Faculty of Engineering

DISTRACTED DRIVER DETECTION USING DEEP LEARNING

Koh Qi Zhe

Bachelor of Engineering

Electrical and Electronics Engineering with Honours

2023

UNIVERSITI MALAYSIA SARAWAK

Grade: ____________

Please tick (√)
Final Year Project Report ☑
Masters ☐
PhD ☐

DECLARATION OF ORIGINAL WORK

This declaration is made on the ......24......day of......July......2023.

Student's Declaration:

I ------ Koh Qi Zhe, 70101, Faculty Engineering ------
(PLEASE INDICATE STUDENT'S NAME, MATRIC NO. AND FACULTY) hereby declare that the work entitled ---- Distracted Driver Detection Using Deep Learning ---- is my original work. I have not copied from any other students' work or from any other sources except where due reference or acknowledgement is made explicitly in the text, nor has any part been written for me by another person.

24th July 2023
Date submitted

Koh Qi Zhe (70101)
Name of the student (Matric No.)

Supervisor's Declaration:

I---- Ts Dr Kho Lee Chin ---- (SUPERVISOR'S NAME) hereby certifies that the work entitled ---- Distracted Driver Detection Using Deep Learning ----(TITLE) was prepared by the above named student, and was submitted to the "FACULTY" as a * partial/full fulfillment for the conferment of Bachelor of Electrical and Electronics Engineering with Honours (PLEASE INDICATE THE DEGREE), and the aforementioned work, to the best of my knowledge, is the said student's work.

Received for examination by: ____________________
(Name of the supervisor)

Date: 24 July 2023

I declare that Project/Thesis is classified as (Please tick (√)):

☐ **CONFIDENTIAL** (Contains confidential information under the Official Secret Act 1972)*****
☐ **RESTRICTED** (Contains restricted information as specified by the organisation where research was done)*****
☑ **OPEN ACCESS**

Validation of Project/Thesis

I therefore duly affirm with free consent and willingly declare that this said Project/Thesis shall be placed officially in the Centre for Academic Information Services with the abiding interest and rights as follows:

- This Project/Thesis is the sole legal property of Universiti Malaysia Sarawak (UNIMAS).
- The Centre for Academic Information Services has the lawful right to make copies for the purpose of academic and research only and not for other purpose.
- The Centre for Academic Information Services has the lawful right to digitalise the content for the Local Content Database.
- The Centre for Academic Information Services has the lawful right to make copies of the Project/Thesis for academic exchange between Higher Learning Institute.
- No dispute or any claim shall arise from the student itself neither third party on this Project/Thesis once it becomes the sole property of UNIMAS.
- This Project/Thesis or any material, data and information related to it shall not be distributed, published or disclosed to any party by the student except with UNIMAS permission.

Student signature ____________________ (24th July 2023)

Supervisor signature: ____________________ (24th July 2023)

Current Address:
1-2-3 (Townhouse), Jalan Chan Siew Teong, 11200 Tanjong Bungah, Pulau Pinang, Malaysia.

Notes: * If the Project/Thesis is **CONFIDENTIAL** or **RESTRICTED**, please attach together as annexure a letter from the organisation with the period and reasons of confidentiality and restriction.

[The instrument is duly prepared by The Centre for Academic Information Services]

DISTRACTED DRIVER DETECTION USING DEEP LEARNING

Distracted Driver Detection Using Deep Learning

KOH QI ZHE

A dissertation submitted in partial fulfilment
of the requirement for the degree of
Bachelor of Engineering
Electrical and Electronics Engineering with Honours

Faculty of Engineering

Universiti Malaysia Sarawak

2023

ACKNOWLEDGEMENT

First of all, I am grateful to have the opportunity to choose Ts. Dr Kho Lee Chin as my Final Year Project supervisor. I would like to thank my supervisor for her support, guidance, and supervision in helping me finish my final year project. She is an excellent supervisor who always shows me a lot of support, has a lot of patience, and helps me plan my project. Next, I want to express my gratitude to my family, especially my parents, for being there for me constantly, for their encouragement, and for their support while I studied. Finally, I would want to thank my friends and classmates from University Malaysia Sarawak for their never-ending support and encouragement.

