International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Real-Time Driver Sentiment Analysis Using Hybrid Deep Learning Algorithm

[1]Dr Shilpa Gite, [2]Deepak Mane, [3]Mr. Ishan Doshi, [4]Ms. Prachi Mohite, [5]Ms. Bhavya Upadhyay, [6]Ms. Siddhi Joshi

Submitted: 26/09/2023 **Revised**: 17/11/2023 **Accepted**: 29/11/2023

Abstract: Everyday, over 1.3 million people are killed in traffic accidents around the world. The vast majority of these accidents are caused by human error which are caused by many things. Traffic mistakes are usually only caused by drivers who are upset while they are behind the wheel. This can result in bad driving, risky manoeuvres, or even crashes in the worst cases. There are a variety of solutions to this problem which can help people prevent these risky maneuvers and traffic accidents, but the best and most effective one is to keep the driver awake and in a safe driving condition which is the main focus of the project. In the author's framework, this research paper identifies the drivers face in the current frame at regular intervals and recognizes the drivers' emotions by snapping a photograph of the driver and using image processing techniques to extract symptoms of different emotions. The author finetune a typical pre-trained deep neural network model, CNN, on facial expression data to extract characteristics from the image of the face which is captured in a car for expression identification. The author intends to utilise the following algorithms in this research paper: VGG16, AlexNet, and VGG19. Among these, VGG16 is a widely used and straightforward Convolutional Neural Network (CNN) Architecture for ImageNet, a substantial visual database project utilised in software research for visual object recognition, boasting an accuracy of 96.63%. Furthermore, the user is permitted to load a pre-trained version of the network into VGG19.

***Keywords**: Deep neural networks, facial emotion recognition, Driver sentiment analysis, Road accidents, Emotions, FER Dataset*

1. Introduction

Advanced Driver Assistance System(ADAS) intelligent internal systems that provide the primary driver with a variety of assistance [17]. These systems could be used to deliver crucial information about traffic, upcoming road closures and blocks, levels of congestion, suggested routes to avoid congestion, etc. These systems may also be employed to detect human driver distraction and fatigue and issue cautionary alerts or to evaluate driving performance and offer recommendations [17]. These systems are capable of taking over control when a threat is detected, performing simple tasks (like cruise control) or complex manoeuvres (like overtaking and parking). The ability to communicate between various vehicles, vehicle infrastructure systems, and transportation management centres is the main benefit of using assistance systems [17] [29].

Emotions can be expressed visually, orally, or through other physiological means like hiding the feelings and keeping to ourselves. There is increasing evidence that emotional intelligence is part of what is known as intelligence [1]. Humans use different forms of dispatches such as speech, hand gestures and passions [2]. Understanding one's passions and decoded heartstrings is critical for a relevant and correct understanding. Humans use their facial expressions to express their emotions.

These emotions tell a lot about the things or conditions that the person is going through; because of the person's actions/reactions on their emotions can sometimes decide what their future. For instance, when a person is driving a vehicle, be it a car, truck, bus or any automobile, one wrong maneuver and it can end a driver's life within minutes [28]. A model-based non-rigid face tracking algorithm is used in the respective author's real-time system to extract stir features and facial emotions, which are then fed into a classifier that recognises various facial expressions.

The author's system will catch real- time video of the facial expressions of the automobile driver and with help of that it will show the driver and also other passengers in the automobile what emotion the driver is expressing through which the driver can alter his emotions and drive

[1]*Symbiosis Institute of Technology Symbiosis Centre for Applied A Symbiosis International (Deemed University), Pune,Indi shilpa.gite@sitpune.edu.in*
[2]*Department of Computer Engineering Vishwakarma Institute c Technology, Pune 411037, Maharashtra India dtmane@gmail.com*
[3]*Information technology Symbiosis Institute of Technology Symbiosi International (Deemed University), Pune, Indi ishan.doshi.btech2019@sitpune.edu.in*
[4]*Computer Science & Engineering Dept Symbiosis Institute of Technolog Symbiosis International (Deemed University), Pune, Indi prachi.mohite.btech2019@sitpune.edu.in*
[5]*Information Technology Department Symbiosis Institute of Technolog Symbiosis International (Deemed University), Pune, Indi bhavya.upadhyay.btech2019@sitpune.edu.in*
[6]*Computer Science & Engineering Department Symbiosis Institute of Technology Symbiosis International (Deemed University), Pune, India siddhi.joshi.btech2019@sitpune.edu.in*

