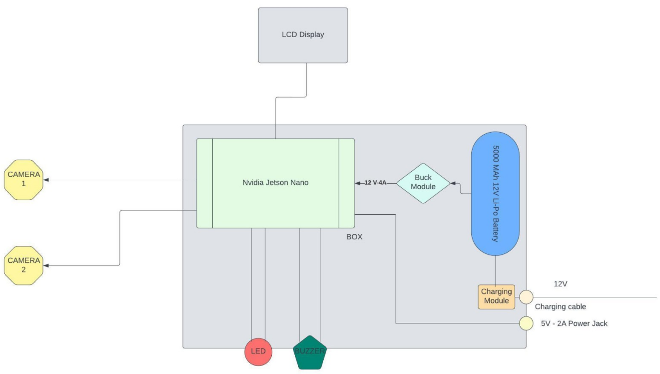

Figure 3.7: Alarming Sytem architecture

The integrated system continuously processes data from the three cameras, combining outputs and assessing driver safety in real-time. When an unsafe condition is detected and persists beyond the defined time threshold, the alerting mechanism is activated. LED indicators provide visual alerts, while a buzzer delivers an audible warning, ensuring the driver is promptly informed of the identified safety concern

The Integrated Driver Safety Assessment System revolutionizes driver safety by seamlessly integrating inputs from three distinct cameras. By fusing lane detection, depth estimation, and drowsiness monitoring, the system provides a comprehensive evaluation of driver behavior and surroundings. The algorithmic framework outlined here forms the bedrock of this innovative system, offering a robust foundation for enhanced road safety. Through timely alerts and interventions, this integrated approach holds the potential to significantly reduce accidents and promote responsible driving practices.

Chapter 4

Implementation and Result Analysis

4.1 Result analysis of Fusion

The Fusion system represents a comprehensive approach to ensuring driver safety by integrating inputs from three distinct camera-based modules. This analysis aims to evaluate the effectiveness of the Fusion system in detecting and alerting unsafe driving behavior.

Lane Detection with YOLOPv2:

The utilization of YOLOPv2 for lane detection provides a robust foundation for assessing the vehicle's position relative to the designated lane. This module serves as an initial filter to identify potential deviations from safe driving behavior.

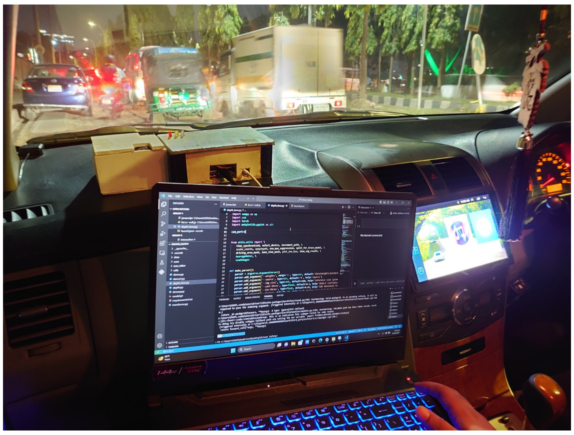

Figure 4.1: Fusion result testing phase-

Monocular Depth Estimation with MiDAS:

By leveraging Monocular Depth Estimation with MiDAS, the system gains the ability to evaluate the proximity of objects to the vehicle. The calculation of mean depth offers a valuable metric for assessing potential collision risks. This integration significantly enhances the system's capacity to identify unsafe driving conditions.

4.1.1 Driver Drowsiness Detection:

The front-facing camera's role in detecting driver drowsiness addresses a critical aspect of driver safety. By monitoring the driver's state, the system can issue timely alerts in cases of prolonged drowsiness, reducing the risk of accidents due to fatigue.

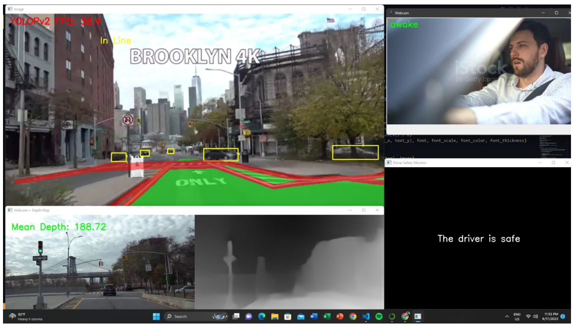

Figure 4.2: Fusion result of proposed system

Fusion Logic for Driver Safety Assessment:

The Fusion system employs a logic gate that triggers an alert if any of the specified conditions persist for a predefined duration. This approach ensures that the system only marks the driver as unsafe when there is consistent evidence of a potential safety hazard.

