بسم الله الرحمن الرحيم

BACHELOR OF SCIENCE IN ELECTRONIC AND TELECOMMUNICATIONS ENGINEERING

Unsafe Driving Behavior Detection By Using CNN Deep Learning Method.

Submitted by

Shihab Adnan (T191010)

Supervised by

Prof. Engr. Md. Razu Ahmed

Chairman, Department of Computer and Communication Engineering, IIUC.

Department of Electronic and Telecommunication Engineering
International Islamic University Chittagong
Kumira, Chittagong–4318
February, 2024

Declaration

The work in this thesis, with the exception of properly cited quotations and summaries, is our own, we hereby declare.

Shihab Adnan

ID: T191010

Certificate of Approval

The thesis entitled as **“Unsafe Driving Behavior Detection By Using CNN Deep Learning Method”** Submitted by **Shihab Adnan**, bearing ID No: T191010 , to the Department of Electronic and Telecommunications Engineering (ETE) of International Islamic University Chittagong (IIUC) has been accepted as satisfactory for the partial fulfillment of the requirements for the Degree of Bachelor in Electronic and Telecommunications Engineering and approved as to its style and contents for the examination held on _____________2024.

Approval of Supervisor

The thesis of this Study is **" Unsafe Driving Behavior Detection By Using CNN Deep Learning Method".** This study is submitted by **Shihab Adnan (T191010)** to the Department of Electronic and Telecommunication Engineering (ETE) of International Islamic University Chittagong (IIUC). It has been satisfactory for the partial fulfillment of the requirements for the bachelor's degree of Science in Electronic and Telecommunication Engineering (ETE). It was approved by our supervisor.

____

Approved by

Prof. Engr. Md. Razu Ahmed

Supervisor

Chairman
Department of Computer and Communication Engineering
Faculty of Science and Engineering
International Islamic University Chittagong
Kumira, Chittagong.

Acknowledgement

With deep appreciation, we thank Almighty Allah for all his favors, including our good health and patience, which helped us to complete our thesis. We would also like to express our gratitude to our research supervisor, **Prof. Engr. Md. Razu Ahmed**, whose continuous support and supervision made this thesis possible. We consider ourselves lucky to have **Prof. Engr. Md. Razu Ahmed** as our research supervisor, who has helped and advised us during the entire thesis. A special thanks to the instructors and employees of the Department of ETE, IIUC for their invaluable contributions. We are incredibly appreciative of our parents unwavering love and support as well as the strength their prayers have given us. We are especially appreciative of our classmates help, they are excellent.

Abstract

The complex nature of human behaviour and the numerous factors contributing to distractions on the road, particularly in terms of the limited ability to provide precise and timely early warnings, are addressed in this paper through the introduction of a novel method. The proposed method aims to enhance the detection of distracted driver behaviour using modified CNN transfer learning. By employing transfer learning within the Convolutional Neural Networks framework, the method seeks to tackle the intricate challenges associated with identifying distracted driving behaviours. The experimental results demonstrate an overall accuracy of approximately 99%, with the highest achieved accuracy reaching 99.46% on a publicly available dataset. This highlights the effectiveness of the proposed approach in significantly improving the precision of distracted driver behaviour detection.

Table of Contents

List of Figure

List of Table

List of Abbreviations

VGG	Visual Geometry Group
SGD	Stochastic Gradient Descent
ADAM	Adaptive Moment Estimation
ReLU	Rectified Linear Unit
RMSProp	Root Mean Square Propagation
AdaGrad	Adaptive Gradient
CNN	Convolutional Neural Network
AI	Artificial Intelligence
ML	Machine Learning
DL	Deep Learning
FC	Fully Connected
Conv	Convolutional Layer
DNN	Deep Neural Network
FCN	Fully Convolutional Network
AP	Average Precision
KNN	K-Nearest Neighbors
XGB	XGBoost
SVM	Support Vector Machine

Chapter 1
Introduction

1.1 Introduction

Getting distracted while operating a vehicle is extremely dangerous, as it takes attention away from the road and increases the risk of accidents. It is estimated that distracted driving contributes to around 25% of all motor vehicle accidents [1]. Common distractions include texting, talking on the phone, eating, and using in-car technologies. The National Highway Traffic Safety Administration (NHTSA) revealed that in 2023, there were approximately 3,000 fatalities and 400,000 injuries caused by distracted driving in the United States alone [2]. There were also significant economic costs associated with distracted driving. According to a study conducted by [3], the annual cost of crashes caused by distracted driving is estimated to be around $40 billion. These costs include medical expenses, property damage, and lost productivity [4]. Additionally, distracted driving can have long-lasting emotional and psychological impacts on the individuals involved in accidents, as well as their families and loved ones [5]. Another important factor to consider is that distracted driving not only puts the driver at risk but also endangers the lives of pedestrians and other motorists on the road. Studies have shown that the use of handheld electronic devices while driving is particularly dangerous, as it takes both visual and cognitive attention away from the task of driving, leading to a higher likelihood of accidents [6]. Many countries have implemented laws that prohibit or restrict the use of handheld electronic devices, but there is still a lack of compliance and awareness among drivers. In this case, technological advancements such as hands-free systems and voice-activated controls can provide safer alternatives for drivers who need to stay connected while on the road. These advancements allow drivers to keep their hands on the wheel and their eyes on the road, reducing the risk of accidents caused by distracted driving [7].

ISO (International Organization for Standardization) has developed a standard that provides guidelines for implementing a road traffic safety management system to help organizations address the issue of distracted driving. This standard aims to reduce the number of accidents caused by distractions and promote safer driving practices worldwide [8].

Another potential solution to address the issue of distracted driving is the application of deep learning technology. Deep learning, a subset of artificial intelligence, has the potential to further enhance road safety by enabling advanced driver assistance systems. These systems can analyze real-time data from various sensors and cameras to detect potential hazards and alert drivers, helping to prevent accidents caused by distractions or human error. Driver monitoring systems [9] measure various parameters, such as eye movement, head position, and even heart rate, to detect signs of distraction. By analyzing this data in real-time, deep learning algorithms can accurately identify when a driver is not fully focused on the road and issue timely warnings or take corrective actions. Deep learning, as a pivotal player in innovative statistical and computational modeling, has reshaped the landscape of data analysis [10]. Its application in the realm of driver behavior analysis holds the potential to not only enhance our understanding of how drivers behave but also to swiftly translate these insights into real-life improvements. An indispensable initial step in this analysis involves the comprehensive comprehension and categorization of individual driver skill levels and habitual driving patterns. The definition of various parameters crucial for such classification and analysis precedes the actual examination. Parameters encompass a wide array of factors, including driving speed, acceleration patterns, lane-keeping behavior, and response times to stimuli. The nuanced understanding of these parameters allows for a more detailed and accurate characterization of driver behavior, contributing to a more holistic analysis [11].

To achieve this, the availability and quality of driving data become paramount. On-board sensors or sensor-equipped smartphones are frequently employed, capturing a diverse and extensive set of data points during the course of driving. Additionally, a vehicle controller serves as a valuable tool to extract compliance-related data. This dataset [12] encompasses not only basic metrics such as steering wheel settings, vehicle speed, and brake pedal position but also sensor data related to the vehicle's speed, direction, and velocity.

Furthermore, the integration of global positioning system (GPS) data provides valuable insights into the spatial context of driver behavior [13]. This allows for the analysis of driving patterns in specific locations or under certain environmental conditions, contributing to a more context-aware understanding of driver behavior. The combination of these diverse data sources forms a comprehensive dataset for deep learning models to learn intricate patterns of driver behavior. The problem of distracted driving has received considerable attention in recent years due to its worrisome impact on the occurrence of motor vehicle accidents and fatalities [14]. Although your thesis examines several facets of distracted driving and suggests remedies like hands-free systems and deep learning technology, there are further layers to be explored within this intricate subject. An important factor to take into account is the impact of cultural and behavioral elements on distracted driving. Despite the acknowledged hazards, numerous persons persist in participating in distracting activities while operating a vehicle as a result of societal norms, perceived exigency, or just routine. Comprehending these fundamental variables is essential for formulating efficient treatments and encouraging the adoption of safer driving practices. The significance of education and awareness initiatives cannot be exaggerated when it comes to addressing distracted driving. Public outreach campaigns might aim to engage both motorists and passengers, highlighting the significance of maintaining concentration while driving and the possible repercussions of being distracted. These efforts can facilitate a change in social norms and decrease the occurrence of distracted driving by promoting a culture that emphasizes responsibility and accountability.

Furthermore, legislative actions are crucial in tackling the issue of distracted driving. Although the implementation of rules that forbid the use of handheld electronic devices is a positive move, the difficulties lie in effectively enforcing these laws and ensuring people's adherence to them. Enhancing the enforcement measures and imposing more severe penalties on violators can work as effective deterrents and promote the notion that distracted driving will not be accepted.

Moreover, incorporating distracted driving awareness and prevention methods into driver education and training programs might have a long-lasting effect. By inculcating safe driving practices from an early age, novice drivers are more adept at resisting distractions and making conscientious decisions while driving. Regarding technology solutions, hands-free systems and deep learning algorithms hold potential in reducing distracted driving. However, continuous research and development are necessary to improve their effectiveness and dependability. Furthermore, it is crucial to thoroughly tackle concerns related to privacy and data security in order to guarantee the widespread usage and approval of these technologies.

In leveraging deep learning for driver behavior analysis, the ultimate goal extends beyond the mere detection of risky behavior. The aspiration is to develop proactive systems that can anticipate and mitigate potential hazards. By understanding the subtle cues and patterns indicative of safe or risky driving, these models contribute to the development of advanced driver-assistance systems (ADAS) that prioritize safety and enhance overall road traffic management.

1.2 Research Background

Distracted driving, at the crossroads of technology, human behavior, and road safety, remains an intricate challenge requiring persistent research endeavors. As the definition of distracted driving expands to encompass a spectrum of behaviors diverting attention from the task of driving, recent advancements and emerging technologies add new layers of complexity to this issue. The advent of mobile devices, particularly smartphones, has played a pivotal role in the surge of distracted driving incidents. Beyond traditional distractions like eating or conversing with passengers, the omnipresence of smartphones introduces novel dimensions, ranging from texting to interacting with in-car technologies. A notable example of recent research in this realm is the State Farm Distracted Driving Behavior dataset [15], which, through the application of Convolutional Neural Networks (CNN), has offered insights into the nuanced aspects of distracted driving behaviors. While legal measures and public awareness campaigns strive to curb this pervasive problem, technological interventions, particularly utilizing artificial intelligence (AI) and CNNs, have demonstrated promise in real-time detection and mitigation. The State Farm Distracted Driving Behavior dataset, coupled with CNN architectures, showcases the potential for proactive interventions by issuing timely warnings when signs of distraction manifest.

However, the multifaceted nature of distractions remains a persistent challenge for accurate identification and prevention. The interplay of visual, manual, and cognitive elements necessitates a nuanced understanding that continues to be a focal point for ongoing research efforts. The adaptation of detection technologies, especially those leveraging CNNs, to diverse driving conditions and their seamless integration into existing vehicular infrastructure present complex challenges that demand further exploration. Legislative initiatives are essential for tackling distracted driving, alongside technological solutions. Nations across the globe are enacting more stringent regulations and sanctions to discourage drivers from participating in distracting activities. Public awareness initiatives additionally aim to instruct individuals about the perils of distracted driving and advocate for proper behavior while operating a vehicle. Moreover, the incorporation of sophisticated driver-assistance systems (ADAS) into automobiles shows potential for reducing distractions. These systems employ sensors, cameras, and AI algorithms to identify possible dangers and issue warnings or take action as needed, enhancing attempts to address distracted driving [16]. Furthermore, the rise of self-driving vehicles offers a fascinating opportunity to tackle the issue of distracted driving. Self-driving cars have the capacity to take over the driving responsibility, thereby eliminating human errors and significantly reducing distractions. Nevertheless, guaranteeing the dependability and security of autonomous technology remains a crucial challenge. In order to effectively tackle the issue of distracted driving, a comprehensive strategy is necessary, which integrates technology advancements, regulatory measures, public awareness campaigns, and shifts in society attitudes. Ongoing cooperation between researchers, policymakers, industry stakeholders, and the public is crucial in order to establish a more secure driving environment for everyone [17].

The consequences of distracted driving extend beyond statistics; they reverberate in human lives impacted by preventable accidents. Despite strides in technology and our comprehension of driver behavior, the dynamic nature of distractions mandates continuous research. Researchers aim to refine existing methodologies, drawing inspiration from recent works like those employing CNNs, and to develop innovative solutions that align with the evolving landscape of distractions. Incorporating new variables and refining algorithms are critical steps in enhancing the efficacy of detection systems, ultimately contributing to a safer driving environment.

1.3 Problem Statement

The extant challenges confronting distracted driving behavior detection systems manifest in their recurrent constraints in accurately discerning and categorizing unsafe driving behaviors. These limitations stem from the inherent variability in human conduct and the intricacy of differentiating between intentional and unintentional distractions. Conventional systems typically rely on the manual analysis and interpretation of sensor data, a procedurally cumbersome and error-prone endeavor.

Furthermore, a number of these extant systems are prone to generating false alarms or evince deficiencies in providing timely alerts, thereby engendering potential risks for accidents and compromising safety. The imperative for more dependable and efficacious driving behavior detection systems is unequivocal, necessitating a concerted effort to address these challenges and augment overall road safety. The act of driving while being distracted presents a substantial danger to the safety of the road. To effectively deal with this problem, it is essential to adopt a comprehensive strategy that surpasses conventional methods of detection. Machine learning algorithms and artificial intelligence have exciting opportunities to improve the precision and dependability of detecting systems. These systems can assess driving behavior more accurately and intelligently by using real-time data streams from several sensors such as cameras, accelerometers, and GPS. Incorporating cognitive models that imitate human decision-making processes can aid in distinguishing deliberate distractions from unintentional lapses in focus. Moreover, it is crucial to promote cooperation among academics, engineers, politicians, and stakeholders in order to create and execute strong measures that successfully reduce the hazards linked to distracted driving. By implementing collaborative initiatives and fostering creativity, we can work towards creating a road environment that is both safer and more secure for everyone.

1.4 Motivation

This research subject was chosen for a number of reasons. Due to safety concerns, automation is becoming more and more necessary for vital vocations like driving. Developments in machine learning and artificial intelligence have opened up new possibilities for innovation in the car industry. A competent technology that can specifically be used to verify and assess the legality of a move while driving is being developed by researchers via advanced study.The necessity for this research subject to develop efficient solutions to improve road safety is prompted by the rise in motor vehicle accidents brought on by driver distraction and inattention. This effort is anticipated to significantly lower accident rates and make the roads safer for all drivers. The importance of this technological innovation in improving safety and reducing the negative effects of human error while driving cannot be overstated. The selection of this research topic arises from a combination of crucial aspects that emphasize the need for automation in vital professions, specifically in the field of driving where safety issues are significant. The emergence of machine learning and artificial intelligence has brought up unparalleled prospects for revolutionary advancements in the automobile sector. In this context, researchers are developing an advanced technological system that can carefully analyze and confirm driving maneuvers accurately. This system aims to meet the urgent requirement for improved safety measures on our roadways. The importance of this research endeavor is emphasized by the increasing occurrence of motor vehicle accidents caused by driver distraction and lack of attention. This research seeks to provide effective solutions using modern approaches to improve road safety and reduce the high incidence of accidents. This program has the potential to significantly reduce accident rates, therefore creating safer roads for all drivers. Furthermore, this research seeks to tackle the root causes of accidents, such as driver distraction and inattention, with the goal of not only avoiding collisions but also reducing the economic and societal consequences that come with them. The profound capacity of this technical advancement to enhance safety and minimize the harmful consequences of human mistakes while driving is of utmost significance. As the research advances, it is expected to bring about a new era of vehicle safety. This will involve the combination of advanced technology and human effort, working together to protect lives and improve the entire driving experience.

1.5 Objective of Research

This research focuses on the development of new techniques for….

1. To create a dataset that includes ten new labels for distracted driver behavior.
2. To modify a CNN model (VGG-16) in order to identify the type of distracted driver behavior.

1.6 Organization of the Thesis

There are five chapters in this thesis report. Chapter One has the following sections: Introduction, Background, Problem Statement, Motivation, and Research Study Objectives. The subsequent section of our literature review concentrates on the essential components of the research papers that are relevant to our investigation and showcases the associated studies. Chapter 3 covers the research work, data processing, convolutional neural network model vgg-16 architectures, model training, testing, validation, and optimizers. It also outlines the study's methodology. Chapter 4 reports on the research work performance, findings, and debate. Chapter 5 describes the study's findings, conclusions, and potential contributions.

Chapter 2
Literature Review

2.1 Introduction

An increasing number of accidents and fatalities worldwide are caused by distracted driving, which is a serious threat to road safety. Developing efficient distracted driving detection systems has become necessary, and recent technological developments, especially in the areas of computer vision and machine learning, have made this possible. This section provides a concise overview of pertinent and notable contributions in the literature concerning the detection of distracted drivers. Distracted driving presents a substantial hazard to global road safety, as evidenced by a growing number of accidents and fatalities directly linked to this problem. The widespread utilization of mobile devices, in addition to other distractions such as eating, engaging in conversation with passengers, or making adjustments to in-car entertainment systems, has worsened the issue. Hence, it is imperative to create efficient detection systems in order to alleviate the hazards linked to distracted driving.

The latest progress in technology, namely in the fields of computer vision and machine learning, have created new opportunities for tackling this problem. These advancements have the capability to identify and intervene in cases of distracted driving immediately, therefore improving road safety and preserving lives.

Within the realm of literature, multiple approaches and methodologies have been suggested and investigated for the purpose of identifying distracted drivers. These methods involve utilizing powerful sensors and cameras mounted in vehicles to observe driver behavior, as well as employing complex algorithms that assess visual signs that suggest distraction. Moreover, the incorporation of machine learning methodologies empowers systems to acquire knowledge from data and consistently enhance their precision and dependability. Significant advancements in this domain involve the creation of algorithms that can identify facial expressions, eye movements, and head positions that suggest distraction. Furthermore, the incorporation of contextual information, such as vehicle dynamics and environmental conditions, strengthens the resilience of detection systems. In summary, the effort to develop effective distracted driving detection systems is a vital undertaking to guarantee the safety of the roads.

Further investigation and advancement in this field have the capacity to greatly diminish the occurrence and consequences of distracted driving-related accidents on a global scale.