safely and reduce kindly fatigue of the automobile driver [9]. This model which the author is aiming to build can be defined as Facial emotion recognition through deep learning. In 1978, Ekman and Freisen were among the first medical researchers to study facial features, leading to the development of the FACS [3], in which facial movements are described by AU (Action Units) movement units. This Action Units(AU) are basically the movement or the fundamental actions which is done by the muscles present in an individual's face, it has been discovered that there are 12 for the upper face and 18 are known for the lower face. The human face is divided into 44 Action Units, each encoded with one or more groups of facial muscles [8]. Automated FER (Facial Expression Recognition) is the best studied by researchers compared to other data-driven modalities performed by researchers, but it is a difficult task as everyone expresses their emotions in their own way and because of this the authors of this paper have created their own dataset so as to maximize the output of the model and collect as many as possible dataset for better proficiency [4].

There are several obstacles and challenges in this area that should not be overlooked, such as how the automobile driver is driving it, at what instance the driver is driving whether it is day time or night time, whether the automobile driver's face is covered by mask or glares and many more.

Therefore, the primary innovation or function of the programme is to accurately and flawlessly determine the driver's emotion at the end. Neural network models with a large number of layers are used to characterise deep knowledge algorithms, which have the potential to automate the point-of-birth process. Convolutional neural networks (CNNs) are a particular kind of deep neural networks (DNNs) among deep knowledge techniques. CNNs perform very well in computer vision applications because they are able to recognise patterns and retrieve attributes from pictures. By allowing "end-to-end" learning from labourers to input photos, deep-knowledge algorithms based on Facial Expression Recognition (FER) approaches lessen the dependency on recognition models and pre-processing techniques, akin to point birth styles [5]. The convolutional neural network (CNN) is the most popular deep-knowledge model among numerous others. It automatically creates a point map by passing various adulterants across the input photos. The emotional expression is identified as being tied to a particular class-based issue, and the point map is integrated with completely connected layers [5]. Recently, a number of research have coupled the deep knowledge- rested model with face characteristics to enhance the performance of facial emotion identification. [6]

The novelty of the software program is to efficiently perceive the emotion of the driving force without fault within side the result. Using DNN helps to get to know algorithms that characterize the use of neural networks whose models are constructed of large quantities of layers, it may automatize the function extraction process. Among deep neural networks getting to know algorithms, there's a selected kind of deep neural networks (DNNs) known as convolutional neural networks (CNNs), that have awesome overall performance on PC imaginative and prescient due to the fact they could discover styles and apprehend traits amongst photos. The convolutional neural network (CNN) is the most famous version amongst numerous DNN getting to know the patterns. It convolves enter photos via many filters and routinely produces a function map. The function map is blended with absolutely linked layers, and the emotional expression is diagnosed as belonging to a selected class-primarily based total output [7]. Recently, diverse research has blended facial capabilities and the DNN getting to know-primarily based totally version to reinforce the overall performance of facial features popularity.

After the pre-processing of the video, divided photos are despatched to the CNN so one can discover the emotion by the use of the shape of the human eye, mouth, the duration of the smile he has, quantity of eye is open within side the picture or body the gadget has given to CNN then it's going to test if the driving force is showing any emotions or not, it's going to additionally be beneficial to understand that if the consumer isn't always blinking properly.This mission will be used within the destiny of the automobiles as an Advanced Driver Assistance System(ADAS) that allows it to be incorporated with the built-in dashboard of automobiles like Tesla, Tata, Hyundai, Toyota, etc.[10]

Paper highlights

1. The dataset creation played a big part in this paper.
2. The architecture of this project explains the overall project as dividing videos to image frames and moving ahead pre-training models for future predictions on real-time videos.
3. Model training with 3 top Imagenet CNN models like AlexNet, VGG16, VGG19 with a highest accuracy of AlexNet with 96.63% and total parameters being 21M.

Rest of the paper is organized as follows:

In Section II, the author is discussing the Literature Review where they have studied various research papers,articles for this project. Then in Section III, the authors are putting in their methods on how they have worked the models and how they have applied their knowledge into it. It.also mentions how different models work and how they are different from each other. Coming to Section IV, the authors have mentioned the limitations they had faced making this project. In Section V, the Results have been discussed which shows how the model is working and shows us how much accuracy does all the models offer and how much data loss it gives. And in the last Section VI, they have mentioned what is ultimate conclusion of making this project and what it gives to the future in the Technological world as future Scope.