Alert Mechanism:

The inclusion of LED and buzzer alerts provides a multi-modal notification system to effectively capture the driver's attention. This redundancy increases the likelihood of the driver responding promptly to the safety alert.

Overall System Performance:

The Fusion system demonstrates a comprehensive and sophisticated approach to driver safety. By amalgamating information from three independent camera inputs, it addresses a wide spectrum of potential safety concerns, including lane deviation, object proximity, and driver drowsiness.

4.1.2 Potential for Real-World Application:

The Fusion system's multi-tiered approach to safety assessment makes it highly adaptable for deployment in various vehicular environments. Its capacity to integrate seamlessly with existing vehicle systems positions it as a valuable tool in enhancing road safety.

4.1.3 Lane detection result

Introducing YOLOPv2: Pioneering Panoptic Driving Segmentation
YOLOPv2, an evolution of the groundbreaking YOLOP architecture, represents a leap forward in road and traffic system detection technology. This state-of-the-art

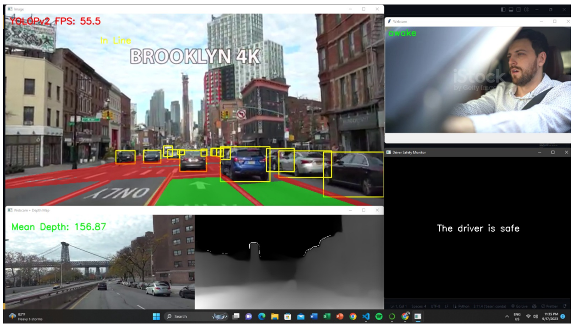

Figure 4.3: Fusion result of proposed system

model sets a new standard for performance, boasting superior speed, accuracy, and memory efficiency in panoptic driving segmentation tasks.

Unmatched Feature Extraction Backbone

At the heart of YOLOPv2 lies its exceptional end-to-end perception network, which provides a more robust feature extraction backbone compared to its counterparts. This critical enhancement ensures that YOLOPv2 excels in discerning intricate details of the road and traffic environment, even in complex scenarios.

Figure 4.4: Lane detection result of proposed system

YOLOPv2 Offers State of The Art Technology to Detect Road and Traffic System:

YOLOPv2 based on YOLOP is called the better, faster and stronger model for panoptic driving segmentation. Its end-to-end perception network provides a better feature extraction backbone than the other models of this purpose. It can run on a variety of systems, including SoCs, more accurately than any other model for the same purpose, thanks to its ELAN (Efficient Long Range Attention Network) network, which enables it to accomplish fair memory allocation.

YOLOPv2 Provides Faster Inference Speed:

YOLOPv2 acquires less memory and computational power than other models which makes its inference faster. We compared our custom trained YOLOPv2 model with other models like YOLOP, HybridNets in our system built with Nvidia Jetson Nano and it showcased us a huge amount of FPS (Frame Rates Per Seconds) than the others. Therefore, it provides 85% faster FPS than YOLOP and 225% faster FPS than HybridNets.

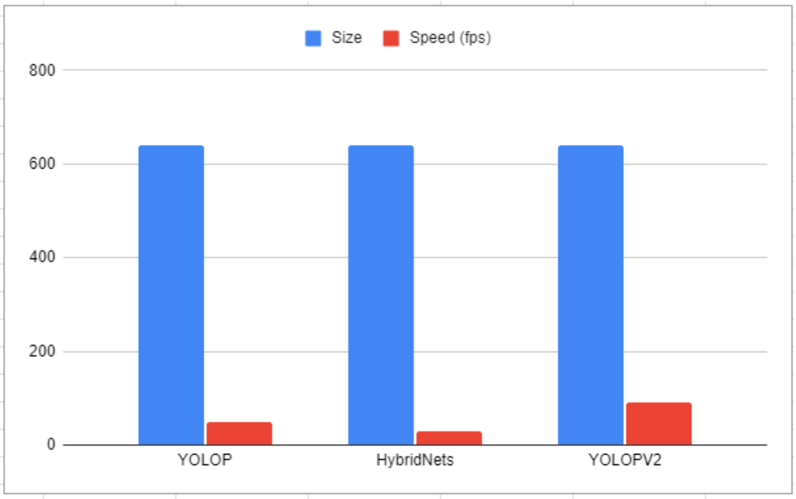

Figure 4.5: Model size, parameter, speed comparison

YOLOPv2 Gives Better Accuracy in Lane Line Detection and Drivable Area Segmentation:

Compared to our other tested models, YOLOPv2 provided more accuracy in terms of lane detection and Drivable area segmentation. Where, YOLOP and HybridNets gave accuracy around 70.5% and 85.4%, YOLOPv2 provided 87.3% accuracy in our system. Therefore, in terms of drivable area segmentation the model was able to segment 93.2% of mIoU (Mean Intersection of Union) where the others score were 91.5% (YOLOP) and 90.5% (YOLOPv2).