2.2 Scope of Research

The analysis of unsafe driving behavior detection has garnered significant attention due to its potential to enhance road safety. With ongoing technological advancements, researchers and companies are actively developing diverse methods and algorithms aimed at accurately identifying and classifying unsafe driving behaviors [18]. These approaches leverage data from sensors, such as accelerometers and cameras, enabling the detection of actions such as speeding, harsh braking, lane weaving, and distracted driving [19]. A substantial focus in future research revolves around the development of real-time monitoring systems capable of providing immediate feedback to drivers, facilitating timely correction of behavior and accident prevention. The integration of artificial intelligence and machine learning algorithms is anticipated to further enhance the accuracy and effectiveness of these systems in detecting unsafe driving behaviors. The detection of unsafe driving behavior is a crucial field of study that continues to garner substantial interest due to its potential to greatly enhance road safety. With the growing dependence on technological breakthroughs, there is a heightened emphasis on creating efficient techniques and algorithms to detect and categorize hazardous driving actions. Recently, there has been a significant increase in research and technical advancements focused on utilizing data from different sensors, such as accelerometers and cameras, to identify and categorize various hazardous driving actions [20]. These behaviors include activities such as exceeding the speed limit, abruptly stopping, changing lanes erratically, and driving while distracted, all of which are well-established factors that contribute to road accidents and deaths. Researchers and corporations are using advanced algorithms and machine learning approaches to improve the precision and dependability of these detection systems. The primary goal of this research is to create real-time monitoring systems that can offer instant feedback to drivers. These systems provide drivers with timely information on their driving patterns, enabling them to make necessary changes, thus decreasing the probability of accidents and encouraging the adoption of safer driving habits. In addition, the use of real-time monitoring allows for the timely detection of potentially dangerous circumstances as they occur, enabling proactive intervention to avert accidents before they happen. The incorporation of artificial intelligence (AI) and machine learning algorithms signifies a notable progress in the domain of identifying hazardous driving behavior. These technologies have shown impressive abilities in analyzing large quantities of data and deriving valuable insights, thus improving the precision and efficiency of detecting systems. AI-powered algorithms can enhance their accuracy in detecting hazardous behaviors and differentiating them from typical driving activities by consistently acquiring knowledge from fresh data and adjusting to changing driving patterns. Furthermore, the possible uses of AI go beyond simple identification to proactive intervention and the prevention of accidents. AI algorithms can evaluate real-time driving habits and offer individualized advice to drivers to reduce harmful behaviors. Moreover, AI-powered predictive analytics can detect regions and scenarios with a high likelihood of risk, allowing authorities to execute focused interventions, such as heightened law enforcement or infrastructure enhancements, to enhance road safety on the whole [21]. Notwithstanding the encouraging progress in technology, there are still significant obstacles in the development and implementation of efficient systems for detecting risky driving behavior [22]. These encompass concerns with the confidentiality of data, partiality in algorithms, and the necessity for uniform evaluation measures to precisely gauge the effectiveness of detection algorithms. To overcome these problems, it is necessary for researchers, policymakers, and industry stakeholders to collaborate across disciplines. This collaboration is crucial to guarantee that technology is implemented in a responsible and ethical manner to improve road safety. Notably, the images in the DDB dataset were uniquely collected from dash cameras to encompass a varied range of driving scenarios, a novel approach. To bolster accuracy, we modified the VGG16 model and introduced a new model into our framework.

2.2 Literature Review

In the early stages of identifying distracted driving, rule-based systems and conventional computer vision techniques were prominently utilized. These systems relied on manually designed characteristics, such as erratic steering or lane deviations, to detect indications of driver preoccupation. While effective in discerning certain forms of distraction, these rule-based systems and computer vision techniques faced challenges in adapting to different driving scenarios. The inherent limitation of these methods prompted the development of more advanced and adaptable approaches. During the initial phases of tackling distracted driving, rule-based systems and traditional computer vision techniques were extensively utilized [23]. These technologies were specifically engineered to identify indications of driver distraction through the analysis of behaviors like as irregular steering or lane deviations. Although they were initially successful in detecting certain types of distraction, they faced challenges in adjusting to various driving conditions and emerging forms of distraction.

An inherent obstacle encountered by these initial approaches was their dependence on manually constructed rules and predetermined characteristics. These guidelines frequently encountered difficulties in accurately representing the intricate and ever-changing nature of distracted driving actions in different situations. Furthermore, traditional computer vision methods encountered challenges in effectively comprehending intricate visual data in real-time, particularly under demanding situations like unfavorable weather or fluctuating lighting conditions.

Upon acknowledging the limitations of these methods, researchers and engineers initiated an investigation into more advanced and flexible solutions. As a result, the field of driver monitoring systems was transformed by the emergence of sophisticated machine learning algorithms, specifically deep learning models. Through the utilization of extensive labeled data, these models have the capability to autonomously acquire and identify complex patterns and characteristics that indicate driver distraction, without requiring explicit rule-based programming. Furthermore, the progress in sensor technology, including the incorporation of several cameras, infrared sensors, and other physiological sensors, has facilitated a more thorough and resilient monitoring of driver behavior. These multimodal systems have the capability to collect a wider variety of information, such as facial expressions, eye movements, and physiological signals. This allows for a more comprehensive comprehension of driver attention and participation.

Ultimately, although rule-based systems and conventional computer vision techniques provided the initial framework for identifying distracted driving, their constraints required the creation of more flexible and advanced methods. By combining sophisticated machine learning algorithms with multimodal sensor technologies, contemporary driver monitoring systems have attained unparalleled levels of precision and adaptability in detecting and reducing the hazards linked to distracted driving.

To identify the characteristics of distracted drivers from photos and videos, including their body language and facial expressions, researchers had to resort to employing pre-trained algorithms. These pre-trained models have demonstrated efficacy in identifying a variety of visual signals associated with distracted driving, having been trained on extensive datasets [24].

Support vector machines (SVMs) and decision trees have been employed to classify driving behavior based on features extracted from in-car sensors. While these methods showed improvements over traditional approaches, their performance was limited by the need for manually engineered features. With an assessment accuracy of 92.24%, the Two-layer Neural Network model the paper [25] showed remarkable performance in distracted driver detection. When it comes to picture categorization, it has benefits over conventional machine learning methods.

Multiple Convolutional Neural Network (CNN) models and identified an optimal ensemble strategy by averaging the probabilities generated by VGG-16, VGG-19, and Inception V3. The culmination of these models resulted in a noteworthy log loss of 0.795, underscoring the effectiveness of their ensemble approach [26].

A deep learning method called Multiple Scale Faster-RCNN (MSFRCNN) for identifying a driver's cell phone use and hands position on the steering wheel, showing improved performance compared to Faster R-CNN, particularly in detecting hands on the wheel [27].

In this study, the authors present the Optimally-weighted Image-Pose Approach (OWIPA), which is an ensemble of ResNets designed for classifying distractions based on both original and pose estimation images. The methodology employs ResNet101 and ResNet50 for image classification, and an optimal weight is determined through grid search. The experimental findings reveal a notable accuracy of 94.28% on the AUC Distracted Driver Dataset [28].

The authors proposed employing deep learning in their Hybrid CNN Framework (HCF) to identify instances of inattentive driving. Concatenated features, fully linked layers, and a cooperative pretrained model were used in this framework to successfully filter out anomalies. There was an improved dropout method implemented to avoid overfitting. With an impressive categorization accuracy of 96.74%, the HCF may be able to help drivers develop safe driving practices. In the context of road safety, the study emphasizes how important it is to combat distracted driving [29].

Using deep learning architectures like VGG-16, GoogleNet, AlexNet, and ResNet, the authors provide a distraction detection system. Trials on a testbed for aided driving revealed improved performance when compared to baseline models.

The system, running on a Jetson TX1 board, showed an accuracy range of 86–92%, with an accuracy of 89% for the GoogleNet model. There is potential for reducing distracted driving accidents if this technology is put into actual cars [30].

The goal of this research was to create a CNN-based system that would detect and notify distracted drivers, which is an essential step in reducing the increasing number of traffic accidents. The performance of the VGG-16 architecture was greatly improved by using regularization techniques. The testing results demonstrated an excellent 96.31% accuracy and a GPU processing speed of 42 images per second, validating the efficacy of their technique. In addition, the improved system demonstrated a noteworthy 95.54% classification accuracy, demonstrating that optimization is feasible even with a reduced number of parameters.

In the conducted research, a pioneering CNN-based multi-model fusion approach for detecting distracted driving behavior in intelligent cockpit settings, referred to as CF-Net, was introduced. The model was constructed using classification networks, and an open-source dataset was employed for the analysis.

To enhance the model's performance, the XGBoost learning tool was integrated. The experimental results showcased improved detection capabilities, with a test accuracy of 93.45% and minimal verification loss. Within the field of vehicle safety, the identification of distracted driving behaviors has become a significant cause for concern, especially with the introduction of intelligent cockpit systems. The research undertaken in answer to this difficulty provided a novel solution called CF-Net, which utilizes Convolutional Neural Network (CNN)-based multi-model fusion approaches. CF-Net is a notable breakthrough in the industry since it combines classification networks designed expressly to identify distracted driving behaviors in a unique way. CF-Net utilizes many modalities to assess intricate connections between driver actions and environmental signals, resulting in enhanced detection capabilities. These findings underscored CF-Net as a significant advancement in the realm of intelligent cockpit environments.

TABLE I. KEY FINDINGS OF THE PREVIOUS LITERATURES

Author and year	**Title**	**Methods/Algorithms**	**Findings**
Moukafih [31] 2019	Aggressive Driving Detection Using Deep Learning-based Time Series Classification.	LSTM-FCN architecture.	In order to determine if a driving session contains aggressive conduct, the study presented a deep learning-based approach for classifying driver behavior. A time series categorization was used to formulate the issue. They obtained a 5-minute window length F-measure score of 95.88%.
Kim, Hyungil [32] 2020	Toward Real-Time Estimation of Driver Situation Awareness: An Eye-tracking Approach based on Moving Objects of Interest	four linear regression models and real time video-based driving simulation approach	In this paper, they proposed a technique to operationalize driver eye-movement data analysis based on moving objects of interest (OOI) for driver awareness of road hazards. Their models accounted for about 50% of the variance in the data.
Jabbar, Ratebateb. [33] 2020	Driver Drowsiness Detection Model Using Convolutional Neural Networks Techniques for Android Application.	CNN	This paper's main objective was to employ neural network-based techniques to identify such micro sleep and sleepiness. They utilized facial landmarks which are detected by the camera and that is passed to a Convolutional Neural Network (CNN) to classify drowsiness. They has achieved an average of 83.33% of accuracy for all categories data where the maximum size of the model did not exceed 75KB.90%.
Tran, Duy [34] 2020	Real-time detection of distracted driving using dual cameras.	VGG-16, ResNet, and MobileNet-v2	In order to identify distracted driving behaviors, the study provides a deep learning technique that uses a synchronized picture recognition system and achieves 96.7% identification accuracy at 8 frames per second.

A. Muhammad, T. Ali , M. Irfan, et al. [35] 2020	Driver Drowsiness Detection Model Using Convolutional Neural Networks Techniques for Android Application.	CNN Block I and CNN Block II	They obtained an F-measure score of 95.88% within a 5-minute time frame. Dlib was found to be 80.25% for eye recognition accuracy and 78.60% for sleepy accuracy.
Tamagusko, Tiago, [36] 2022.	Deep Learning applied to Road Accident Detection with Transfer Learning and Synthetic Images	MobileNetV2 and EfficientNetB1	In contrast to the model based on MobileNetV2, which took 38 seconds to obtain a mAP of 0.87 and an MCC of 0.68, the EfficientNetB1 model took 72 seconds to obtain these values.
Vitorino, João, [37] 2023	Adversarial Robustness and Feature Impact Analysis for Driver Drowsiness Detection	KNN, XGB and SVM	XGB 150 distinguished itself by maintaining an F1-Score of 79.19% during the assault and effectively identifying each produced example. In addition to XGB 150.Even though it was still fooled by a respectable amount of cases and saw a decline in its score to 72.62%, KNN 180 trained with adversarial training managed to achieve an F1-Score of 83.33% on the initial holdout set.
F. Sajid, A. R. Javed [38] 2023	Driver Drowsiness Detection System	OPENCV, SVM	Creating a temporary condition monitoring system for autos throughout time was the major goal of this work. For both eye detection and tiredness, Dlib's average real-time test accuracy was found to be 80.25% and 78.60%, respectively.

The key findings of the earlier literature are summarized in Table I. Every new article makes room for further research. We were able to assess all of this data and decide whether or not our system and research were feasible.

2.4 Summary

Reviewed several current papers that were pertinent to this work. Our Literature Review chapter includes a review of a driving behavior detection-based work. We studied the terminology carefully and used them in our work, as well as reviewing some relevant earlier work that inspired and encouraged us to work through our approach. The Literature Review chapter provides a crucial framework for comprehending the development of driving behavior detection and emphasizes the originality and significance of our study in this field.

Chapter 3
Methodology

3.1 Introduction

This section includes a summary of the main findings and research methodology, along with brief reports on the actions of distracted drivers. The process of training the algorithm to reliably identify distracted driving behaviors based on the gathered data is also covered, along with algorithm fine-tuning.

3.2 Data Collection and Data Set

To implement the deep learning model, two distinct datasets were utilized. The initial dataset originated from our proprietary driving video recording system, encompassing diverse driving scenarios captured from consistent angles, while the second dataset was sourced from Kaggle [15] called State Farm's distracted driver detection dataset, a publicly accessible repository. An exhaustive examination was undertaken initially to identify the fundamental causes of vehicular collisions, involving an in-depth analysis of factors such as driver conduct, road conditions, and vehicle maintenance. The investigation pinpointed driver behavior as the primary catalyst for crashes, notably instances of distraction. Subsequently, during the data collection phase, a Kaggle dataset featuring ten categories of distracted driver behavior emerged, offering valuable insights into prevalent distractions such as texting, phone conversations, and drinking while driving. In addition to the Kaggle dataset [15], we generated an independent primary dataset resembling the Kaggle counterpart. This involved the recording of distracted driving scenarios through dashcam videos, allowing for the extraction of frame-by-frame data. This approach facilitated a granular analysis of specific moments of distraction and their discernible impact on driving behavior. The deployment of the deep learning model for detecting distracted drivers involves a comprehensive strategy that combines sophisticated machine learning methods with domain expertise and ethical considerations. This aims to improve road safety and decrease the occurrence of car accidents caused by driver distraction.

The following (Figure 1 and Figure 2) presents exemplar data for each class sourced from both datasets.

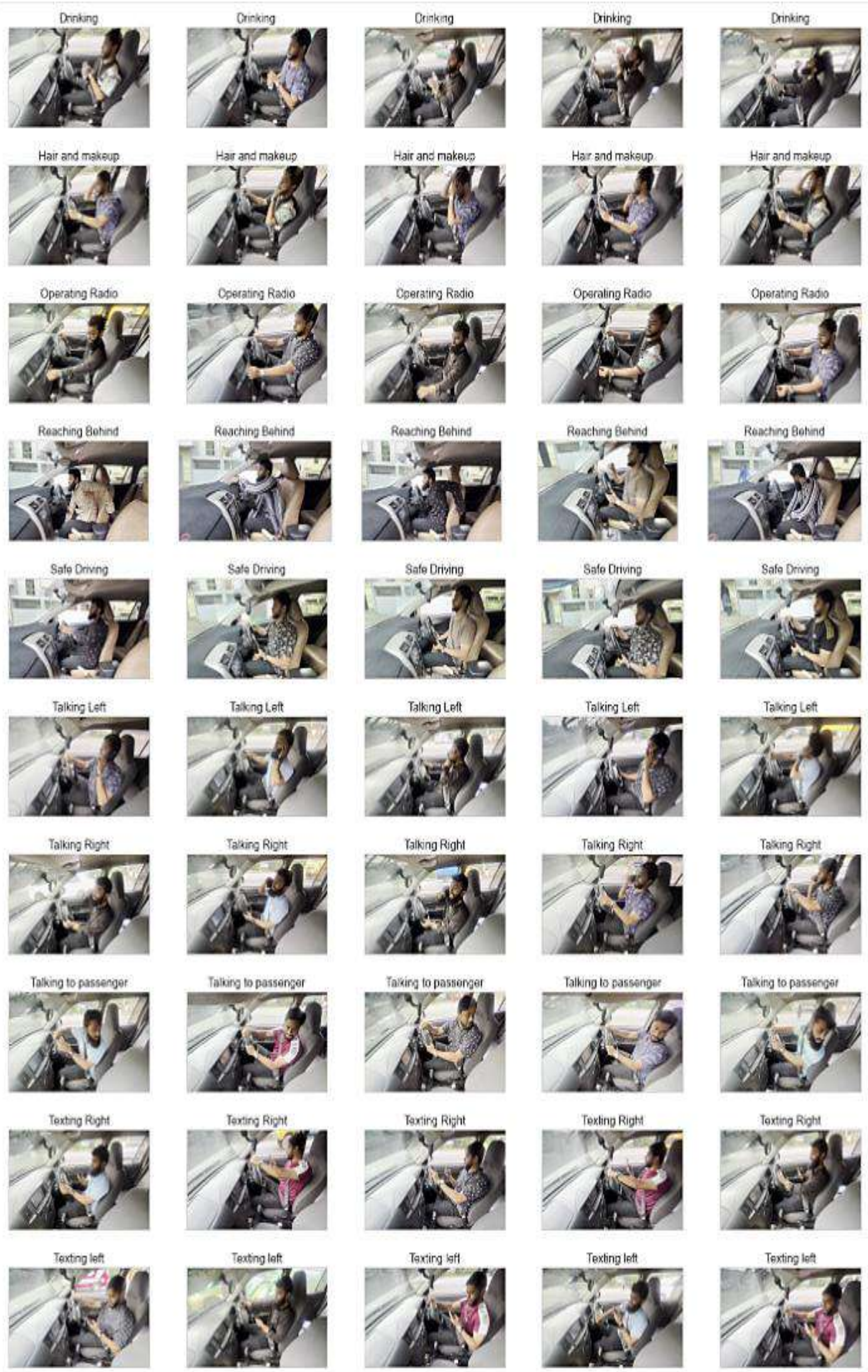

Figure 1 Sample Images from each label in our DDB Dataset.