2. Literature Review

Section two is about the literature review. In this section the initial stage is to go through some previously written literature review research papers to get the knowledge and more clarification about the topic with the help of algorithms used. To compare different research papers and find the proper conclusion; the paper is subdivided into three sections.

Table 1: Table 1 shows the list and details of reffered research paper

Sr. No	Title	Author Name & year published	dataset	Dataset Links	Algorithm Used
1	Identifying the drivers of negative news with sentiment, entity and regression analysis[ ]	Fahim K Sufi, 2022	This dataset contains an aggregation engine that collects and stores global media from some sources that convert all major social media.	https://github.com/jgolbeck/fakenews/blob/master/FakeNewsData.zip	Deep Learning with CNN and RNN
2	Sentiment Analysis Based on Deep Learning:A Comparative Study	Nhan Cach Dang 1, María N. Moreno-García,2 and Fernando De la Prieta 3,2020	Gathered the dataset from various resources based on their availability and accessibility	http://help.sentiment140.com/site-functionality https://www.kaggle.com/crowdflower/twitter-airline-sentiment http://alt.qcri.org/semeval2017/ https://www.kaggle.com/c/word2vec-nlp-tutorial/data	Deep Learning CNN and RNN, machine Learning, LSTM
3	Sentiment Analysis: A Literature Survey	Pushpak Bhattacharyya, Subhabrata Mukherjee, 2013	They collected the skewed dataset and lator balanced it.	https://www.researchgate.net/figure/Comparison-of-different-classifiers-based-on-accuracy-for-ASA_fig2_360084773	Scoring Algorithm

4	Sentiment Analysis for Driver Selection in Fuzzy Capacitated Vehicle Routing Problem with Simultaneous Pick-Up and Drop in Shared Transportation	Pankaj Gupta,Mukesh Mehlawat,Anisha Khaitan,Witold Pedrycz, 2020	They combined the existing and created dataset.	https://www.researchgate.net/figure/An-example-of-fuzzy-sets-for-a-driving-feature_fig1_341524956	CNN and RNN
5	The Future of Autonomous Driving: Driver Sentiment Monitoring Analysis	Hasib Hassan 2021	Gathered the information and created the dataset based on it.	https://www.avl.com/?avlregion=GLOBAL&groupId=10138&lang=en_US	AI-based software models deliver accuracy and performance
6	Sentiment and Emotion Classification of Epidemic Related Bilingual data from social media	Muhammad Zain Ali, Kashif Javed, Ehsan ul Haq, *Anoshka Tariq* 2021	They combined the existing and created dataset.	https://www.kaggle.com/code/impratiksingh/unsupervised-learning/data	Machine Learning
7	A Neural Group-wise Sentiment Analysis Model with Data Sparsity Awareness	Deyu ,Meng ,Linhai ,Yulan 2021	Three real - world Datasets	-	Text Classification & Sentiment Analysis
8	Driving skill analysis using machine learning The full curve and curve segmented cases	Naiwala P. Chandrasiri, K. and Nawa, Akira Ishii, Shuguang Li 2012	Captured the information from internet and created their own dataset	http://drive.nissan-carwings.om/WEB/index.html	Driver Simulator and Confusion Matrix
9	Driving Behaviour Analysis Using Machine and Deep Learning Methods for Continuous	Nikolaos Peppes, Theodoros Alexakis, Evgenia Adamopoulou,Konstantinos 2021	Selected dataset from Kaggle and other helping internet websites	https://www.kaggle.com/datasets/outofskills/driving-behavior	Clustering algorithm, Logistic Regression, SVM, MLP, RNN

	Streams of Vehicular Data				
10	Detecting Human Driver Inattentive and Aggressive Driving Behavior Using Deep Learning: Recent Advances, Requirements and Open Challenges	MONAGI H. ALKINANI 1 , WAZIR ZADA KHAN 2 , 2020	Gathered readymade labelled and unlabelled dataset	https://archive.ics.uci.edu/ml/machine-learning-databases/bag-of-words/vocab.nips.txt	CNN, RNN, LSTM, DBN

1. 2.1 Driving Behaviour Analysis Using Machine and Deep Learning Methods for Continuous Streams of Vehicular Data [18].