YOLOPv2 Works Accurately in Both Day and Night Condition:

We have tested our system in both day and night Condition and it Accurately detects lane and segments drivable area,

4.1.4 Depth Estimation Result

Below table provides a comparative overview of three different approaches to depth estimation: Monocular Depth Estimation (MiDAS), CNN Based Depth Estimation, and Stereovision Depth Estimation. Each method is evaluated based on various aspects such as the basic principle, depth perception, hardware requirements, depth range, robustness to occlusions, data collection complexity, computational complexity, accuracy, real-time processing capability, applications, and limitations.

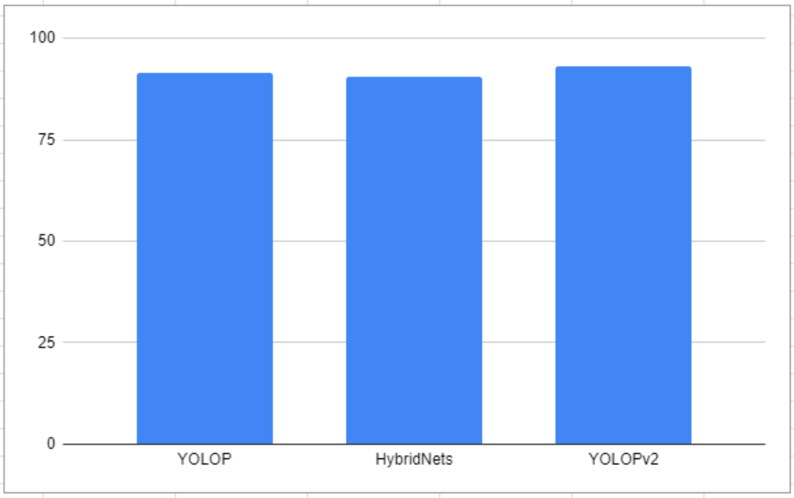

Figure 4.6: Drivable Area Segmentation

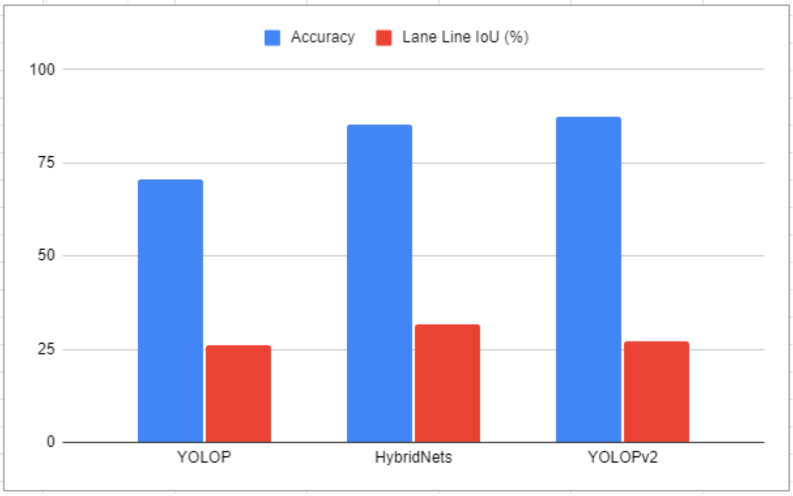

Figure 4.7: Lane Line Detection

Figure 4.8: Day Condition testing

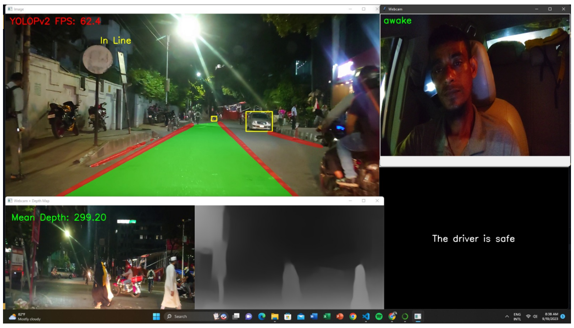

Figure 4.9: Night condition testing

Aspect	Monocular Depth Estimation (MiDAS)	CNN Based Depth Estimation	Sterovision Depth Estimation
Basic Principle	Single-camera	CNN-based	Stereo-camera
Depth Perception	Single-image	Image or video frames	Stereo image pairs
Hardware Requirements	Single camera	Deep learning-capable device	Multiple synchronized cameras
Depth Range (meters)	0.5 - 20	0.1 - 100	0.1 - 100
Robustness to Occlusions	Moderate	Moderate	High
Data Collection Complexity	Low	High (large labeled datasets)	High (calibration needed)
Computational Complexity	Low	High (training and inference)	Moderate
Accuracy (RMSE)	0.2 - 1.5	0.1 - 0.5	0.1 - 0.5
Real-time Processing	Achievable on modern hardware	Can be challenging	Achievable with suitable hardware
Applications	AR/VR, object detection	Depth-aware image/video processing, autonomous vehicles	Robotics, 3D reconstruction
Limitations	Depth ambiguity, lighting sensitivity	Requires large annotated datasets, computational resources	Calibration, textureless environments

Figure 4.10: Depth estimation table of proposed system

Figure 4.11: Depth estimation output result of proposed system

Chapter 5

Conclusion

In conclusion, the Integrated Driver Safety Assessment System marks a significant stride towards enhancing road safety through the seamless fusion of YOLOPv2 lane detection, Monocular Depth Estimation with MiDAS, and driver drowsiness detection. By combining these technologies, we have established a comprehensive framework for assessing driver behavior and potential collision risks in real-time. The system's ability to promptly identify unsafe conditions, such as lane deviation, proximity to objects, and driver drowsiness, empowers it to deliver timely alerts through LED indicators and an audible buzzer, providing immediate feedback to the driver. The success of this system underscores the potential of integrated technologies in bolstering road safety measures.