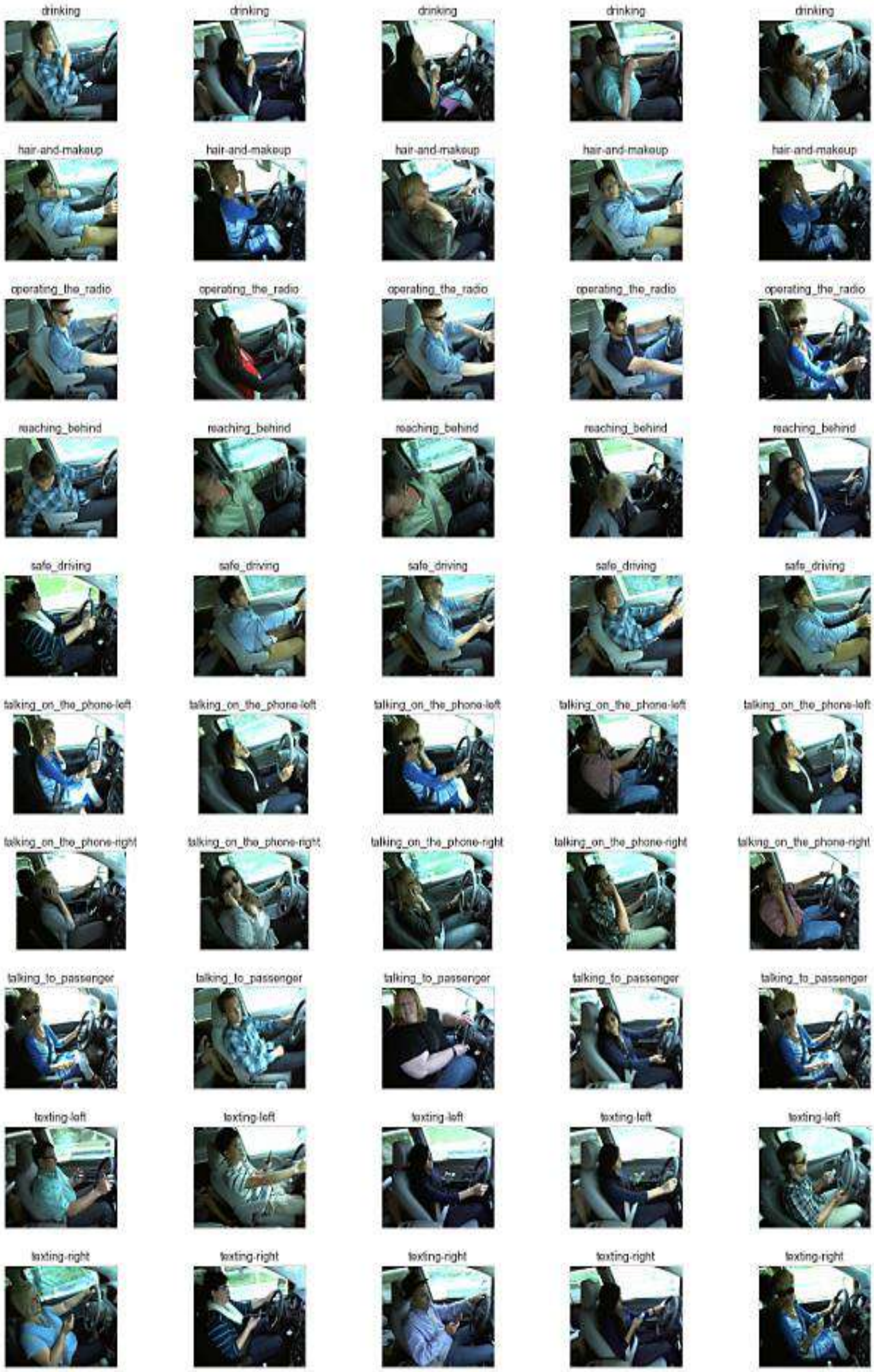

Figure 2 Sample Images from each label in Kaggle Dataset [8].

3.3 Proposed Methodology

This section details data collection, labeling, deep learning model design, and system training. The data collection process involves capturing images and videos of distracted driving in various scenarios, such as texting, talking on the phone, or eating while driving. These images are then labeled as with the corresponding Kaggle dataset [8] to create a comprehensive dataset. The deep learning model design includes fine-tuning the VGG-16 architecture to adapt it specifically for detecting distracted driver behaviors. System training involves training the modified VGG-16 classifier using the labeled dataset to improve its accuracy in predicting different types of distracted driver behaviors. A fuller explanation and visual representation of the whole distracted driver behavior detection system are provided by the work flow diagram that follows.

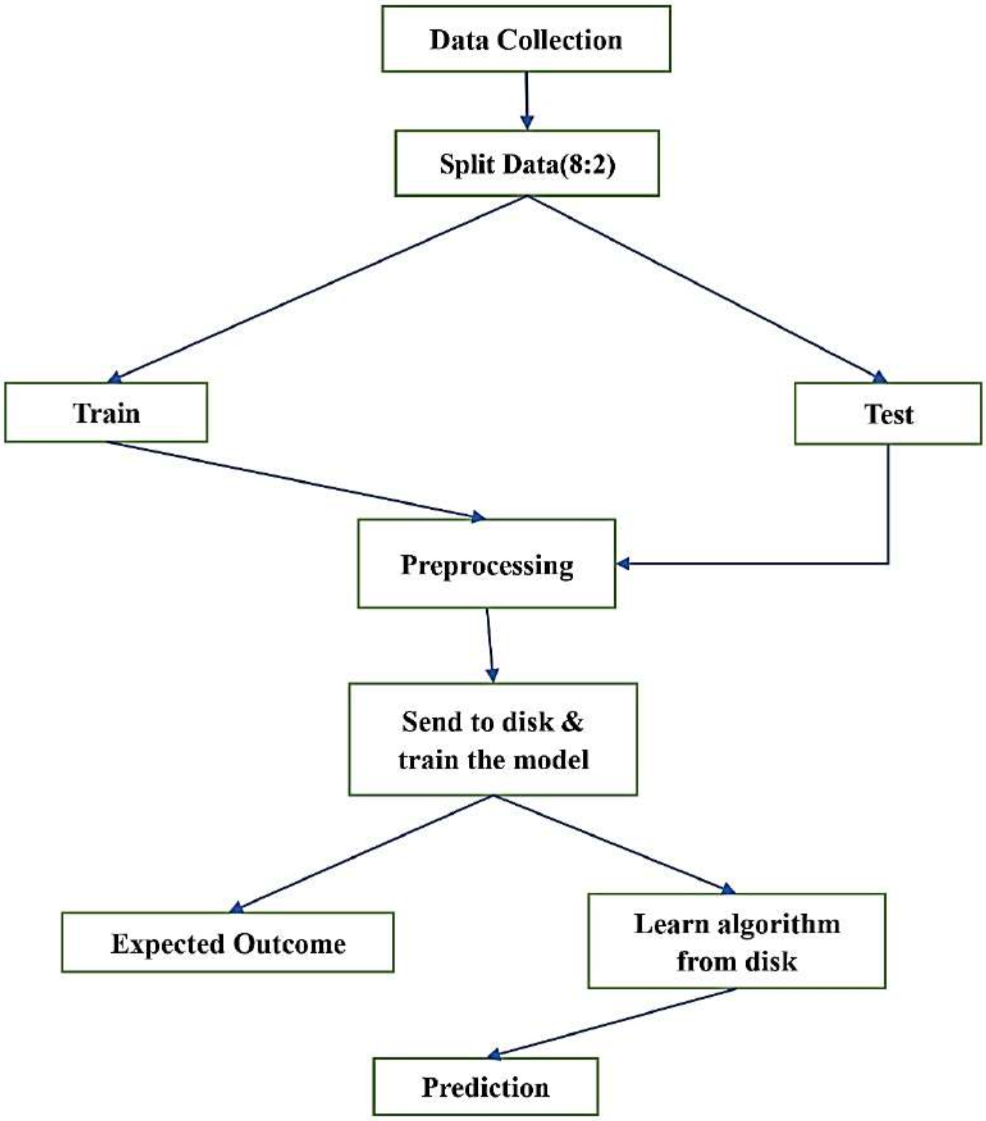

Figure 3 Flow Chart of proposed CNN Model

3.3.1 Statistical Analysis of Datasets

For this research, two different datasets were used. The first dataset, which we will refer to as the DDB Dataset, was created by taking pictures of our dashcam. Ten categories were created from a total of 4941 photos, which were then split into two folders: one for the training set (Figure 4) with 3747 images and another for the testing set (Figure 5) with 1194 images.

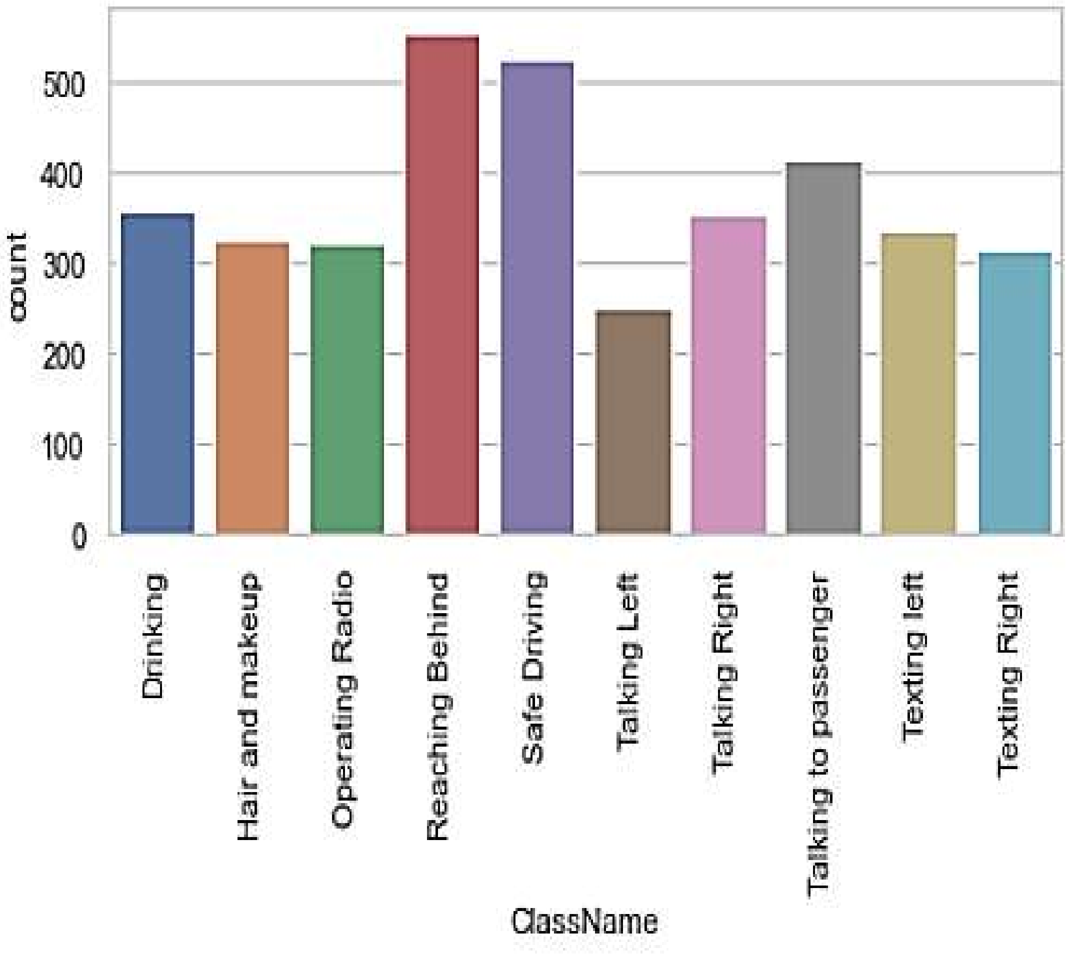

Figure 4 Statistic visualization of training data of our DDB Dataset.

This partition guarantees that the model can be trained on a significant amount of the data while also being assessed on unknown data to evaluate its capacity to generalize and its performance. Researchers intend to utilize this dataset, along with its training and testing subsets, to create and verify machine learning or deep learning algorithms that can effectively identify and categorize different driving behaviors. These algorithms show potential for improving driver assistance systems, autonomous vehicle technologies, and overall road safety.

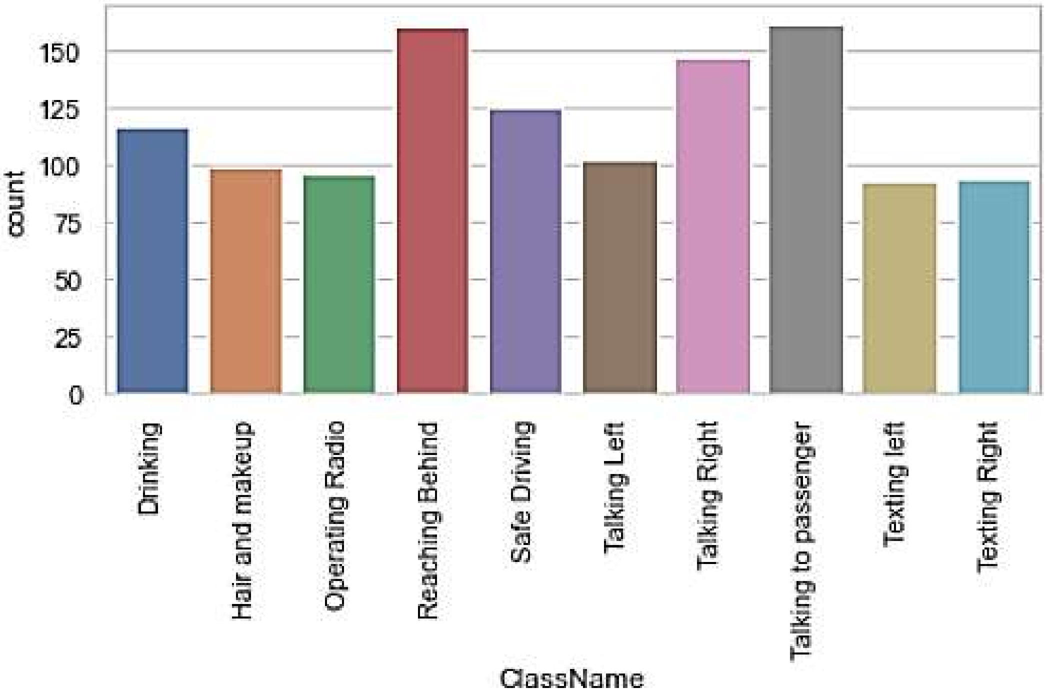

Figure 5 Statistic visualization of testing data of our DDB Dataset.

The State Farm Distracted Driver Detection (SF3D) dataset [15], the second dataset, was made publicly accessible on Kaggle in 2016 [15]. The 640 × 480 RGB images in this collection were captured with a 2D dashboard camera that was fixed in place. The training set of this dataset has 22424 data points. Research and development of algorithms for detecting distracted drivers frequently make use of the SF3D dataset.

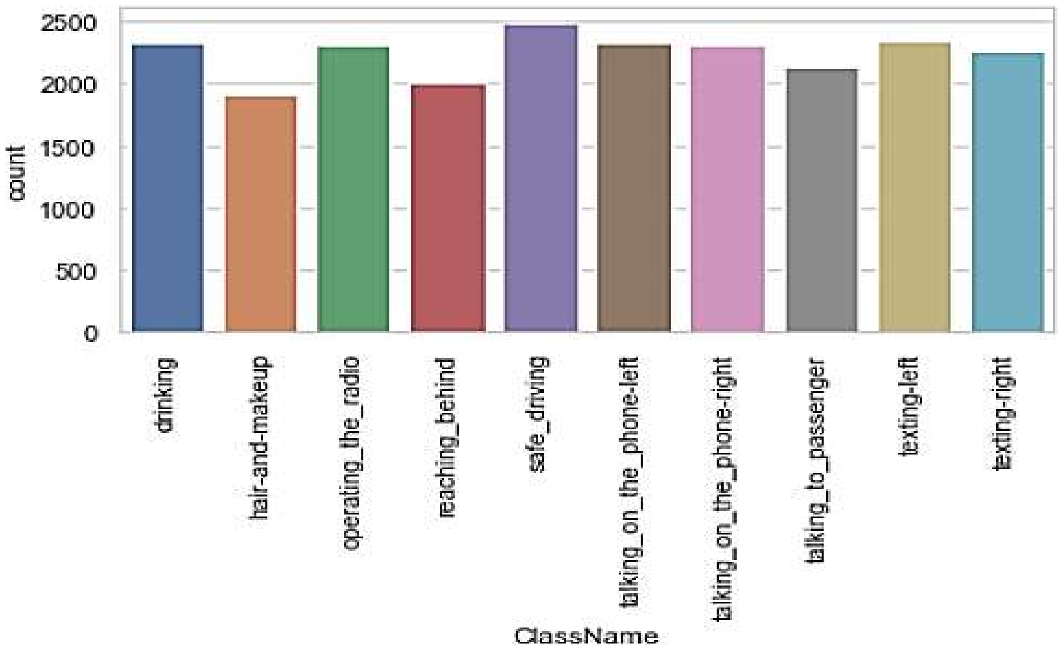

Figure 6 Statistic visualization of training data of Kaggle Dataset [8].

3.3.2 Data Pre-Processing

Following the acquisition of the data, we categorized the collected information into ten classes based on their respective categories. We undertook a comprehensive data preprocessing phase with the objective of refining and adapting the gathered data into a format conducive to analysis. This process encompassed the removal of any missing values, the standardization of numerical variables, and the conversion of categorical variables into numerical representations through encoding techniques. We implemented feature scaling to homogenize the scale of all variables, ensuring uniformity and precision in subsequent analysis and modeling endeavors. We also conducted feature scaling to ensure all variables were on a similar scale for accurate analysis and modeling. Here data preprocessing scheme showing below:

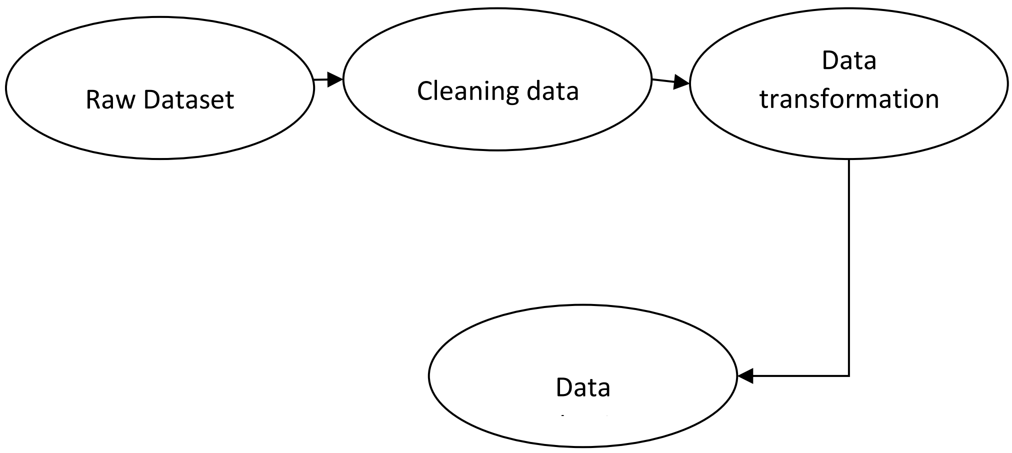

Figure 7 Data preprocessing stage in our model.