ABSTRACT

Driving involves a wide range of complex operations and the coordination of multiple senses, making it a task that requires utmost attention and focus. Various factors, such as cell phone use, adjusted audio equipment, smoking, consumed food and drinks, conversed with passengers, or experienced fatigue, distracted drivers and jeopardized their safety, resulting in car collisions and injuries. The rising prevalence of distracted driving poses significant risks to road safety, leading to increased accidents and fatalities. Various studies explore different approaches to detect driver distraction, from traditional machine learning to advanced deep learning. Deep learning, especially CNN-based methods, show better accuracy and real-time performance. In this project, a lightweight deep learning detection model based on MobileNetV2 was proposed to address the detection of distracted driver actions without causing discomfort to the driver, as some current technologies involve wearing sensors that can be uncomfortable. The proposed model was enhanced by incorporating attention mechanisms like the SE module. The model was trained and tested using the American University in Cairo (AUC) distracted driver dataset, which encompassed 10 distraction categories. Techniques like hyperparameter tuning, data augmentation, and class weighting were utilized to optimize the model, achieving an impressive accuracy of 93% with the configuration of batch size 32, learning rate 0.0001, 21 epochs. Evaluation metrics, including confusion matrix, FPS, accuracy, precision, recall, and F1 score, were employed to assess the effectiveness of the model. Additionally, the proposed method was compared to MobileNetV2 model and other existing architectures in terms of accuracy and parameters. It outperformed unmodified deep learning models and maintained a balance between accuracy and parameter utilization, while some other modified models perform slightly better. The proposed method exhibiting promising potential in accurately detecting distracted drivers with efficiency.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AdaGrad	-	Adaptive Gradient
ADB	-	Adaptive Boosting
API	-	Application Programming Interface
AUC	-	American University in Cairo
BiFPN	-	bi-directional feature pyramid networks
BiLSTM	-	Bidirectional Long Short-Term Memory
BN	-	Batch Normalization
c	-	Number of Output Channel
CAM	-	Class Activation Mapping
CDCL	-	Cross Domain Complementary Learning
CNN	-	Convolutional Neural Network
CPU	-	Central Processing Unit
DT	-	Decision Tree
DMD	-	Driver Monitoring Dataset
FC	-	Fully Connected
FN	-	False Negative
FNN	-	Feed forward Neural Network
FP	-	False Positive
FPS	-	Frame per Second
GB	-	Gradient Boosting
GPU	-	Graphics processing unit
GTX	-	Giga Texel Shader Extreme
HCF	-	Hybrid CNN Framework
HMLSTM	-	Hierarchical Multiscale Long Short-Term Memory
HOG	-	Histogram of Gradient
HRNN	-	Hierarchical Recurrent Neural Network
HS	-	h-swish,
HSDDD	-	Hybrid Scheme Detection of Distracted Driving
KFPN	-	Key Point Feature Pyramid Network
KNN	-	k-nearest Neighbours

LR	-	Logistic Regression
LWANet	-	Lightweight Attention-based Network
mAP	-	Average Mean Accuracy Rate
n	-	Number of times
NBN	-	No Batch Normalization
NB	-	Naïve Bayes
MTRO	-	Adam Momentum Training Rate Optimization
NAS	-	Network Architecture Search
NHTSA	-	National Highway Traffic Safety Administration
NL	-	Nonlinearity,
PAFs	-	Part Affinity Fields
PCA	-	Principal Component Analysis
RAM	-	Random Access Memory
RCNN	-	Region-based Convolutional Neural Network
ReLu	-	Rectifier Linear Unit
RF	-	Random Forest
RGB	-	Red Green Blue
RNN	-	Recurrent Neural Network
RMSProp	-	Root Mean Squared Propagation
ROI	-	Region of Interest
RPN	-	Region Proposal Network
RWSE	-	Residual Weighted Squeeze-and-Excite
s	-	Stride
t	-	Expansion Factor
SE	-	Squeeze-and-Excitation
SGD	-	Stochastic Gradient Descent
SVM	-	Support Vector Machine
TN	-	True Negative
TP	-	True Positive

CHAPTER 1

INTRODUCTION

1.1 Background

According to World Health Organization's estimation, a significant number of individuals, ranging from 30 to 50 million, sustain injuries and approximately 1.3 million individuals succumb to fatalities as a result of traffic accidents annually [1]. While there are various factors that contribute to these statistics, distracted driving is considered a primary contributing factor. The National Highway Traffic Safety Administration (NHTSA) defines distracted driving as any behaviour that shift concentration away from the act of driving, such as utilizing electronic devices, consuming food or beverages, engaging in conversations with passengers, manipulating in-vehicle technology, or engaging in any other activities [2].