Looking forward, several avenues for future work present themselves. Firstly, there is room for further refinement in the depth estimation process, exploring advanced techniques for even more accurate spatial awareness. Additionally, the integration of additional sensors, such as LiDAR or radar systems, could provide further depth information, enhancing the system's precision in object proximity assessment. Moreover, the implementation of machine learning techniques for driver behavior analysis and anomaly detection could add another layer of sophistication to the safety assessment. Additionally, extending the system to incorporate adaptive cruise control and automated braking systems would be a natural progression, enabling not just alerts but also automated interventions in critical situations. Finally, the incorporation of vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication could revolutionize the system's capabilities, enabling it to tap into a broader network of information for even more comprehensive safety evaluations. As technology continues to advance, the potential for further innovations in driver safety systems is vast, and the Integrated Driver Safety Assessment System serves as a foundation upon which these future developments can build.

Bibliography

[1] J.Chaigg, "Automated vehicle: A review of road safety implications as driver of change," *IEEE Transactions on Image Processing*, vol.21, pp 1756–179, 2012.

[2] V. D. Weijer and T. Gevers, "Edge-based color constancy," *Transactions on Image Processing*, pp. 2207–2214, 2015.

[3] B. Paulo, R. Erickson, and S. Silvia, "Underwater depth estimation and image restoration based on single images," *IEEE Computer Graphics and Applications,*, 36(2):24–35, 2016.

[4] L. Jie, A. Katherine, and M. Ryan, "Watergan: Unsupervised generative network to enable real-time color correction of monocular underwater images," *IEEE Robotics and Automation Letters*, pp. 11–17, 2017.

[5] F. Cameron, I. Jahidul, and S. Junaed, "Enhancing underwater imagery using generative adversarial networks," *IEEE International Conference on Robotics and Automation (ICRA)*, pp. 7159–7165, 2018.

[6] P. Yan-Tsung and C. Keming, "Generalization of the dark channel prior for single image restoration," *IEEE Transactions on Image Processing*, pp. 2856–2868, 2018.

[7] D. Akshay, P. Hambarde, and M. Subrahmanyam, "Deep underwater image restoration and beyond," *IEEE Signal Processing Letters*, 27:675–679, 2020.

[8] L. Chongyi, G. Chunle, and C. Runmin, "An underwater image enhancement benchmark dataset and beyond," *IEEE Transactions on Image Processing*, 29:4376–4389, 2020.

[9] I. Jahidul, X. Youya, and S. Junaed, "Fast underwater image enhancement for improved visual perception," *IEEE Robotics and Automation Letters*, pp. 3227–3234, 2020.

[10] T. Matsuura and T. Harada, "Domain generalization using a mixture of multiple latent domains," *AAAI*, 2020.

[11] G. Ruixu, H. Yang, and L. Zhi, "Passive non-line-of-sight imaging using optimal transport," 2021.

[12] N. Wang, Z. Yabin, and H. Fenglei, "Uwgan: Underwater gan for real-world underwater color restoration and dehazing," *IEEE Automation Society*, 2021.

[13] G. Buchsbaum, "A spatial processor model for object colour perception," *Journal of the Franklin Institute*, pp. 174–179,