Figure 8 exhibits a compilation of sample photos following the preprocessing phase in our approach. Preprocessing is an essential stage in machine learning and computer vision pipelines. It encompasses a range of modifications that are applied to raw input data to enhance its applicability for future analysis and model training. The preprocessing stage usually involves a variety of actions customized to the individual needs of the dataset and the goals of the model. Every one of these preprocessing procedures is essential for improving the quality, consistency, and relevance of the data utilized for model training and analysis. In the preprocessing stage, many operations are commonly carried out to ready the data for subsequent analysis and model training. These actions are customized to tackle unique issues inherent in the dataset and to enhance the efficiency of the machine learning or computer vision model being constructed.

Figure 8 Sample Images after preprocessing stage in our model

3.3.3 Modified VGG16 Architecture

VGG-16 comprises a total of 13 convolutional layers and 3 dense layers. The architecture of VGG-16 integrates interleaved convolutional and pooling layers, forming the basis for the CNN. The model incorporates five convolutional layers followed by two fully connected layers. The convolutional layers utilize a filter size of (3,3), while the pooling layer adopts a size of (2,2) to scale down feature maps, reducing computational costs and parameter counts while preserving crucial information. The convolutional layers employ a variable number of filters, ranging from 64 to 512. Except for the first eight layers, set as trainable, the remaining layers are configured as untrainable. This approach facilitates fine-tuning of the model while retaining lower-level features. Specifically, eight convolutional layers are maintained in a frozen state, and the remaining layer is replaced with a sequential model consisting of seven new layers: a GlobalAverage pooling layer, a batch normalization layer, dense layers with 256 neurons each, connected to ReLU activation layers, two dropout layers with a 0.2 rate for preventing overfitting and ensuring model accuracy, and a final output layer utilizing a softmax activation function. By tailoring the structure to meet the specific demands of the task, professionals can create deep learning models that are both more efficient and more successful in various computer vision applications.

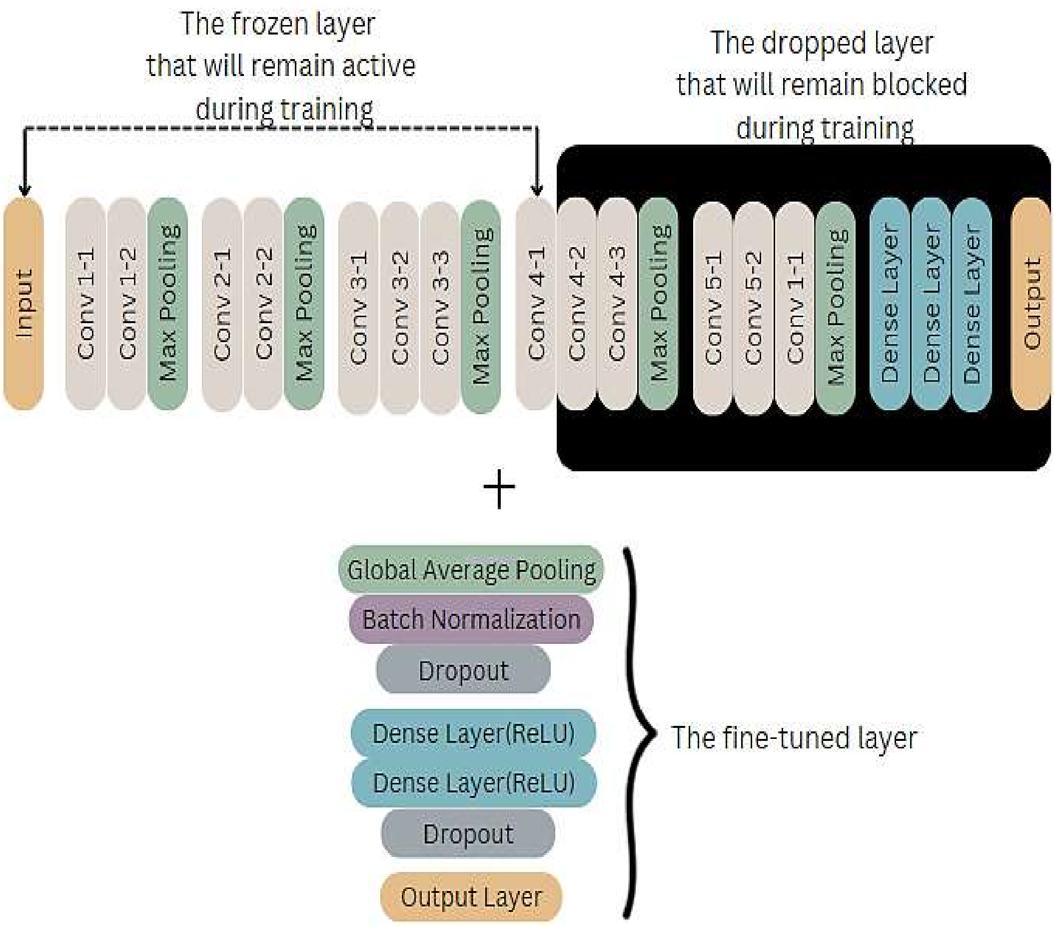

Figure 9 Modified VGG-16 Architecture

Figure 9 depicts a visual representation of the altered VGG16 model. The modification probably involves making changes to the original VGG16 design in order to better meet the specific requirements of the current assignment. This may entail modifying the layer count, adjusting the size of convolutional filters, or introducing new layers to improve the model's performance. Figure 9 presents a comprehensive illustration of the architecture's composition, showcasing the organization of convolutional layers, pooling layers, and fully connected layers. Each layer is commonly depicted as a block, with connections between layers illustrating the movement of data inside the network. Figure 9 provides researchers and practitioners with valuable information on the architectural improvements used in the VGG16 model. Analyzing these changes can help understand how they enhance the model's functionality and efficacy. Figure 9 and Figure 10 provide visual assistance in understanding the architecture and layout of the updated VGG16 model. This aids in the analysis, interpretation, and optimization of the neural network for the specific task or application.

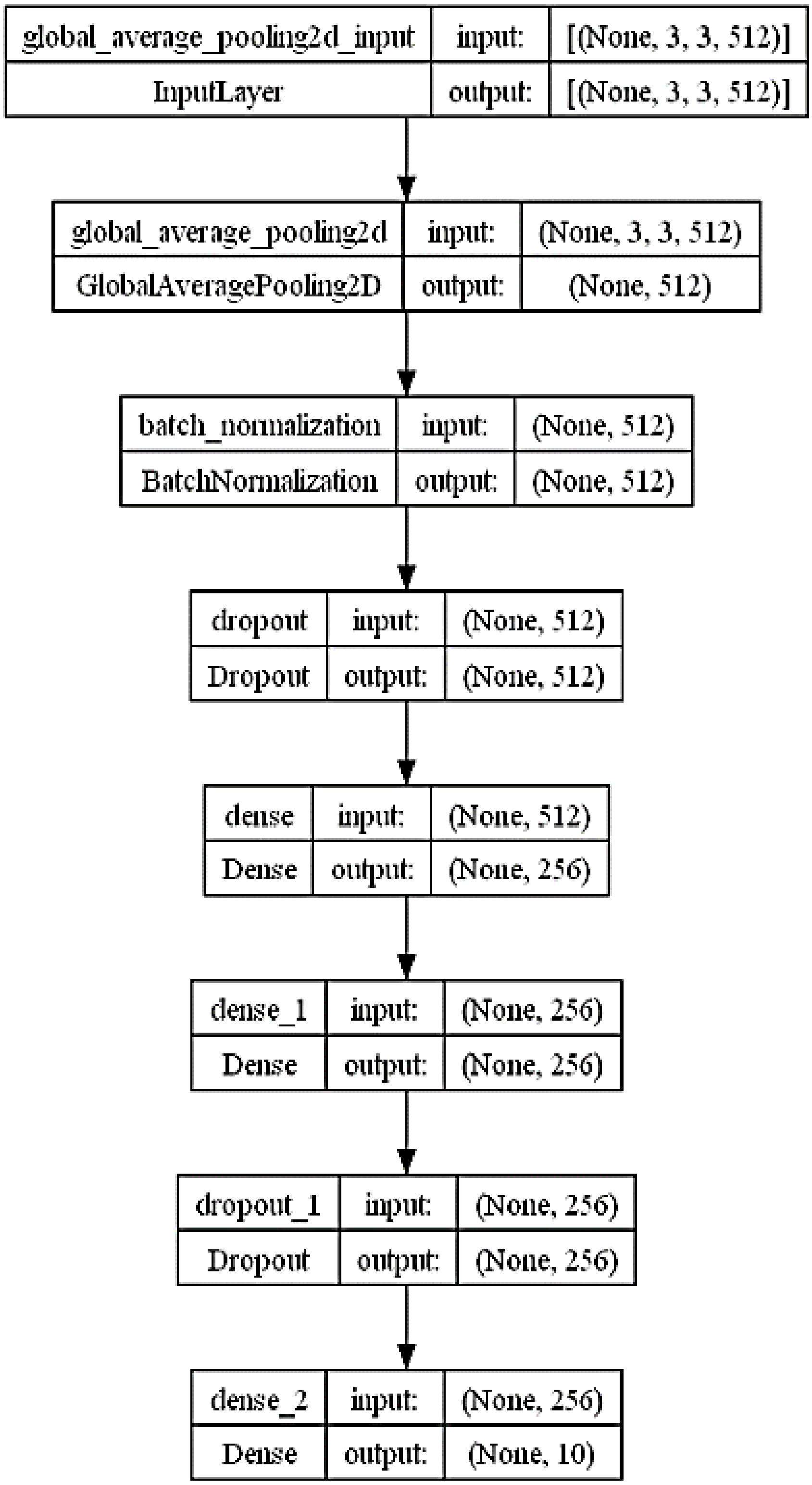

Figure 10 Modified VGG-16 Architecture

Figure 9 and Figure 10 shows the modified VGG16 model. The summary of the module, or the layers added to the neural network for effective processing, is shown in Figure 11. In order to send the values of this one-dimensional array as input to the neural network, a flatten layer is added to the network.

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 global_average_pooling2d (G  (None, 512)              0
 lobalAveragePooling2D)

 batch_normalization (BatchN  (None, 512)              2048
 ormalization)