The transportation industry is moving towards a new, more intelligent, and efficient age as a result of the fast growth in the demands for both people and products in terms of transportation as well as the development of information and communication technologies (ICT). Without a doubt, the two most crucial things for protecting the environment are lowering CO_2 emissions and minimising environmental impact. Given this, it is generally accepted that a vehicle's fuel consumption and petrol emissions are closely related to the way it is driven. It is therefore more possible than ever to propose solutions that seek to both monitor and enhance drivers' behaviour from an environmental perspective, given that modern vehicles are outfitted with sensors that can gather a variety of data, including speed, acceleration, fuel consumption, direction, etc. The method described in this study offers a comprehensive integrated platform that gathers, stores, processes, analyses, and correlates various data flows coming from automobiles [30] by combining open-source-based tools with well-known machine and deep learning [31] methods. In particular, processing and analysis are done on data streams coming from various vehicles.

In particular, they assessed the state-of-the-art of machine learning techniques employed for DBA (Database Management System) in order to identify any trends and gaps. This study concluded that the ML techniques for DBA represent a growing trend in recent years. The techniques employed for DBA can be divided into supervised or unsupervised and can be applied in real-time or not. For example, in [7–9] authors employed unsupervised learning methods while they employed supervised learning in. In the last few years, real-time assessment of the driving behaviour became a more demanding requisite since this can provide more safety to both drivers and traffic by providing the driver with feedback in real-time. For example, the authors proposed a real-time assessment of the driving behaviour by employing a support vector machine with a time window in which the features are extracted. The features are typically extracted by different sensors which can be placed inside a vehicle or at the environment. However, in the last few years we have seen a trend towards the development of low-cost sensors which can be placed in the environment. For example, the authors proposed a method for DBA by employing vehicle trajectory analysis.

OBD (ON-BOARD DIAGNOSTICS (OBD) CONNECTOR), is basically an electronic automotive system which mainly provides self-diagnosis reporting capabilities for technicians. Also, the combination of these methods can better detect the driving behaviour. This can be achieved by analyzing the data collected from the vehicle’s OBD II connector in a cloud-based environment. In order to achieve the aim of this study, as presented in the next paragraph, a detailed analysis was performed and different methods, algorithms and technologies were examined in order to propose a system which can be used for the needs of this study.This paper proposes a novel DBA framework which can collect and analyze data, and help to assess the driving behaviour of the car's driver. This framework is composed of four main components, i.e., the car's OBD II connector, the smartphone, the cloud-based environment and the data analysis. The next Section of the study presents these components and how they work together in order to collect and analyze driving data.The car's OBD II connector is a small device that is integrated in the car. Through This device, data coming from the car's sensors can be read and sent to a mobile device.

2. 2.2 Detecting Human Driver Inattentive and Aggressive Driving Behavior Using Deep Learning: Recent Advances, Requirements and Open Challenges.

In this paper, Author proposes a novel approach to detect human driving behavior using a deep learning framework. In addition to the existing work on human driving behavior detection, the author aims to address the problem of detecting human driving behavior in a more comprehensive manner. The paper proposes to use a multimodal deep learning approach to learn the features of human driving behavior from both the driver's face and the vehicle's dynamics. The proposed framework consists of two main components:

(1) a deep learning framework.

(2) A Survey on Human Driver Abnormal Driving Behavior Detection Techniques Abnormal human driving behavior is characterized by various risky driving actions and behaviors.

There are four major categories of Human Driver Abnormal Driving Behavior (HIADB) that are commonly observed and characterized in the literature: Distraction, Fatigue/Drowsiness, Alcohol Intoxication, and Aggressive Driving Behaviors. The term human driver distraction includes several types of activities that may cause a human driver to take his attention away from the road. It can be external distractions such as talking to passengers or on the phone, or internal distractions such as thinking about something else or being tired. Fatigue/Drowsiness is a major cause of human driver abnormal driving behavior. Human driver fatigue/drowsiness can be caused by a driver's lack of sleep,illness, or medication. Alcohol intoxication is also a major cause of human driver abnormal driving behavior. Alcohol intoxication can cause a human driver to become impaired in his ability to drive a vehicle safely. Aggressive driving behavior is characterized by a human driver's careless and dangerous driving actions.