1.1.1 Types of Distraction

Driver distraction can manifest in three forms: visual, manual, and cognitive distractions, where visual distractions have involved diverting one's gaze from the road, manual distractions have involved disengaging one's hands from the steering wheel, and cognitive distractions have involved not being mentally focused on the task of driving. [3].

Visual distractions happen for a variety of reasons, such as looking at a phone, adjusting the radio, or even looking at something outside the vehicle. Visual distraction is particularly dangerous because it can cause a driver to miss important visual cues, such as traffic signals, pedestrians, or other vehicles.

Manual distraction happens when a driver disengaging one's hands from the steering wheel. This can happen for a variety of reasons, such as reaching for an object, eating, or even changing the radio station [4]. Manual distraction is dangerous because it can cause a driver to lose control of the vehicle. Drivers are now exposed to extra distraction-inducing variables thanks to the widespread use of on-board devices like smart phones and navigation systems. Several top automobile manufacturers, such as Bayerische Motoren Werke, Mercedes-Benz, Ford, Toyota have integrated advanced infotainment, control interfaces, and display systems into their cars. However, manipulating these in-car features while driving can cause significant distractions that may lead to traffic incidents. [5]

Cognitive distractions can be defined as not being involved mentally concentrated on the task of driving [6]. This can happen for a variety of reasons such as having an intense conversation, being preoccupied with personal problems, daydreaming, or even feeling stressed or tired. This type of distraction can be particularly dangerous as it can cause a driver to miss important visual cues or not react quickly in an emergency.

In our fast-paced world, driver distractions have become a critical concern, contributing to numerous accidents and endangering lives on the road. Identifying and mitigating distractions while driving has been a significant challenge for traditional methods, which often rely on manual assessment or limited sensor-based approaches. However, the advent of deep learning has opened new avenues to tackle this pressing issue effectively.

1.1.2 Deep Learning

Deep learning is a specialized area of machine learning that utilizes neural networks for studying and resolving difficult problems. It is based on the idea that machines can learn to perform tasks by analysing large amounts of data and finding patterns, without being explicitly programmed to do so. This allows deep learning algorithms to learn and improve over time, making them highly adaptable and able to handle a wide range of tasks.

Deep learning is particularly effective at handling large amounts of data, such as images, videos, and audio. It is a technology that can be applied to a wide range of tasks such as forecasting, speech recognition, image identification, and natural language understanding. One of the significant advantages of deep learning is its ability to learn and adapt without human supervision, which makes it ideal for tasks that are too complex for humans to perform manually.

Deep learning algorithms utilize artificial neural networks as their foundation. These neural networks are modelled after the human brain and consist of layers of interconnected nodes designed to identify features within data. Each layer within the neural network is capable for recognizing a distinct level of abstraction, such as edges, shapes, and objects, in the data [7].

Convolutional neural networks (CNN), recurrent neural networks (RNN), and feedforward neural networks (FNN), are the types of deep learning methods. Each type is optimized for a specific task and has its own unique characteristics and uses. Convolutional neural networks are intended to process and analyse images, and they are frequently employed for tasks including image segmentation, object recognition, and classification. [8] Recurrent neural networks are frequently employed for applications like speech recognition and natural language processing because they are built to handle sequential input, such as time series and natural language. [9].

1.1.3 Convolutional Neural Network

An example of deep learning algorithm is Convolutional Neural Networks (CNNs) which are specialized type of neural network that are specifically designed for image processing and recognition. These networks are made up of several layers of interconnected nodes, each layer tasked with identifying increasingly complex characteristics in the input data. The input layer, also known as the convolutional layer, receives the raw image and passes it through a series of operations such as convolution, pooling and activation functions. These layers use filters to convolve over the input data, identifying important features and reducing the dimensionality of the data. The final layers, known as fully connected layers, use the extracted features to make predictions about the input image [10]

Inception [11], ResNet [12], DenseNet[13], and AlexNet [14], among others, have all been suggested in recent years as solutions to issues like image classification and recognition. CNNs have the ability to learn and identify important features from images which is one of the main advantages of using this type of neural network for image processing tasks, rather than relying on hand-engineered features as in traditional machine learning approaches. These techniques are capable of analyse features from photos and classifying them with them. Because of its ability to handle massive data and extract knowledge from complicated systems, deep learning (DL) is essential to the distracted driving system [15]. In order to get the optimal result, DL searches for the network elements or features in relation to the input data by imitating how the human brain works [16].