 dropout (Dropout)           (None, 512)               0

 dense (Dense)               (None, 256)               131328

 dense_1 (Dense)             (None, 256)               65792

 dropout_1 (Dropout)         (None, 256)               0

 dense_2 (Dense)             (None, 10)                2570

=================================================================
Total params: 201,738
Trainable params: 200,714
Non-trainable params: 1,024
_________________________________________________________________
```

Figure 11 Modified VGG16 Model Summary

3.3.4 Model Training Process

In the course of this training process, a transfer learning approach was employed to expedite and enhance the accuracy of predictions within a short duration, replacing the conventional method of constructing a CNN model from the ground up. VGG16, an already pre-trained convolutional neural network (CNN) on ImageNet, was adapted as a feature extractor. The initial 8 layers of this pre-trained model were frozen, and the remaining layers were appended to the end of the CNN. Subsequently, the surplus layers of VGG16 were excised, and a Global Average Pooling layer was affixed, followed by batch normalization and a dropout layer tailored to the specific task. These parameters were fine-tuned using gradient descent with the categorical cross-entropy loss function. The fully connected layer comprised two dense layers, each housing 256 neurons with a rectified linear unit (ReLU) activation function. The hyperparameter for the dropout layer was set at 0.2. A final dense layer, accommodating the requisite number of classes, was introduced, with softmax activation employed to normalize the output layer values into probabilities for the ten classes. The model was configured with a 100x100 input size, 3 channels, and a random state of 42. The ensuing training process ensued following these architectural adjustments.

3.4 Summary

The Deep Convolutional Neural Network (CNN) VGG16 is distinguished by pre-trained layers that exhibit a robust understanding of the intricate features present in images, encompassing aspects such as shape, color, and structure. This deep architecture, particularly the extensively layered VGG16, has undergone training on an extensive array of diverse images, equipping it with a formidable capability for intricate classification tasks. The selection of VGG16 for our application is grounded in its proven track record of high performance across challenging classification scenarios. This choice is especially pertinent as we are dealing with two datasets that comprise distinct types of images, and the utilization of the VGG16 model is instrumental in achieving accurate and reliable validation results. The decision to employ VGG16, a deep Convolutional Neural Network (CNN), in our thesis is based on its remarkable ability to accurately recognize and classify images. The architecture of the model consists of pre-trained layers, which have undergone thorough training on a wide range of images. This training has enhanced the model's capability to accurately identify complex aspects such as shape, color, and structure. In addition, the several layers of VGG16 contribute to a profound comprehension of image representations, rendering it a favored option for addressing the inherent intricacy of our datasets. As we explore picture categorization, the strong and dependable nature of VGG16 plays a crucial role in guaranteeing precise validation results for different sorts of images. This enhances the credibility and trustworthiness of our research findings. In addition, utilizing the established capabilities of VGG16 not only simplifies our methodology but also places our work within the wider framework of cutting-edge deep learning approaches, highlighting the meticulousness and complexity of our thesis endeavors. Through the utilization of transfer learning, we leverage the information stored in VGG16's weights to accelerate the convergence of the model and improve its efficiency. Moreover, the adaptability of VGG16 extends beyond the limitations of our particular datasets, providing the possibility of being useful for a wide range of image analysis jobs in different fields.

Chapter 4
Results and Discussion

4.1 Introduction

The comprehensive evaluation of the presented model relies on conducting experiments centered around predicting categories of Distracted Driver behavior. This assessment is performed on two datasets through the utilization of a modified VGG16 model, as outlined in the preceding chapter.

4.2 Programming Language and Necessary Dependencies

Python 3.9 was used as the programming language due to its extensive libraries. TensorFlow 2.10.0 was the core library used for the tasks. The model construction was facilitated through the Keras Sequential API, and essential data manipulation and analysis were performed with the assistance of NumPy, Scikit-Learn, and Pandas. All dependencies were carefully sourced from the Python Package Index (PyPI), ensuring that we incorporated the latest versions to leverage the most recent advancements and bug fixes. To visualize and interpret our data effectively, Matplotlib was employed for general data visualization, while Seaborn was utilized for more specialized and statistical visualizations.

4.3 Experimental Analysis

The fine-tuned VGG16 model was employed to train and validate its accuracy and performance in predicting distracted driver behavior categories using two datasets characterized by imbalanced classes. The first dataset named DDB Dataset, comprises images taken from various angles, while the second dataset includes additional data sourced from diverse drivers and vehicles. The primary goal of applying these datasets to the adapted VGG16 model is to validate the effectiveness and resilience of our approach in accurately categorizing distracted driver behavior. In comparison with existing state-of-the-art models, our modified VGG16 model demonstrated superior accuracy and robustness. Extensive experiments and analyses were conducted to ensure that our model is impartial to any specific driver or vehicle type, enhancing its reliability for real-world applications. The model was trained using a specific range of epoch counts to achieve optimal performance in accurately categorizing distracted driver behavior. The ideal number of epochs for all optimizations was 100, with higher accuracy attained around 30 epochs. This indicates the model's efficiency for real-time applications, as it does not necessitate prolonged training periods. Additionally, we conducted comparisons with existing literature to validate the model's performance on the Kaggle dataset, demonstrating consistent accuracy across various scenarios and affirming its reliability and generalizability. For training the model, optimization algorithms such as Adam, SGD, RMSProp, and AdaGrad were employed. Each algorithm underwent separate testing, and their performance was assessed based on accuracy and convergence speed. The results indicated that Adam and SGD consistently outperformed RMSProp and AdaGrad in both accuracy and convergence speed, establishing them as the preferred choices for optimizing our model. The model's performance was further assessed using the confusion matrix, revealing insights into specific error types. It showed a higher inclination to misclassify certain classes, signaling potential areas for improvement in subsequent iterations. Precision, recall, and F1-score were computed to offer a more thorough evaluation, considering both false positives and false negatives. These metrics affirmed the overall efficacy of the model while pinpointing particular areas for attention. The calculation of precision, recall, and F1-score provide a thorough assessment of the model's performance, providing valuable insights beyond basic accuracy measurements. Precision quantifies the ratio of accurate positive predictions to all positive predictions, reflecting the model's capacity to minimize false positives.

Precision is defined as the ratio of true positive predictions to the total instances predicted as positive, expressed as:

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \quad \text{.......................................(1)}$$

Recall, on the other hand, assesses the accuracy of positive instances among actual positives, calculated as:

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \text{..........................(2)}$$

The **F1-score**, a blend of precision and recall, is given by:

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \ldots\ldots\ldots\ldots\ldots\ldots\ldots (3)$$

These metrics facilitated a nuanced understanding of the model's strengths and weaknesses, guiding targeted enhancements in future iterations.

4.4 Result Analysis of DDB Dataset

The performance of our model was evaluated using four different optimization algorithms: Adam, SGD, RMSProp, and Adagrad. Each algorithm was applied to train the model and measure its performance. The results obtained from these experiments provide valuable insights into the effectiveness of each optimization algorithm in improving the model's performance. We have created a confusion matrix to evaluate the model's performance for each optimization algorithm. The confusion matrix made it possible to analyze the accuracy, precision, recall, and F1 score of the model for different classes. By comparing the outcomes of every optimization algorithm, we found which one yields the best overall performance and identified any potential trade-offs or limitations of each algorithm.

4.4.1 Performance Analysis of ADAM Optimizer in DDB Dataset

The performance of Adam is shown in Figure 12 The data showed that Adam's performance improved throughout a 100-period training plot. The figure displayed a significant decrease in the loss function, indicating that the model was effectively acquiring and optimizing its parameters. Upon analyzing the effectiveness of the ADAM optimizer on the DDB dataset, it is imperative to further investigate the ramifications of these discoveries. In addition to acknowledging the reduction in the loss function, it is crucial to examine the effect of this optimization on the model's predictive ability. An in-depth analysis of the subtle variations observed in the 100-period training plot can provide valuable insights into the rate at which the model converges and its overall stability. Furthermore, investigating any possible obstacles or constraints faced throughout the optimization process can offer significant insights for future research attempts. Moreover, doing comparative evaluations with various optimization approaches could provide a broader viewpoint on the effectiveness of ADAM in this specific scenario. Examining the practical consequences of these improvements in actual applications helps underscore the importance of such optimizations in advancing machine learning methodology. By doing a full analysis of these factors, the conversation can be enhanced, providing a more thorough comprehension of the success of the ADAM optimizer in optimizing parameters inside the DDB dataset.

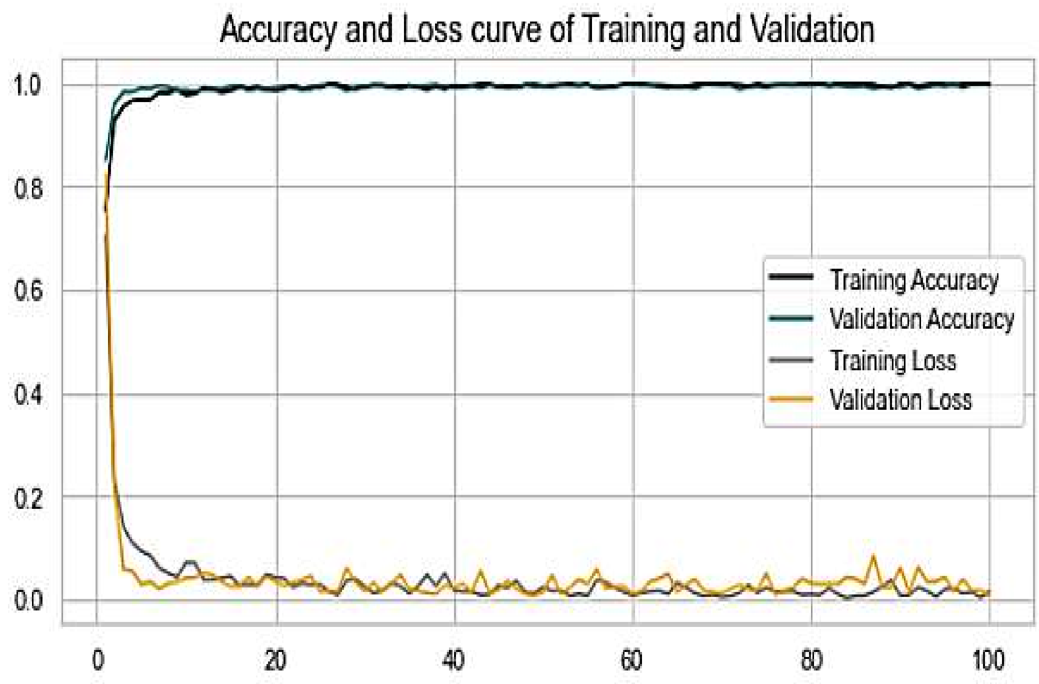

Figure 12 Accuracy and Loss curve of Adam Optimizer of our model in DDB Dataset

The details of Adam's training parameter are shown in Table II below.

TABLE II. FOR "ADAM" OPTIMIZER PARAMETER VALUE AND HYPER PARAMETER

Hyper parameter	**Parameter value**
Number of Epoch	100
Batch Size	16
Optimizer	Adam
Objective function	Categorical Crossentropy
Hidden layer	ReLU
Output layer	Softmax