RL can be used to solve the problem of learning styles.In the problem of learning styles, an agent (human learn-er) interacts with the environment (learning environment)to learn a skill (e.g. learning to read). The agent is rewarded for its performance. If the agent performs well, it is rewarded.If it performs badly, it is punished. The agent needs to learn which action to take to maximize reward. In this way, the agent can learn to perform better. The goal of the agent is to learn how to learn. In this way, the agent can learn the best learn-ing style for itself.DRL can be used to solve the problem of learning styles.In the problem of learning styles, an agent (human learn-er) interacts with the environment (learning environment)to learn a skill (e.g. learning to read). The agent is rewarded for its performance. If the agent performs well, it is rewarded.If it performs badly, it is punished. The agent needs to learn which action to take to maximize reward. In this way, the agent can learn to perform better. The goal of the agent is to learn how to learn. In this way, the agent can learn the best learn-ing style.

3. 2.3 Driving skill analysis using machine learning The full curve and curve segmented cases.

Before starting the experiment, the driver's familiarity with the simulator was tested. If the driver was not familiar with the simulator, It was asked to drive around the test track several times to get familiarized with the simulator. The total duration of simulator training was kept below 30 minutes. After the driver was familiar with the simulator, the experiment was started. The driver was instructed to drive the vehicle for 6 runs through a course which consisted of 6 different curvatures (1st 3 curves were right hand and other 3 curves were left hand) with different intersection angles. The driver was instructed to drive at the speed limit and not to change the lane during the experiment. The distance between the vehicle and the vehicle ahead in the curve was kept around 3 seconds. The total duration of a run was around 10 minutes. All the data were stored in a structured format.

3. Research and Methodology

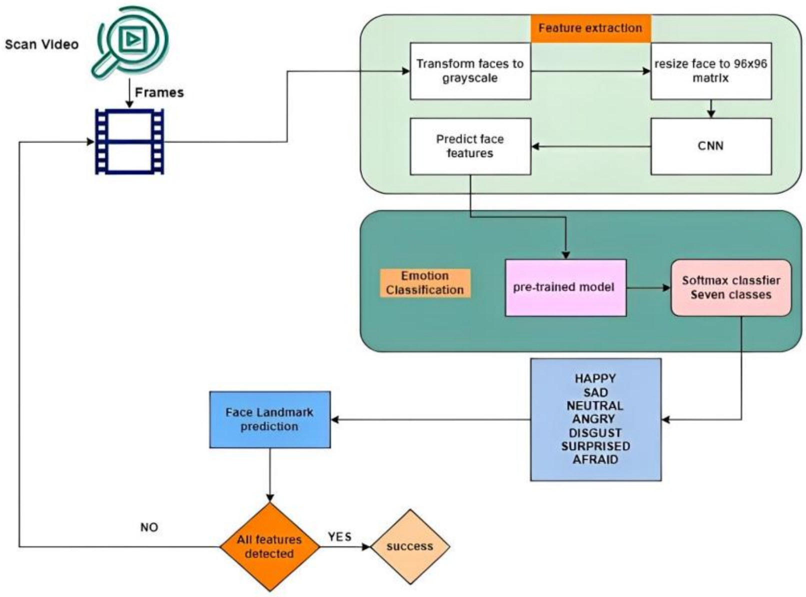


Fig 1. Block diagram for proposed system

The ImageNet database contains a pre-trained version of the network trained on more than a million images. Various image types have been represented by the network using rich feature representations. The AlexNet architecture has five convolutional layers, three max-pooling layers, two normalisation layers, two fully connected layers, and one softmax layer. Each convolutional layer consists of a nonlinear activation function and a convolutional filter.Using the pooling layers, maximum pooling is carried out.

The algorithms used in our project are, AlexNet and VGG16 both are simple and widely used CNN Architecture used for ImageNet which is a large visual database project used in visual object recognition software research, can also load a pre-trained version of the network.

The algorithms AlexNet and VGG16 analysis with respect to CNN -

AlexNet: Able to stack a pre-trained form of the organization prepared on more than a million pictures from the ImageNet database. The organization has learned wealth including representations for a wide run of pictures. Five convolutional layers, three max-pooling layers, two standardization layers, two completely associated layers, and one softmax layer make up the AlexNet engineering. Nonlinear enactment work ReLU and convolutional channels make up each convolutional layer. Max pooling is carried out utilizing the pooling layers.Because totally associated layers are show, input estimate is fixed.The input measure is ordinarily expressed as 224x224x3, be that as it may due to cushioning, it really comes out to be 227x227x3.There are 60 million parameters in AlexNet total.