Deep learning approaches have recently gained significant attention for visual interpretation tasks and have proven to perform exceptionally well. Particularly in image identification applications, the convolutional neural network (CNN) algorithm has made tremendous development. However, designing the optimal CNN architecture remains a challenging task, leading to the development of various architectures such as GoogLeNet (also known as Inception), AlexNet, VGGNet, and the most recent, the deep residual network (ResNet).

Recently, there has been a lot of research interest in developing a system that can classify reckless driving in order to reduce accidents and improve safety on the road. The design of a distraction detection algorithm with the potential to be used in actual vehicles with less processing time is what inspired this study. This paper focuses on d the driver activities detector based on a simple and efficient Convolutional Neural Network.

1.2 Problem Statement

The increasing prevalence of distracted driving poses a significant risk to road safety, leading to a rise in accidents and fatalities as driving involves complicated manipulations and the coordination of numerous senses. Several things can have an impact on the driver, including using a cell phone, adjusting audio equipment, smoking, drinking, eating, conversing with a passenger, or being sleepy. Therefore, it becomes vital to create driver assistance applications to lower hazards and provide early warning of distracted driving.

There is currently an ongoing development of diverse wearable technologies that incorporate integrated sensors with the purpose of monitoring the driver's physiological state, including electroencephalogram and electrocardiogram. It is worth noting that while this method has demonstrated higher reliability in contrast to behavioural and visual approaches, it is essential to highlight that these devices have the potential to cause discomfort such as restricted movement. Besides, this method can be susceptible to interference caused by bodily phenomena such as direct contact with the skin may lead to heat and increased sweating, causing distraction for the driver. [17].

Furthermore, researchers have created a variety of methods for the recognition of distracted driving behaviour based on deep learning during the past few decades. However, the cost of computing is seen as a challenge for model edge-oriented migration. Although the complexity of the model has been greatly decreased throughout the years of research, but compared to mobile phones, most automobiles' onboard electronics have significantly less powerful microcontrollers.

This project focuses on two main problems. The first problem arises from the inability to deploy trained models from traditional networks on edge devices due to the architecture's large number of parameters and the associated computation costs. The second issue pertains to the relatively lower accuracy of lightweight networks, even though they are suitable for deployment on edge devices. Parameters in neural networks are internal variables that the model learns and optimizes during training to make accurate predictions, and the parameter count determines the model's complexity and capacity to learn from data.

As a result, it is essential that assistant technologies to detects inattentive drivers be created. The model must also be compact, efficient, fast processing time due to its limited and restricted processing power of the edge devices in the car which based on a lightweight deep learning model which refers to the processes of condensing CNN models into smaller ones that yet deliver equivalent performance to the original models.

The MobileNetV2 network has been determined as an optimal choice for this application as it possesses a balance of low parameter count and a sufficient level of accuracy as shown in Table 1.1. With a total of 3,500,000 parameters which has advantages like faster training, lower memory usage, reduced overfitting and the ability to distinguish between 1000 different classifiers, the MobileNetV2 network is a lightweight network that is sufficiently robust to handle the AUC Distracted Driver Dataset.

Table 1.1: Convolutional Neural Network Comparison

Deep Neural Network	**Parameters**	**Accuracy**	**Size (MB)**	**Classification (Class)**	**Dataset**	**Application**
AlexNet	60,000,000	80.30%	240	1000	ImageNet	Image Recognition
GoogLeNet	4,000,000	93.33%	NA	1000	ImageNet	Image Recognition
VGG-16	134,000,000	92.70%	528	1000	ImageNet	Image Recognition
VGG-19	143,700,000	90.00%	549	1000	ImageNet	Image Recognition
MobileNetV1	4,200,000	70.60%	16	1000	ImageNet	Image Recognition
MobileNetV2	3,500,000	72.00%	14	1000	ImageNet	Image Recognition
MobileNetV3-Small	2,500,000	67.40%	N/A	1000	ImageNet	Image Recognition
MobileNetv3-Large	5,400,000	75.20%	N/A	1000	ImageNet	Image Recognition