According to the confusion matrix shown in Figure 13, the ADAM Optimizer in our model achieved 100% accuracy for almost all classes, except for one class where it achieves 98% accuracy. This indicates that the ADAM Optimizer is highly effective in accurately classifying the majority of classes in our dataset.

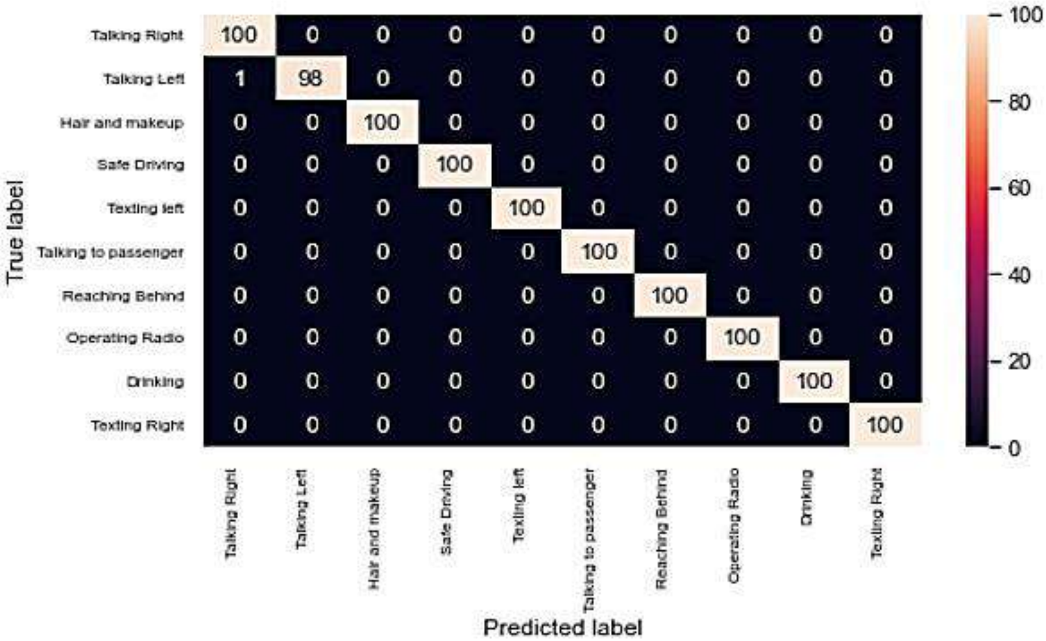

Figure 13 Confusion Matrix of Adam Optimizer in DDB Dataset

4.4.2 Performance Analysis of SGD Optimizer in DDB Dataset

The Figure 14 training plot shows a steady decrease in loss over time, indicating the model is learning and improving its predictions, and a decrease in validation loss, indicating no overfitting.

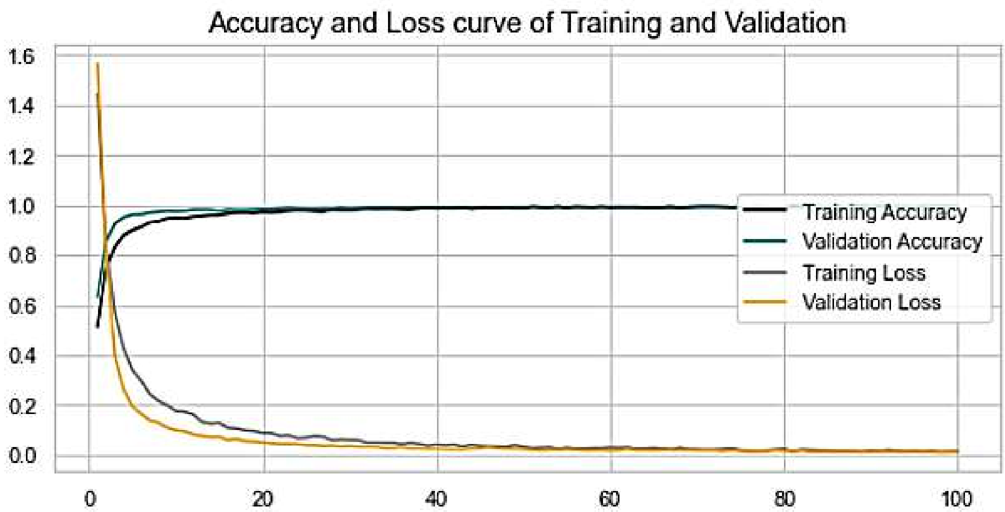

Figure 14 Accuracy and Loss curve of SGD Optimizer of our model in DDB Dataset

The training parameter details for SGD Optimizer are displayed in Table III below.

TABLE III. FOR "SGD" OPTIMIZER PARAMETER VALUE AND HYPER PARAMETER

Hyper parameter	Parameter value
Number of Epoch	100
Batch Size	16
Optimizer	SGD
Objective function	Categorical Crossentropy
Hidden layer	ReLU
Output layer	Softmax

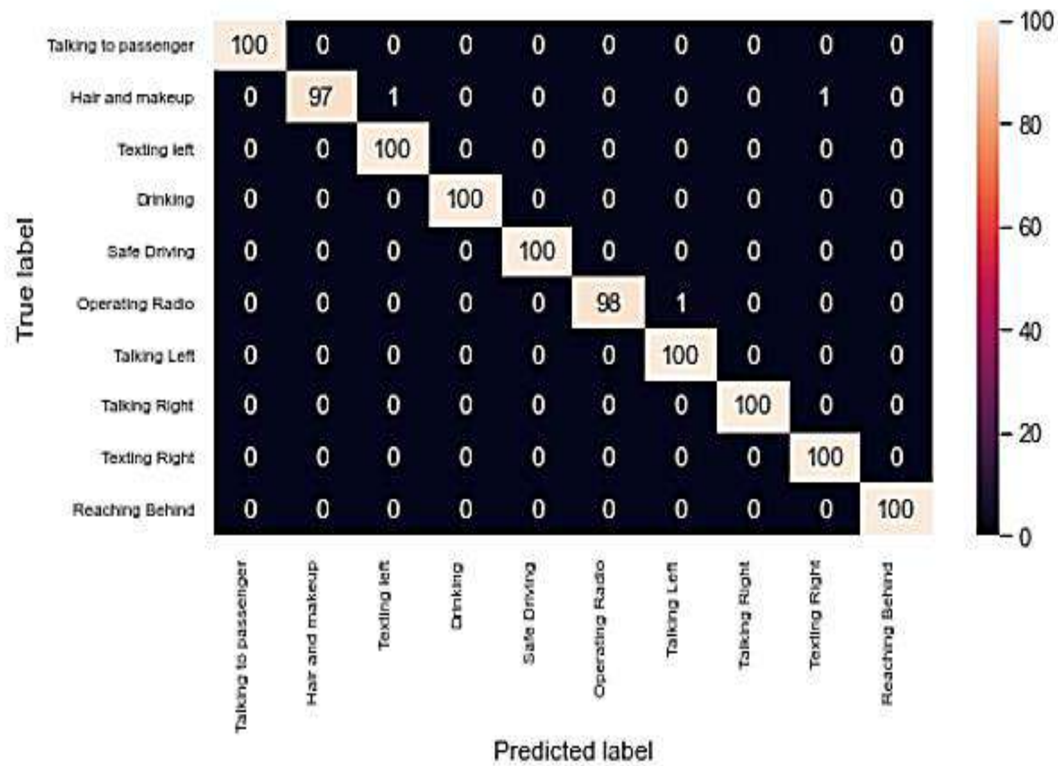

Figure 15 Confusion Matrix of SGD Optimizer in DDB Dataset

From the Figure 15 confusion matrix analysis, we observed that the model performed exceptionally well in correctly classifying most of the samples. Employing the Stochastic Gradient Descent (SGD) optimizer on our dataset, we got a decent accuracy rate of 99.60%. However, there were a few instances where the model misclassified some samples, resulting in a small number of false positives and false negatives.

4.4.3 Performance Analysis of RMSPROP Optimizer in DDB Dataset

In areas with flat losses, RMSProp permits faster mobility while attenuating oscillations on steep grades. RMSProp does not accumulate previous squared gradients; instead, it uses an exponential decay. This approach helps to prevent the learning rate from becoming too large and causing instability. The use of exponential decay ensures that older gradients have less influence on the current update, making the algorithm more adaptive to changing conditions. Figure 16 below shows that as the number of epochs increased, the accuracy increased gradually in the starting epochs, and then it was stable with better accuracy, but for a certain period the loss was decreasing, then the loss started to unstable but it didn't impact the overall accuracy. This indicates that the optimization algorithm was able to find a good balance between updating the model parameters and avoiding overfitting. It is crucial to investigate the underlying mechanisms that are responsible for the observed steady accuracy and diminishing loss. Examining the convergence characteristics of the optimization technique could yield significant insights. It would be beneficial to explore the impact of the algorithm's learning rate, batch size, and regularization techniques on the observed behavior.

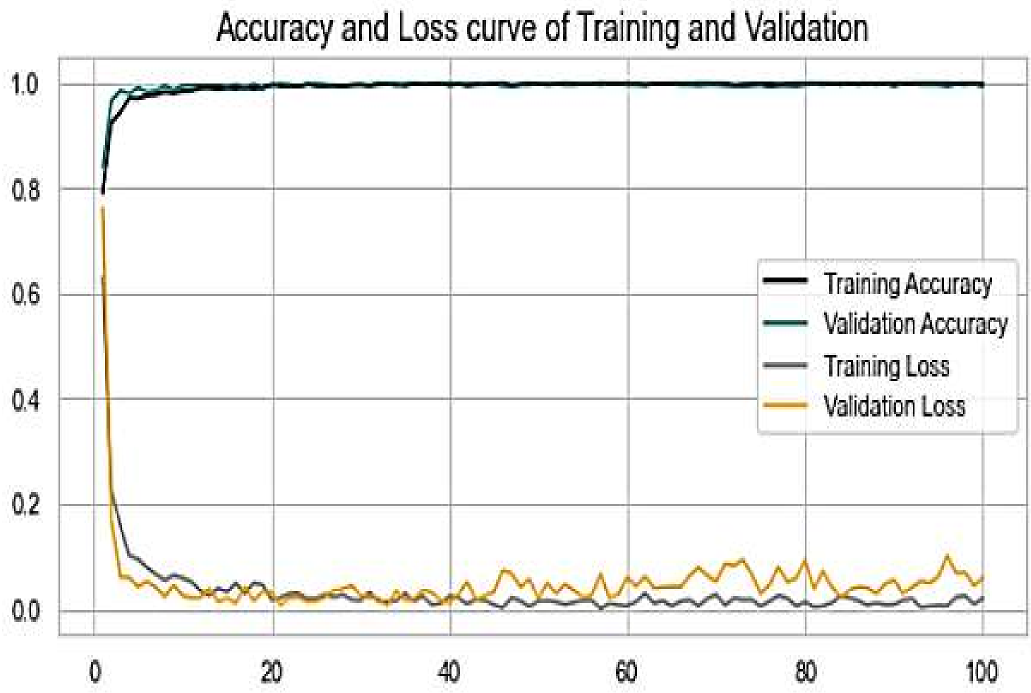

Figure 16 Accuracy and Loss curve of RMSProp Optimizer of our model in DDB Dataset

TABLE IV. FOR "RMSPROP" OPTIMIZER PARAMETER VALUE AND HYPER PARAMETER

Hyper parameter	Parameter value
Number of Epoch	100
Batch Size	16
Optimizer	RMSProp
Objective function	Categorical Crossentropy
Hidden layer	ReLU
Output layer	Softmax

Table IV contains information on the RMSProp optimizer parameter. And the Figure 17 below shows the result of confusion matrix.

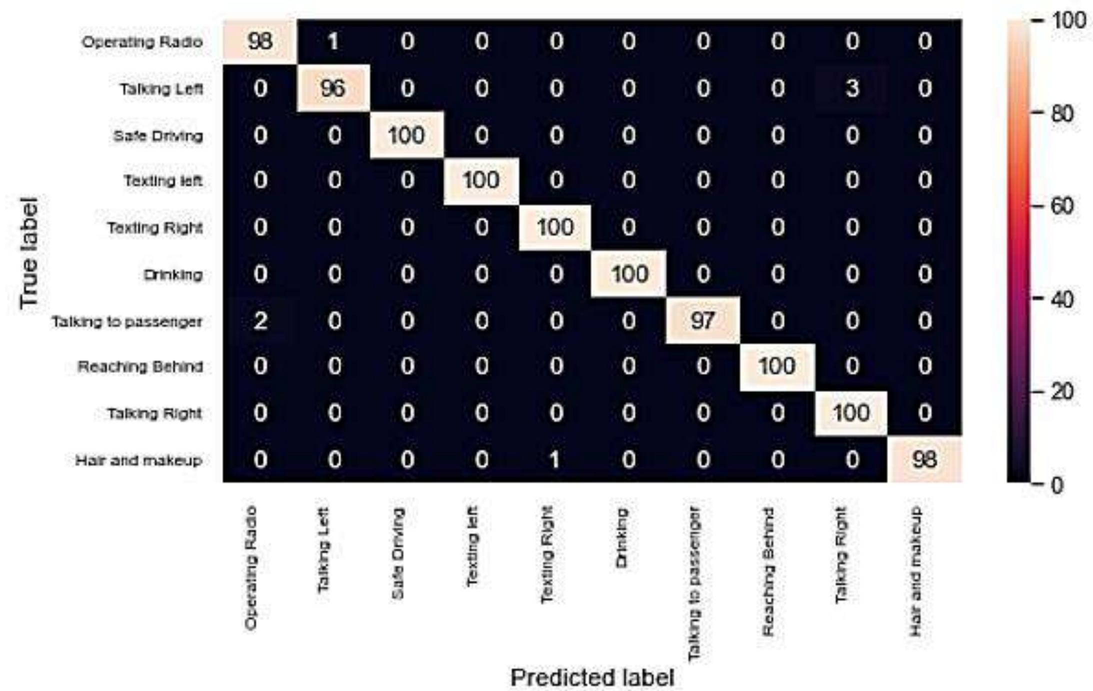

Figure 17 Confusion Matrix of RMSProp Optimizer in DDB Dataset

4.4.4 Performance Analysis of ADAGRAD Optimizer in DDB Dataset

AdaGrad is a sub-gradient stochastic optimization algorithm family that adjusts learning rates based on historical gradients, with smaller gradient parameters being prioritized for better convergence and performance. In Figure 18 below shows how, as the number of epochs increases, accuracy rises and loss falls.

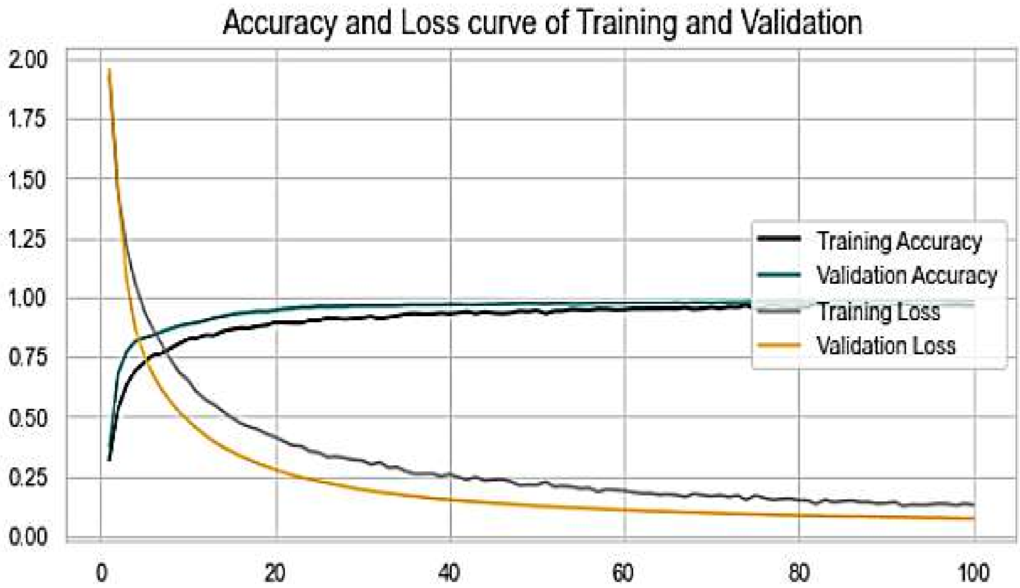

Figure 18 Accuracy and Loss curve of AdaGrad Optimizer of our model in DDB Dataset

TABLE V. **F**OR “**A**DAGRAD” **O**PTIMIZER **P**ARAMETER **V**ALUE **A**ND **H**YPER **P**ARAMETER

Hyper parameter	**Parameter value**
Number of Epoch	100
Batch Size	16
Optimizer	AdaGrad
Objective function	Categorical Crossentropy
Hidden layer	ReLu
Output layer	Softmax

Table V provides information on the AdaGrad optimizer parameter. When compared to other optimization algorithms, AdaGrad had a slow convergence rate. The provided confusion matrix in Figure 19 supports the observation. Once the slow convergence rate of AdaGrad has been established in comparison to other optimization techniques, it is crucial to examine the ramifications of this fact. Firstly, analyze the probable consequences on the training and performance of the model. A sluggish rate of convergence may need extended training durations, augmented processing resources, or more advanced methodologies to address the problem. Furthermore, investigate the factors contributing to AdaGrad's reduced convergence speed, including its adaptive learning rate mechanism and its treatment of sparse gradients. This analysis could yield valuable insights regarding the applicability of AdaGrad in different optimization problems. Furthermore, it is advisable to explore potential substitutes or improvements to AdaGrad that can overcome its convergence restrictions while preserving its adaptive features. By undertaking this rigorous examination, you will enhance your thesis by gaining a more profound comprehension of optimization methods and their practical significance in machine learning applications.

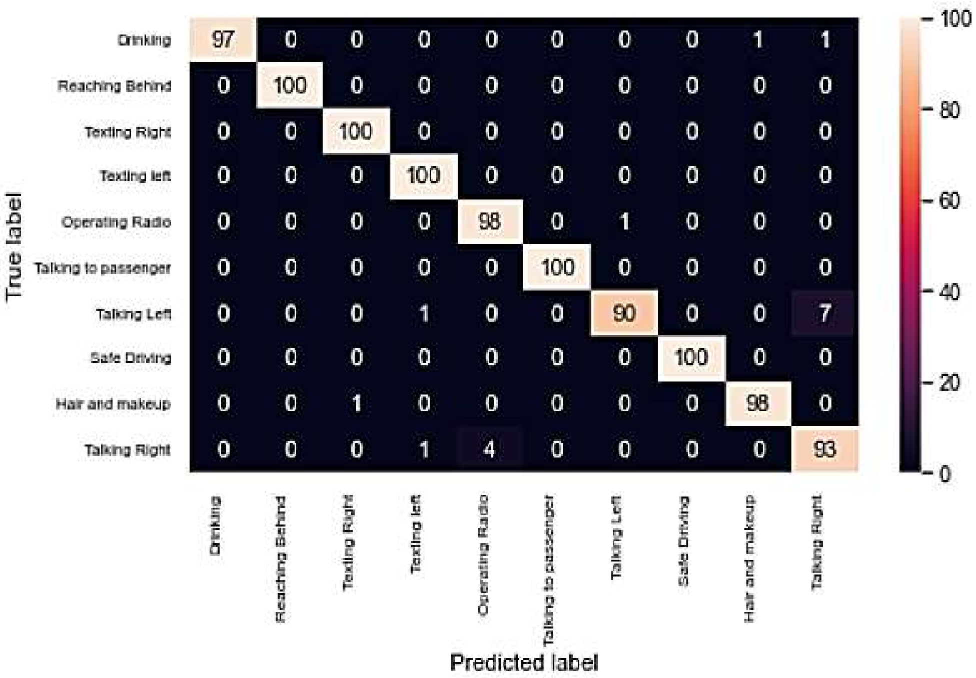

Figure 19 Confusion Matrix of AdaGrad Optimizer in DDB Dataset

4.4.5 Performance Comparison for All Adaptive Optimizers in DDB Dataset

TABLE VI. COMPARISON TABLE OF ABOVE FOUR OPTIMIZERS OF OUR DDB DATASET

Optimizer	Accuracy	Precision	Recall	F1 score
Adam	0.998667	0.998688	0.998667	0.998666
SGD	0.996000	0.996067	0.996000	0.995991
RMSprop	0.992000	0.992125	0.992000	0.992004
AdaGrad	0.982667	0.982897	0.982667	0.982618

These measures offer valuable information regarding the performance of each optimizer in terms of accuracy and the equilibrium between precision, recall, and F1 score. The Adam optimizer demonstrates superior performance on this dataset compared to the other optimizers, as evidenced by its highest values across all measures.

4.5 Kaggle Dataset Result Analysis

The Kaggle dataset [15] named State Farm Distracted Driver Detection is a comprehensive collection of data that provides ten different classes of distracted driving behaviors. As this dataset is in the public domain, many researchers have used it to develop models and algorithms for detecting distracted driving. We employed our model with four optimizers to evaluate its performance on this dataset. The optimizer functions used were Adam, RMSprop, AdaGrad, and SGD. By testing our model with multiple optimizers, we aimed to determine the most effective one for accurately detecting distracted driving behaviors in real-time scenarios. We analyzed the results obtained from each optimizer to identify any patterns or trends that could further enhance the model's performance.

4.5.1 Performance Analysis of ADAM Optimizer in Kaggle Dataset

The observation of loss of Adam in Figure 20 gave same behavior as we seen in our DDB dataset.

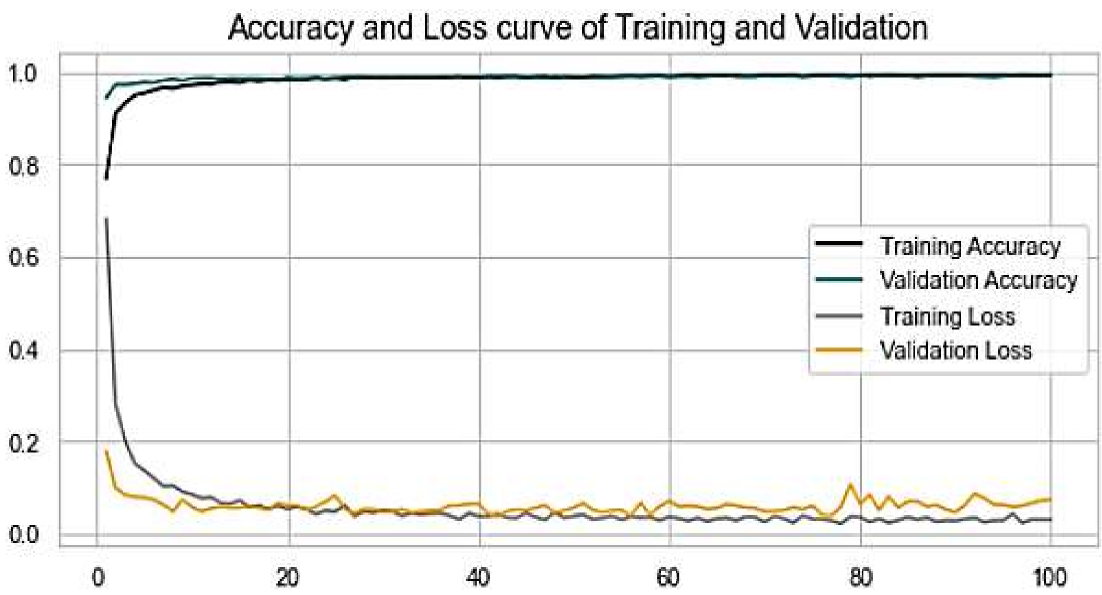

Figure 20 Accuracy and Loss curve of Adam Optimizer in Kaggle dataset

The parameter value and hyper parameter Details provided in following Table VII.

TABLE VII. EXPERIMENTS ON KAGGLE DATASET, ADAM OPTIMIZER'S PARAMETER VALUE AND HYPER PARAMETER

Hyper parameter	Parameter value
Number of Epoch	100
Batch Size	16
Optimizer	Adam
Objective function	Categorical Crossentropy
Hidden layer	ReLU
Output layer	Softmax

The confusion matrix (Figure 21) made it evident that the model can accurately identify the majority of positive cases, as evidenced by its high true positive rate and low false positive rate. Hence, although the elevated true positive rate is encouraging, it is imperative to conduct a thorough examination of the confusion matrix in order to derive significant insights on the model's efficacy and its practical implications.

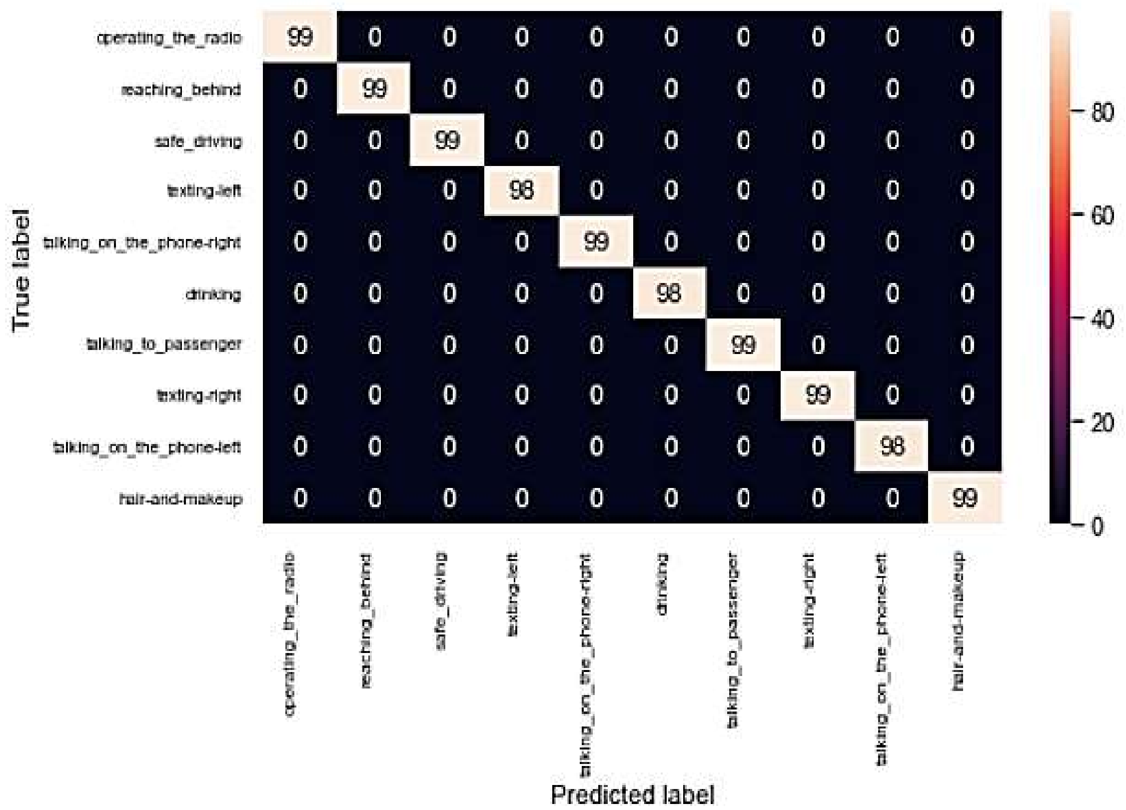

Figure 21 Confusion Matrix of Adam Optimizer in Kaggle dataset

4.5.2 Performance Analysis of SGD Optimizer

Following Figure 22, training accuracy and training loss presented smooth trends throughout the entire training process. The training accuracy steadily increased while the training loss consistently decreased, indicating that the model was learning and improving over time.

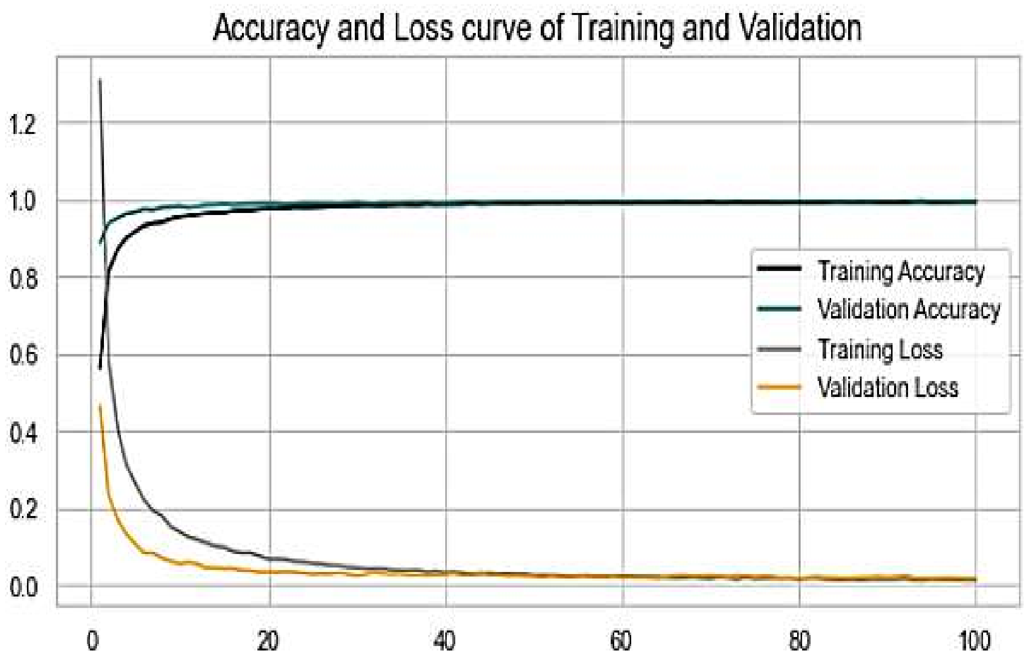

Figure 22 Accuracy & Loss of SGD Optimizer in Kaggle dataset

TABLE VIII. EXPERIMENTS ON KAGGLE DATASET, SGD OPTIMIZER'S PARAMETER VALUE AND HYPER PARAMETER

Hyper parameter	**Parameter value**
Number of Epoch	100
Batch Size	16
Optimizer	SGD
Objective function	Categorical Crossentropy
Hidden layer	ReLU
Output layer	Softmax

Using the SGD optimizer on the Kaggle dataset, we got the desired convergence rate for our model. The accuracy was 99.46%, which is an excellent result. The confusion (Figure 23) matrix also showed high performance with a low number of misclassifications. These results indicate that our model trained effectively and is capable of accurately predicting the classes in the dataset.

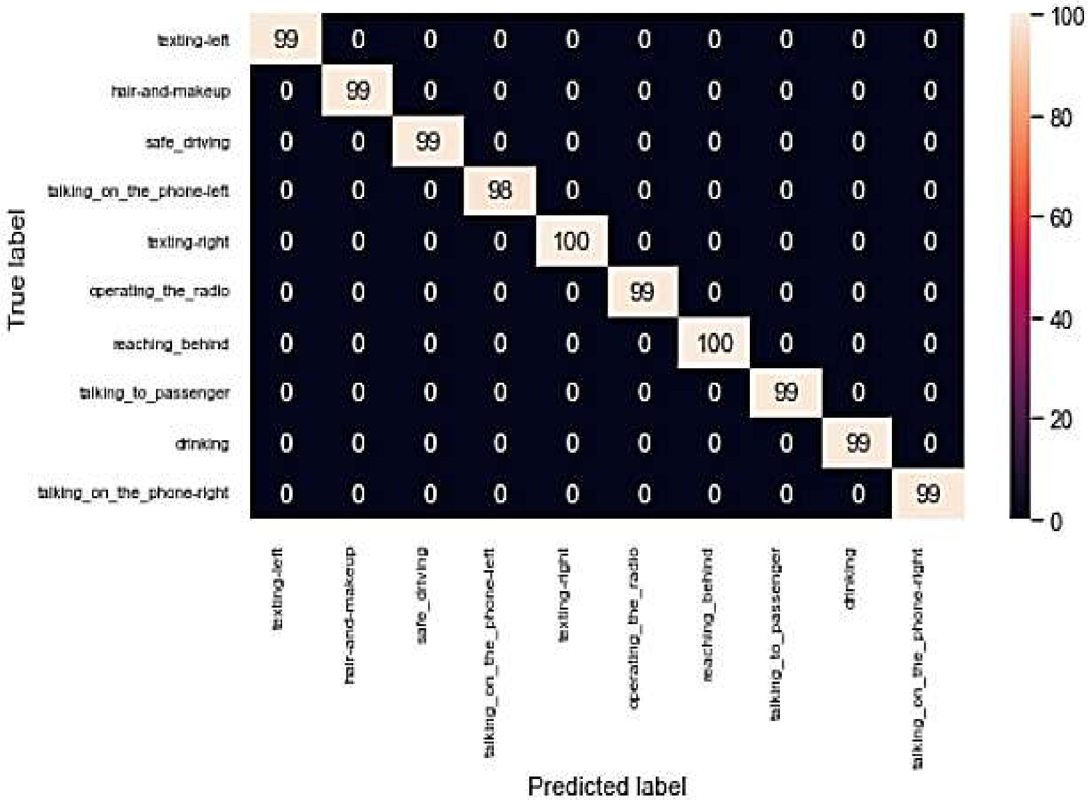

Figure 23 Confusion Matrix of SGD Optimizer in Kaggle dataset

4.5.3 Performance Analysis of RMSPROP Optimizer

To observe what would happen after 100 epochs, we ran the model with RMSProp on the Kaggle dataset. We witnessed (Figure 24) a steady increase in the performance metric up until the approximately 25th epoch. However, after that point, there was a sudden increase in the loss metric, but accuracy remained relatively stable. Regularization techniques, such as early stopping for loss minimization, could be a solution to address the sudden increase in the loss metric. We used early stopping for accuracy monitoring as well, which helped stabilize the accuracy metric even after the sudden increase in loss. Moreover, it would be advantageous to evaluate the efficacy of the RMSProp optimizer by comparing it with alternative optimization methods to determine if comparable challenges exist when employing different optimization strategies.

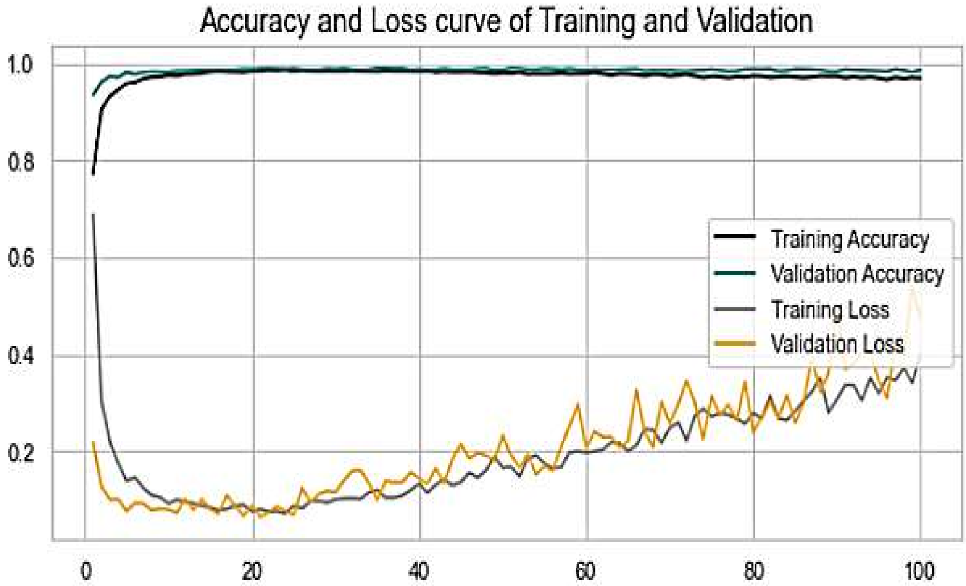

Figure 24 Accuracy and Loss curve of RMSProp Optimizer in Kaggle dataset

TABLE IX. **E**XPERIMENTS **O**N **K**AGGLE **D**ATASET, **R**MSPROP **O**PTIMIZER'S **P**ARAMETER **V**ALUE **A**ND **H**YPER **P**ARAMETER

Hyper parameter	**Parameter value**
Number of Epoch	100
Batch Size	16
Optimizer	RMSProp
Objective function	Categorical Crossentropy
Hidden layer	ReLU
Output layer	Softmax

The RMSProp analysis's confusion metrix (Figure 25) confirmed that the method worked well in terms of stability and speed of convergence. 98.73% accuracy was obtained, however there was a rise in loss after certain epochs. Nevertheless, the increase in loss reported after specific epochs necessitates further examination. Possible reasons that may contribute to this issue could be overfitting caused by an excessively complex model or inadequate scheduling of the learning rate. Furthermore, evaluating the model's ability to generalize on unfamiliar data sets could offer additional understanding of its behavior beyond the training data. Subsequent studies could prioritize the refinement of hyperparameters or the investigation of alternate optimization algorithms to address the observed increase in loss while still achieving high levels of accuracy.

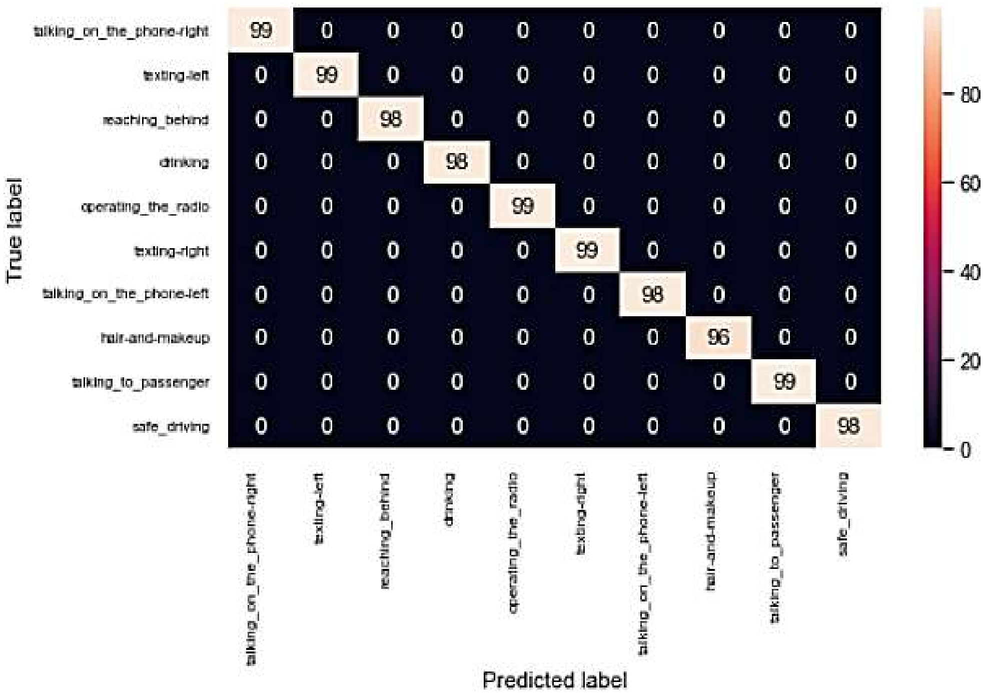