VGG16: VGG16 is a simple CNN algorithm and has a large visual database project used in visual object recognition software research. The VGG16 model achieved 92.7% top-5 test accuracy in ImageNet, which is a dataset of over 14 million images belonging to 1000 classes.It will be providing the real time images for face features extraction to both the algorithms for training the model and then the classifier will classify the expression according to the sentiment of the driver in 7 classes.

It will also be identifying the eye region from the frames to detect the driver's eye using a facial landmark detection algorithm that helps to automatically identify the locations of the facial key landmark points on a facial image or from a video.

VGG16 may be a basic CNN calculation and encompasses a huge visual database extent utilized in visual object recognition software investigation. The VGG16 demonstrated 92.7% top-5 test exactness in ImageNet, which may be a dataset of over 14 million pictures having a place in 1000 classes. They will be giving the genuine time pictures for confrontation highlights extraction to both the calculations for preparing the demonstration and after that the classifier will classify the expression agreeing to the assumption of the driver in 7 classes. Also, they will be recognizing the eye locale from the outlines to distinguish the driver's eye employing a facial point of interest discovery calculation that makes a difference to consequently distinguish the areas of the facial key point of interest focused on a facial picture or from a video.

VGG19: VGG-19 is the name of a convolutional neural network with 19 layers. A pre-trained version of the network, trained on over a million photographs, is available in the ImageNet database. A pre-trained network can classify images into 1000 distinct item classes. Used just as an effective classification architecture for a wide variety of datasets; moreover, because the models' creators made them publicly accessible, other projects that are comparable may utilise them unaltered. Transfer learning may also be used in face recognition applications.

Other frameworks, like Keras, provide easily available weights that the user may modify and apply as they see suitable. There are three fully connected layers and sixteen convolutional layers in this 19-layer convolutional neural network. It has the ability to categorise images into 1000 distinct item classes. The VGG19 algorithm is trained using the ImageNet database, which contains one million images organised into 1000 categories. There is a perception that face recognition software is only used in expensive, high-security applications. The integration and greater processing power have led to the evolution of today's core technologies and a remarkable fall in equipment cost. There are currently several applications for face recognition technologies that are accurate and reasonably priced. Therefore, there are no financial or technical obstacles to growing beyond the pilot project.

In this study, several image processing techniques are used to gain information on the face state. The primary parts of the system are the smartphone hardware system and software implementations for eye tracking and real-time video pictures of the driver's face. The following goals may be accomplished by developers using their active system: Real-time eye and facial monitoring, minimising external light interference and employing the capabilities of smartphones such as camera and tiny hardware which can be mounted in the automobile which make the driver feel comfortable.

A. Videos to Frames - To split real-time web camera videos captured by our system into independent frames at the same rate using the OpenCV library. Now, the videos have been split into photo frames to create the right database for all eight categories of softmax emotions. Then this database is also used for face recognition and guessing.

B. Face Detection - After creating the correct data, use to detect faces using the face recognition algorithm and create a rectangle next to the detected surface.

C. Emotional Discovery - In the next step, fine-tune AlexNet to two facial databases to learn specific facial features, features, and improve the accuracy of recognition. Proper tuning of a pre-trained neural network with a few images is actually faster than training the network from scratch with random weights, which not only saves time but also improves accuracy. After fine-tuning our data set, in each Interest District, extract features (ROI). The last two layers of AlexNet are ready for a new split problem.