1.3 Objectives

This project aims to propose an approach to detect distracted driver. The following goals must be attained in order for this project to achieve its intended goal:

1. To propose a deep learning lightweight model based on MobileNetV2 for the identification of distracted driver actions.
2. To optimise the performance of the proposed model that builds upon the MobileNetV2.
3. To compare the accuracy and processing time of the proposed method with other architecture.

1.4 Project Significance

Based on the problem stated in problem statement which is the high model complexity although results in higher accuracy, but it not applicable for real time application. On the other hand, lightweight networks have relatively low accuracy which might produce some false alert. To address this problem, the proposed method aims to improve the performance of deep learning model by adding "Squeeze and Excitation (SE) block" and modifying the layers of MobileNetV2. The algorithm utilizes the MobileNetV2, a convolutional neural network-based model, to identify patterns of distracted driving and caution driver to refocus on the road with satisfy accuracy and increase the processing time. The reason MobileNetV2 is used because of it is a light weighted architecture and it is designed for mobile and embedded devices which could be a potential to be used in actual vehicles with less processing time. Besides, the America University in Cairo (AUC) dataset is used to training the model. The dataset contains 31 drivers of various genders and races from seven different nations which is a significant advantage over other datasets.

1.5 Report Organisation

Introduction, literature reviews, methodology, results, and discussion, as well as conclusion and recommendations, comprise the five chapters of this report. The project's background, current reckless driving issues, objectives, problem statement, project significance and project outlines are all covered in Chapter 1. The objectives of this project explain the goal. The background study related to distracted driving based on deep learning is presented in Chapter 2. It reviews the current techniques for detecting distracted driver and the collision system, the various methods that have been employed, as well as the findings from earlier studies based on publications in journals, articles, and conference papers. The approach taken in this project is covered in Chapter 3. It explains the steps taken to implement the algorithm, and a flowchart showing the project's workflow is included. Chapter 4 presents the findings of the training process results, where the proposed approach is used on test data, and the outcomes are recorded. The project's conclusion, recommendations, and limitations are outlined in Chapter 5.

1.6 Summary

The problem of detecting distracted driving is crucial due to its impact on road safety and the rising number of accidents. Complex deep learning model and wearable technologies can help in monitoring driver behaviour, but some limitations exist, including computation costs, discomfort, deploying complex models on edge devices and achieving high accuracy remain challenges. A lightweight deep learning-based methods have been proposed, aims to create a compact, efficient, and accurate lightweight deep learning model, utilizing MobileNetV2, to detect distracted driver behaviours. The proposed method improves model performance and uses the AUC Distracted Driver Dataset for training.

Chapter 2

LITERATURE REVIEW

2.1 Overview

This chapter covers a discussion about a general review of distracted driver detection based on machine learning. Both techniques on detecting distracted driver will be discussed and compared based on architecture, dataset, accuracy, Frame per second (FPS) and other parameters. The previous studies for distracted driver detection to lower the car collision cases have been done and discussed in this chapter. The research gap will be identified at the end of this chapter.

2.2 Learning Approach

Various methods can identify distraction signs of the driver, with machine learning and deep learning being popular approaches. These subfields of artificial intelligence train algorithms to learn from data and make predictions without explicit programming. Machine learning focuses on pattern recognition and decision-making, categorized into supervised, unsupervised, and reinforcement learning. Deep learning, a subset of machine learning, mimics the human brain using artificial neural networks. It excels at learning hierarchical representations from complex, unstructured data like images, text, and audio. Convolutional Neural Networks (CNNs) are commonly used for image recognition tasks.

2.2.1 Machine Learning Approaches

In this thesis, distracted driving behaviours were effectively identified through the analysis of images capturing drivers inside their vehicles [18] . A range of machine learning techniques, including naive bayes, decision tree, linear SVM, softmax, and two-layer neural network, was thoroughly investigated to accurately classify the images into distinct categories representing various forms of distracted activities or safe driving. Furthermore, comprehensive discussions were provided on the essential preprocessing steps required for