Figure 25 Confusion Matrix of RMSProp Optimizer in Kaggle dataset

4.5.4 Performance Analysis of AdaGRAD Optimizer

AdaGrad often shows slower convergence compared to RMSProp, but it is more robust to noisy or sparse data. Additionally, AdaGrad tends to perform well in scenarios where the learning rate needs to be automatically adjusted based on the gradient magnitudes. From Figure 26 of training plot and Figure 27 of confusion metrix we observe the performance of AdaGrad on Kaggle dataset with sparse features. The algorithm achieved an accuracy of 97.10%. AdaGrad's performance warrants further exploration in diverse real-world contexts due to its resilience to noisy or sparse data and its capacity to autonomously adapt the learning rate based on gradient magnitudes. Examining its efficacy across various types of datasets, namely those distinguished by high dimensionality or imbalanced class distributions, could yield profound understanding of its suitability and possible benefits. Moreover, doing comparative evaluations with other optimization algorithms like Adam or SGD with momentum could provide a full comprehension of AdaGrad's respective advantages and disadvantages. Furthermore, exploring the computational efficiency and scalability of AdaGrad, particularly in relation to extensive datasets or intricate neural network structures, would enhance the discourse on its practical usefulness.

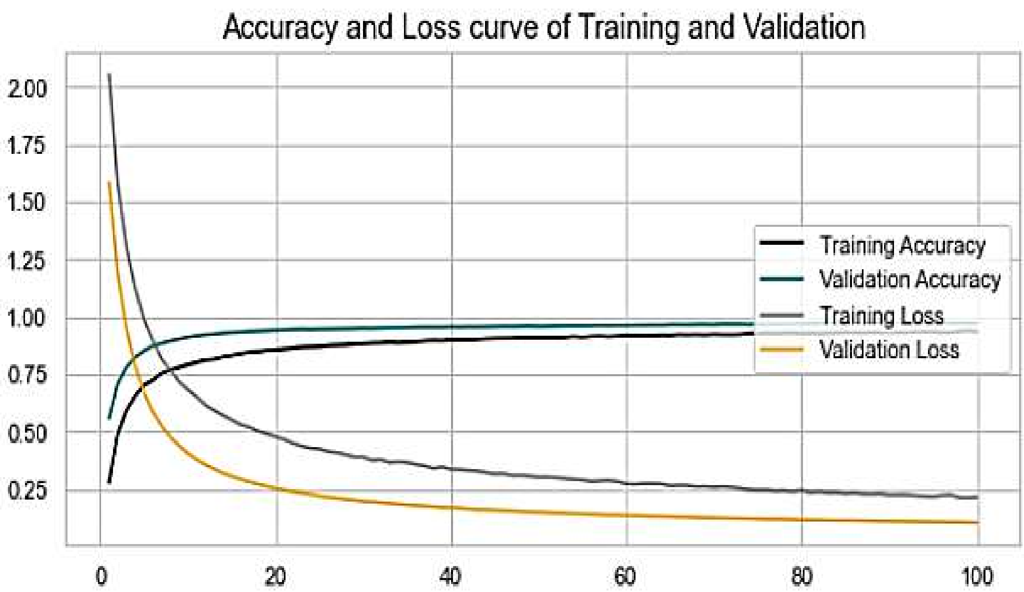

Figure 26 Accuracy and Loss curve of AdaGrad Optimizer in Kaggle dataset

TABLE X. EXPERIMENTS ON KAGGLE DATASET, ADAGRAD OPTIMIZER'S PARAMETER VALUE AND HYPER PARAMETER

Hyper parameter	**Parameter value**
Number of Epoch	100
Batch Size	16
Optimizer	AdaGrad
Objective function	Categorical Crossentropy
Hidden layer	ReLu
Output layer	Softmax

The Figure 27 in the confusion matrix offers vital insights into the performance of AdaGrad on the Kaggle dataset with sparse features. Through the examination of the confusion matrix, we may get a more profound comprehension of how AdaGrad's optimization technique manages the complexities of the dataset's class distributions. More precisely, analyzing the allocation of true positive, true negative, false positive, and false negative predictions can clarify AdaGrad's capacity to precisely categorize instances, especially when dealing with imbalanced classes or noisy data. Furthermore, examining the details of the confusion matrix, including precision, recall, and F1-score for each class, might provide a nuanced viewpoint on the efficacy of AdaGrad in various categories within the dataset. Comprehending these metrics can aid in identifying any possible vulnerabilities or places that need enhancement, hence directing future research paths or optimization tactics. Moreover, conducting a comparison between the performance illustrated in Figure 27 and that of alternative optimization algorithms, such as Adam or SGD with momentum, might offer significant insights for assessing AdaGrad's advantages and disadvantages in relation to its competitors. Through a comprehensive comparative analysis, we can determine whether AdaGrad's ability to handle noisy or sparse data leads to better classification accuracy, or if there are situations where other optimization strategies may perform better. By utilizing the information obtained from Figure 27 of the confusion matrix, in addition to other performance indicators, we may enhance our comprehension of AdaGrad's effectiveness in dealing with real-world datasets that include sparse features. By placing its performance in the context of optimization techniques and evaluating its computing efficiency and scalability, we may make educated judgments regarding its practical usefulness in many fields.

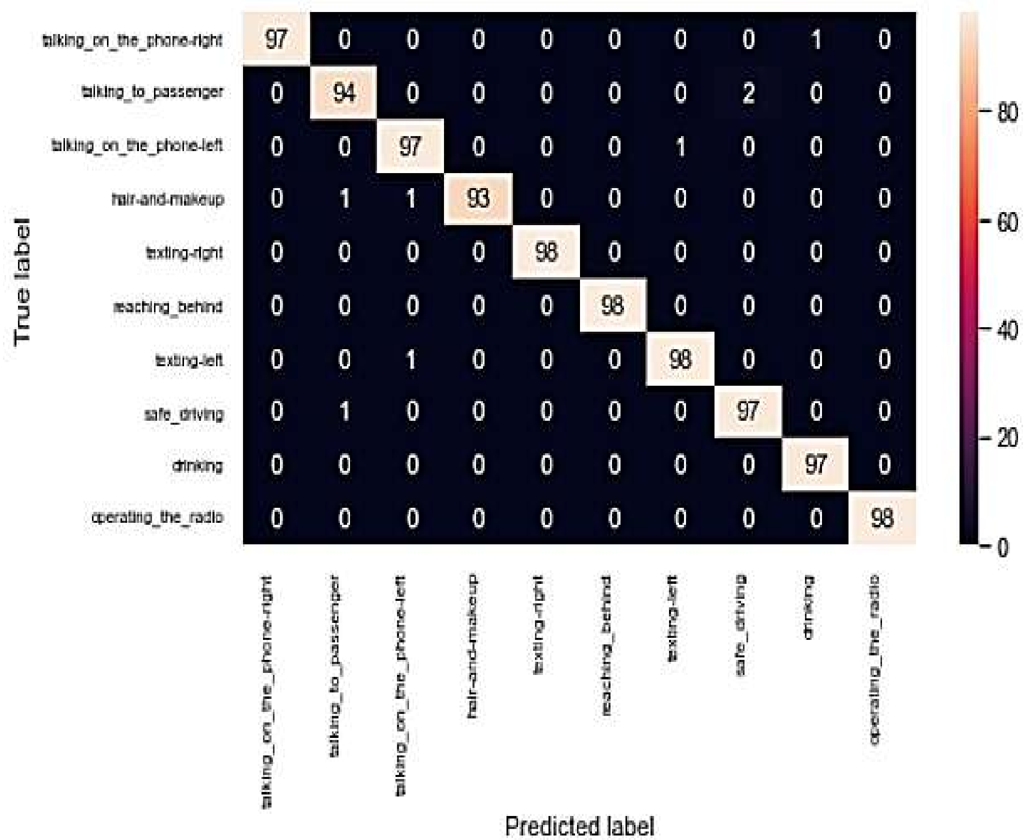

Figure 27 Confusion Matrix of AdaGrad Optimizer in Kaggle dataset

TABLE XI. COMPARISON TABLE OF ABOVE FOUR OPTIMIZERS OF OUR DDB DATASET

Optimizer	Accuracy	Precision	Recall	F1 score
Adam	0.991304	0.991318	0.991304	0.991300
SGD	0.994649	0.994665	0.994649	0.994652
RMSprop	0.987291	0.987376	0.987291	0.987297
AdaGrad	0.971014	0.971040	0.971014	0.970981

4.6 Comparison of Four Optimizer in Two Datasets Optimizer

The performance of four optimization algorithms (Adam, SGD, RMSProp, and AdaGrad) in two different datasets (DDB Dataset and Kaggle Dataset) was evaluated in terms of their final accuracy. Adam and SGD showed faster convergence and achieved higher final accuracy compared to RMSProp and AdaGrad in both datasets. However, Adam achieved higher final accuracy on the DDB Dataset while SGD performed better on the Kaggle Dataset [15]. The accuracy comparison plot in Figures 28 and Figure 29 provides a visual representation of how these algorithms perform on both datasets, allowing for a direct comparison between them.

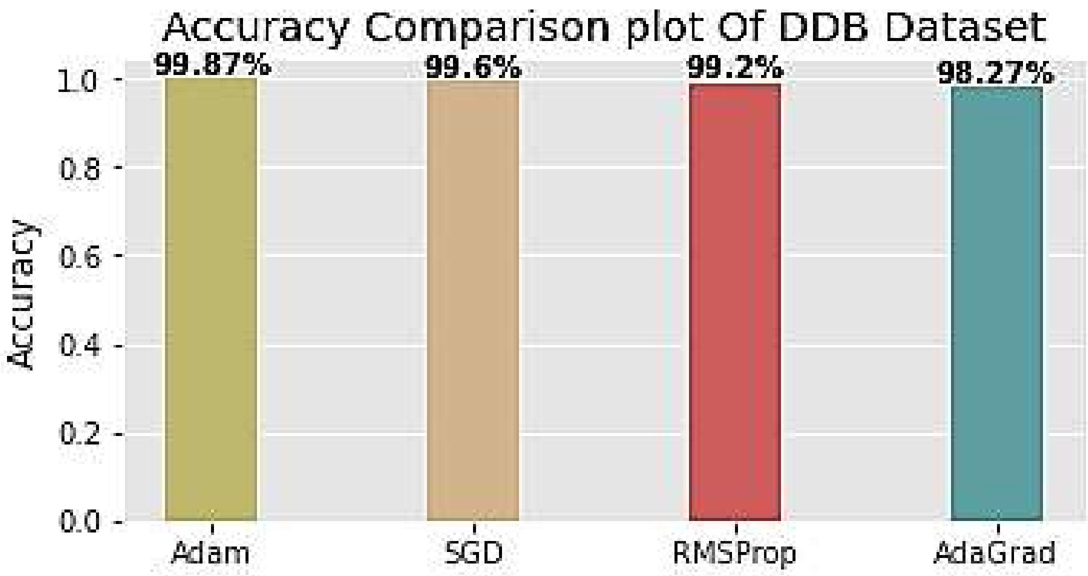

Figure 28 Comparison accuracies of 4 adaptive optimizers in our DDB Dataset.

Figure 29 visually demonstrates the convergence of each method towards ideal accuracy on the Kaggle Dataset, providing a dynamic picture of their paths. Through the examination of this visual representation, it is possible to identify subtle differences in the speed at which the algorithms come together, their capacity to remain stable, and their general efficiency in addressing the particular difficulties presented by the Kaggle Dataset. Furthermore, Figure 29 has the potential to illuminate any discrepancies in performance patterns when compared to the DDB Dataset, thereby offering a full comprehension of algorithmic behavior across various data environments.

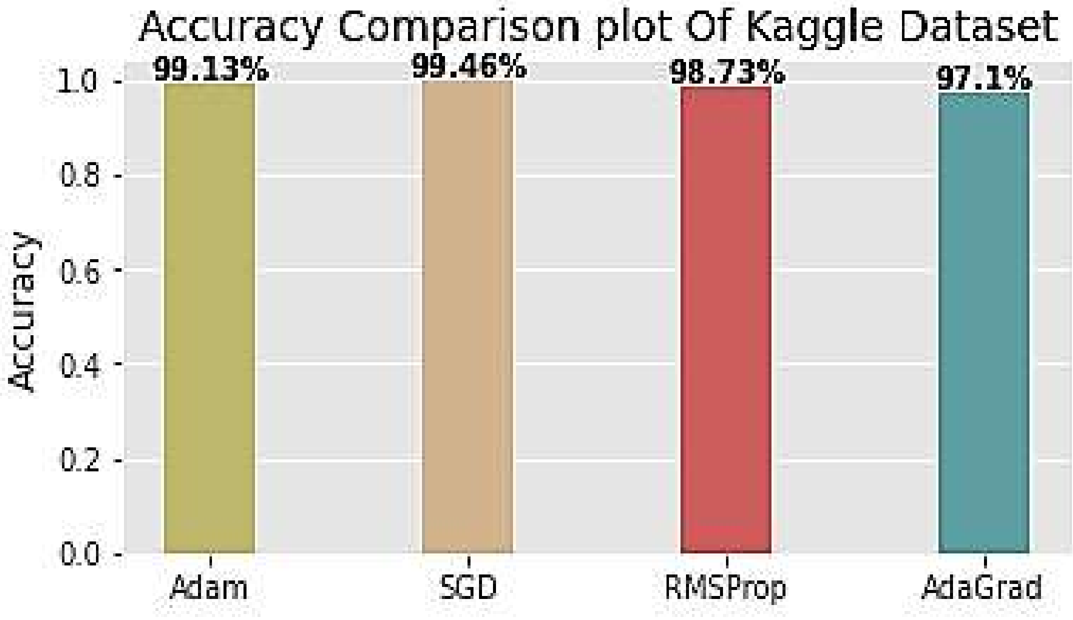

Figure 29 Comparison accuracies of 4 different optimizers in Kaggle Dataset.

4.7 Comparison Result of Kaggle Dataset with Existing Work

We compared our model with the other existing models [38] [39] [40] using SF3D Kaggle dataset that were used on the existing models and found that our model performed better in terms of accuracy and loss. The first dataset is made up from our car dashcam video. Our validation results showed that the model was able to identify driver with high accuracy. The second dataset [8] SF3D was a publicly accessible data consist of ten classes. Again, our model had better accuracy and loss compared to the existing models on this dataset. Existence Kaggle Dataset has total images 22422. We collected that Images from Kaggle exactly like these paper [38] [39] [40] did. In our proposed model in different optimizer like Adam, SGD, RMS prop, AdaGrad validation accuracies 99.13%, 99.46%, 98.73, 97.1% (Figure 29). The comparison of following models for predicting the outcomes of the SF3D datasets shows that the proposed model has achieved significantly better performance (Table XII).

TABLE XII. COMPARISON TABLE OF ABOVE FOUR OPTIMIZERS FOR KAGGLE DATASET

Author	Year	Dataset	Technique Used	Accuracy
F. Sajid, A. R. Javed [38]	2021	SF3D	EfficientDet-D3	99.16%
Vaegae, Naveen Kumar [39]	2022	SF3D	VGG16	93.46%
			ResNet50	93.12%
Hossain, Md Uzzol [40]	2022	SF3D	Simple CNN	97.45%
			VGG-16	94.01%
			ResNet50	94.28%
			MobileNetV2	98.12%
Our Modified VGG16 method	2023	SF3D	Adam	99.13%
			SGD	99.46%
			RMSProp	98.72%
			AdaGrad	97.10%

4.8 Discussion

We curated a dataset on distracted driver behavior following the structure of Kaggle datasets and developed a model adapted from VGG-16. Employing various optimizers in our research, we tested the model's validity using an existing Kaggle dataset. The outcomes from this established dataset revealed enhanced model performance, affirming the success of our modified VGG-16 model in identifying distracted behavior types. Encouraged by these promising results, we are optimistic about the broader applicability of our model to other distracted driver datasets, anticipating improvements in accuracy and overall performance. In our thesis, we undertook an extensive exploration to tackle the pressing problem of distracted driving using machine learning techniques. Our main goal was to create a strong model that could precisely detect different forms of inattentive behavior displayed by drivers on the road. In order to accomplish this, we adopted a practical approach by carefully selecting and organizing our own collection of data. This process entailed recording genuine driving situations using a dashboard camera put in a vehicle that was driven along the Bayezid Link Road in Chattogram. The rigorous process of dataset curation ensured the comprehensive collection of a wide array of distracted driving behaviors, encompassing activities such as texting, eating, adjusting the radio, and engaging in conversations, among others. Adhering to the typical format observed in Kaggle datasets, we meticulously arranged and preprocessed our data to facilitate training and validation. Our methods mostly involved modifying the VGG-16 model, a well-known convolutional neural network architecture that is highly regarded for its success in image identification tasks. By utilizing this framework, we carefully optimized our model to accurately detect distracted drivers by adapting it to their subtle variations. During our research, we examined different optimization strategies to improve the performance of the model. We thoroughly tested its validity using our carefully selected dataset. The results obtained from this method were quite encouraging, demonstrating a substantial enhancement in the model's capacity to reliably detect and categorize distracted driving behaviors. The effectiveness of our technique is demonstrated by the accomplishment we have reached using our modified VGG-16 model on our well-established dataset. Based on these results, we are optimistic that our model may be applied to other datasets involving distracted drivers. We expect that our model will not only enhance accuracy but also make significant contributions to the overall progress in tackling the crucial problem of distracted driving on the roads. It is crucial to thoroughly examine the significance of our research findings within the wider context of road safety and the application of machine learning. It is crucial to engage in a discussion regarding the prospective real-world ramifications of our concept. Possible approaches include incorporating it into intelligent vehicle systems to promptly notify drivers displaying signs of distraction, partnering with transportation authorities to proactively implement measures to prevent accidents resulting from driver inattentiveness.

4.9 Summary

With its pre-trained layers, the VGG16 Deep Convolutional Neural Network (CNN) has strong capabilities. An summary of the dataset, experimental setups, and environmental details utilized to put our suggested models into practice are provided in this chapter. This chapter concludes with a comprehensive discussion of the experimental results obtained from implementing the approaches we have proposed.

Chapter 5
Conclusion and Future Works

5.1 Conclusion

This paper has successfully addressed the intricate challenges posed by the intricate nature of human behavior and the myriad factors contributing to distractions on the road, especially in the context of the limited ability to provide precise and timely early warnings. The introduced novel method, centered on modified CNN transfer learning, has been developed to enhance the detection of distracted driver behavior. By integrating transfer learning within the Convolutional Neural Networks framework, the method has adeptly navigated the complexities associated with identifying distracted driving behaviors. To assess the robustness of the proposed approach, four different optimizers—Adam, SGD, RMSProp, and AdaGrad—were compared on our training model. The experimental results demonstrated an impressive overall accuracy ranging from 98% to over 99%, with the highest achieved accuracy of SGD reaching an outstanding 99.46% on a publicly available dataset. These findings substantiate the efficacy of the proposed methodology in significantly improving the precision of distracted driver behavior detection, marking a substantial advancement in the realm of road safety.

5.2 Contribution of this Thesis

The primary contribution of this thesis lies in the proposal and implementation of a novel deep learning-based distraction prediction technique. This innovative approach aims to identify ten distinct unsafe driving types solely through image analysis. A key aspect of our contribution involves the introduction of a new feature integrated into a pre-trained model. To develop and validate the model, we amassed a comprehensive dataset comprising images extracted from our car's dashcam videos and additional street-view images. The deep learning model underwent extensive training on this large-scale dataset, ensuring its capability to deliver accurate and reliable predictions for enhanced effectiveness in detecting various forms of unsafe driving behaviors.

5.3 Future Works

In the pursuit of further advancements, numerous areas warrant exploration within the scope of this thesis. One avenue for future research involves the curation of a more comprehensive dataset. While images were meticulously collected from our car's dashcam videos, an expansion of the dataset could incorporate additional street-view images to enhance diversity and broaden the scope of the image dataset. Furthermore, future work could focus on optimizing and fine-tuning the deep learning model through continued rigorous training on an even larger-scale dataset. This extended training process aims to fortify the model's predictive capabilities, ensuring heightened robustness and accuracy in its performance. Such endeavors would contribute to the continual refinement and enhancement of the proposed methodology, paving the way for more sophisticated and effective applications in the realm of driving behavior detection. Future study could investigate the integration of further contextual information to enhance the driving behavior recognition system, in addition to expanding the dataset and improving the deep learning model. This may need the integration of data from different sensors installed in the car, including GPS, accelerometer, and gyroscope, in order to provide a comprehensive picture of the driving conditions and the driver's behavior. Moreover, exploring the use of live data streams could facilitate the creation of flexible and responsive algorithms that can adjust to evolving road circumstances and driver actions in real-time. In addition, investigating the incorporation of reinforcement learning methods could enable the creation of a self-enhancing system that consistently acquires knowledge and adjusts its behavior in response to feedback obtained from real-life driving situations. These developments have the potential to greatly improve the accuracy, reliability, and usefulness of driving behavior detection systems, which in turn can boost road safety and aid in the development of autonomous driving technology. Furthermore, it is essential to take into account the societal and ethical ramifications of implementing such technology. Subsequent investigations should focus on inquiries pertaining to the confidentiality of data, prejudiced algorithms, and the possible consequences on driver independence. Engaging in cooperative endeavors with policymakers, ethicists, and stakeholders can guarantee the responsible and fair implementation of driving behavior detection systems. In addition, investigating the incorporation of several data sources, such as audio and natural language processing, could yield further understanding of driver behavior, emotions, and intentions. By adopting an interdisciplinary approach, it is possible to develop more complete and nuanced models that can enhance the system's comprehension of human variables in driving. This, in turn, can improve overall performance and safety on the roads.

References

[1] Groth Law Accident Injury Attorneys., "100 Distracted Driving Facts & Statistics | Groth Law Accident Injury Attorneys," May 5, 2022.

[2] National Center for Statistics and Analysis., "Distracted driving in 2021," National Highway Traffic Safety Administration., Washington, 2023, May.

[3] L. J. M. T. R. Z. E. &. L. B. A. Blincoe, "The Economic and Societal Impact Of Motor Vehicle Crashes, 2010(Revised)," National Highway Traffic Safety Administration, Washington, (2015, May).

[4] A. a. C. T. Lie, "ISO 39001 Road Traffic Safety Management System, Performance Recording, and Reporting.," in In The Vision Zero Handbook: Theory, Technology and Management for a Zero Casualty Policy, Cham: Springer International Publishing, 2022, pp. 1-13.

[5] A. S. B. D. C. C. L. D. H. G. L. K. N. P. H.-Y. W. C. a. Z. S. Aghaei, "Smart driver monitoring: when signal processing meets human factors: in the driver's seat," IEEE Signal Processing Magazine , vol. 33, no. 6, pp. 35-48., 2016.

[6] W. L. L. W. S. a. L. L. Shi, "Autonomous Driving Landscape.," Springer, pp. 1-18, 2021.

[7] P. Z. S. S. H. Z. Z. X. C. J. J. Z. L. e. a. Xiao, "Pandaset: Advanced sensor suite dataset for autonomous driving.," 2021.

[8] S. G. S. T. Bhakti Baheti, "Detection of Distracted Driver Using Convolutional Neural Network," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, pp. 1032-1038, 2018.

[9] D. a. Y. Y. Feng, "Machine Learning Techniques for Distracted Driver Detection.," 2019.

[10] S. S. D. S. R. A. P. Prof. Pramila M. Chawan, "Distracted Driver Detection and Classification," Int. Journal of Engineering Research and Application , vol. 8, no. 4, pp. 60-64, 2018.

[11] T. Y. Z. C. Z. K. L. a. M. S. Hoang Ngan Le, "Multiple scale faster-rcnn approach to driver's cell-phone usage and hands on steering wheel detection.," In Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp. 46-53, 2016.

[12] H. V. J. H. C. C.-O. C. Y.-L. C. a. B. R. Koay, "Optimally-weighted image-pose approach (owipa) for distracted driver detection and classification.," vol. 14, no. 21, p. 4837, 202.

[13] X. W. J. C. S. W. a. Y. Z. C. Huang, "HCF: A Hybrid CNN Framework for Behavior Detection of Distracted Drivers," IEEE Access, vol. 8, pp. 109335-109349, 2020.

[14] D. H. M. D. W. S. H. B. a. G. C. Tran, "Real-time detection of distracted driving based on deep learning," IET Intelligent Transport Systems, vol. 12, no. 10, pp. 1210-1219, 2018.

[15] S. Farm, "State farm distracted driver detection-data.," [EB/OL] (2016).

[16] C. H. ,. Z. X. ,. S. M. Song Liye, "Distracted driving behavior detection based on CNN and multi-model fusion," Proceedings of the SPIE, vol. 12636, no. 126363, p. 7, 2023.

[17] C. Zhang et al., "Real-time Unsafe Driving Behavior Detection Using Convolutional Neural Networks," in IEEE Transactions on Intelligent Transportation Systems, vol. 20, no. 5, pp. 1987-1996, May 2019.

[18] Y. Wang, Z. Zhang, and Y. Liu, "CNN-Based Method for Unsafe Driving Behavior Detection Using In-Vehicle Camera Data," in IEEE Intelligent Vehicles Symposium (IV), Paris, France, 2019, pp. 1903-1908.

[19] L. Zhou et al., "Unsupervised Learning of Unsafe Driving Behavior with Convolutional Neural Networks," in IEEE Access, vol. 7, pp. 106563-106571, 2019.

[20] S. Li, X. Xu, and H. Wang, "Deep Learning-Based Unsafe Driving Behavior Detection System Using Dashboard Cameras," in IEEE Access, vol. 8, pp. 79889-79898, 2020.

[21] G. Chen, Li. U, and Ching Ung, "CNN-Based Approach for Unsafe Driving Behavior Detection from Smartphone Videos," in IEEE International Conference on Multimedia and Expo (ICME), London, UK, 2020, pp. 1-6.

[22] W. Liu, S. Wang, and L. Zhang, "Detecting Unsafe Driving Behavior with Convolutional Neural Networks and In-Vehicle Sensors," in IEEE Sensors Journal, vol. 20, no. 17, pp. 9823-9830, Sept. 2020.

[23] H. Wu, Nu. Wo, and Chi. Chung, "An Efficient Framework for Real-Time Detection of Unsafe Driving Behavior Using Convolutional Neural Networks," in IEEE Access, vol. 8, pp. 204901-204912, 2020.

[24] X. Chen, U. Ju, and Huang, "Improving Unsafe Driving Behavior Detection Accuracy with Deep Convolutional Neural Networks," in IEEE Intelligent Transportation Systems Conference (ITSC), Rhodes, Greece, 2021, pp. 1-6.

[25] J. Yang, Wo. Chen, and Pong. Gen, "Fusion of Sensor Data and CNN Features for Accurate Unsafe Driving Behavior Detection," in IEEE Transactions on Vehicular Technology, vol. 70, no. 2, pp. 1340-1350, Feb. 2021.

[26] Z. Huang, Y. Chen, and X. Li, "Robust Unsafe Driving Behavior Detection Using Multi-Stream Convolutional Neural Networks," in IEEE Access, vol. 9, pp. 28890-28901, 2021.

[27] Y. Zhao, Eung. Pen, and Cung. Ji, "Hierarchical Feature Fusion for Improved Unsafe Driving Behavior Detection with Convolutional Neural Networks," in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toronto, ON, Canada, 2021, pp. 8653-8657.

[28] Q. Liu, Nun. Ji, and Fung. Shi, "Context-Aware Unsafe Driving Behavior Detection Using Temporally Ordered Convolutional Neural Networks," in IEEE Transactions on Intelligent Transportation Systems, vol. 22, no. 10, pp. 5715-5725, Oct. 2021.

[29] S. Zhang, Sun. Chi, and Lu. Wi, "Joint Spatial-Temporal Learning for Unsafe Driving Behavior Detection Using Convolutional Neural Networks," in IEEE Transactions on Multimedia, vol. 23, no. 3, pp. 1053-1063, March 2021.

[30] Y. Wang, Peng. Ki, and Q. Liu, "Real-Time Unsafe Driving Behavior Detection Using Deep Convolutional Neural Networks with Attention Mechanism," in IEEE Transactions on Industrial Informatics, vol. 17, no. 4, pp. 2588-2596, April 2021.

[31] Y. H. H. a. M. G. Moukafih, "Aggressive Driving Detection Using Deep Learning-based Time Series Classification.," IEEE, pp. 1-5, 2019.

[32] H. S. M. A. T. T. M. a. J. L. G. Kim, "Toward real-time estimation of driver situation awareness: An eye-tracking approach based on moving objects of interest," IEEE, vol. 4, pp. 1035-1041, 2020.

[33] R. M. S. M. K. K. A.-K. M. K. a. K. B. Jabbar, "Driver drowsiness detection model using convolutional neural networks techniques for android application.," IEEE, pp. 237-242, 2020.

[34] D. H. M. D. J. L. a. W. S. Tran, "Real-time detection of distracted driving using dual cameras.," IEEE, pp. pp. 2014-2019, 2020.

[35] T. M. G. C. M. A. H. a. A. F. Tamagusko, "Deep Learning applied to Road Accident Detection with Transfer Learning and Synthetic Images," Transportation research procedia, vol. 64, pp. 90-97, 2022.

[36] J. L. R. E. M. I. P. a. A. L. Vitorino, "Adversarial Robustness and Feature Impact Analysis for Driver Drowsiness Detection," in In International Conference on Artificial Intelligence in Medicine, Cham: Springer Nature Switzerland, 2023.

[37] A. J. S. S. J. R. S. a. Y. G. Goel, "Driver Drowsiness Detection System," Journal of Pharmaceutical Negative Results, pp. 1169-1178, 2023.

[38] A. R. J. A. B. N. K. A. A. a. M. R. F. Sajid, "An Efficient Deep Learning Framework for Distracted Driver Detection," in IEEE Access, vol. 9, pp. 169270-169280, 2021.

[39] N. K. K. K. P. K. B. a. O. O. O. Vaegae, "Design of an Efficient Distracted Driver Detection System: Deep Learning Approaches.," IEEE Access, vol. 10, pp. 116087-116097, 2022.

[40] M. U. M. A. R. M. M. I. A. A. M. A. U. a. B. K. P. Hossain, "Automatic driver distraction detection using deep convolutional neural networks.," Intelligent Systems with Applications, vol. 14, p. 200075, 2022.