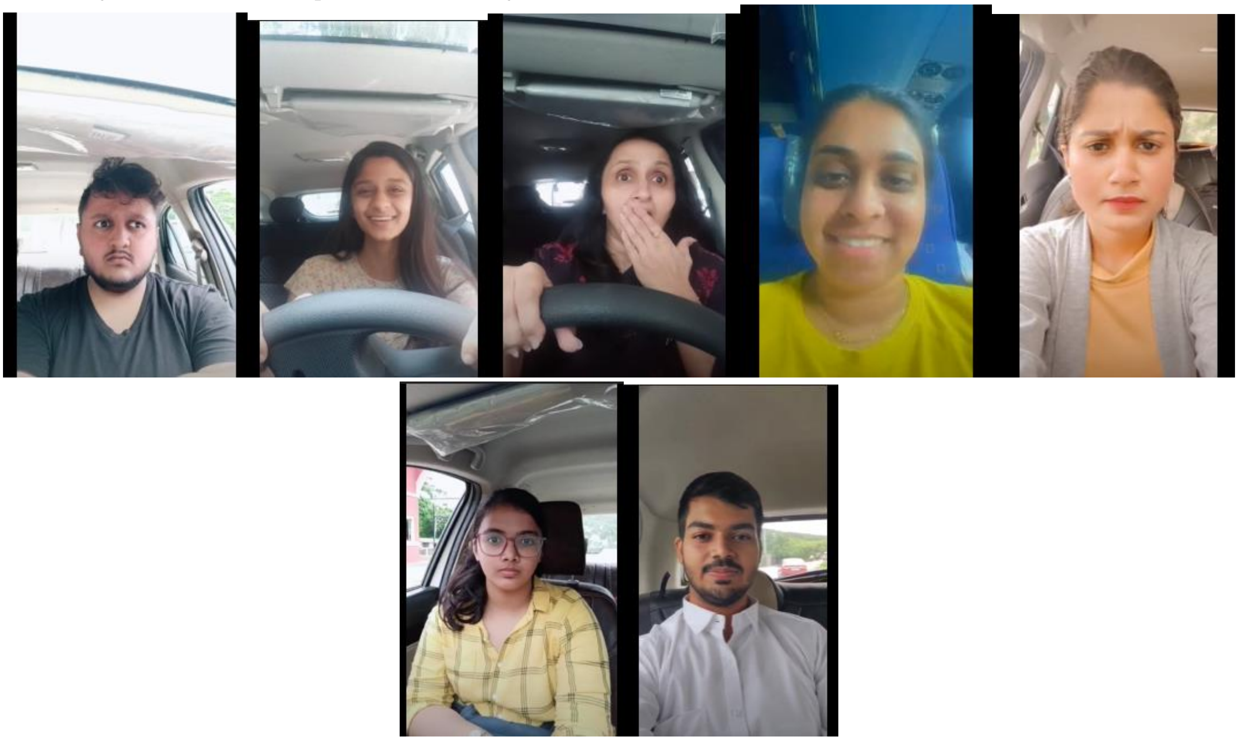


Fig. Emotion recognition Dataset samples.

Dataset Collection:

The goal to collect these images is twofold, first to create a real dataset of images labelled with emotions, videos were collected in real time from family, friends, and colleagues to create the dataset.

The dataset created consists of images for each emotion. These images were collected from the videos they created. Then, using selected images in which faces are visible and emotions are undeniably present, they end up with very heterogeneous images of different ages, races, beards, and non-beards, the same for eyeglasses. It is tried to get the image. Facial expressions can be interpreted as a mixture of emotions divided into seven parts:

To understand the best configuration for each of the pre-trained networks, four configurations were trained and tested with each one of them. For the training of these classifiers, it was necessary to select an optimizer. Initially, first optimizer Alex Net was tried, then VGG16 and VGG19.

After training and fine-tuning using a validation dataset for the best possible results, the model was tested and presented with each of the seven facial expressions.

Facial Expression	Training	Testing
Happy	3109	1600
Sad	2482	1600
Neutral	2937	1600
Angry	2562	1600
Disgust	2521	1600
Surprised	2767	1600
Afraid	2292	1600

Comparison:

AlexNet lets in for multi-GPU education with the aid of placing 1/2 of the version's neurons on one GPU and the opposite 1/2 of on any other GPU. Not simplest does this imply that a larger version may be trained, however it additionally cuts down at the education time.

The achievement of AlexNet is normally attributed to its capacity to leverage GPU for schooling and being capable of teaching those massive numbers of parameters.

VGG16 and 19 are the mainstream CNNs utilized by many researchers with ample of references and code

to aid our task in comparison with the sources of VGG11 and 13 as an alternative scarce.

The rework characteristic might not be the first-rate manner to attain statistics augmentation, however the supplied API

permits greater green statistics augmentation processing. The parameters are arbitrary considering we observed little versions withinside the end result.The fluctuation on deep CNNs doesn't leave with multiplied epoch, it has multiplied the education epoch through 3 fold the end result nevertheless doesn't converge.

4. Limitations

Model accuracy degrades when frames of attention are not clearly captured due to obstacles such as goggles or glasses with reflections. Camera operations such as automatic adjustments related to zoom and rotation are performed when performing experiments not considered. Once the eyes are located, auto zoom helps increase accuracy. Face and mouth detection accuracy will be reduced when the driver's driving force is not pointed at the camera. Many phone-based structures claimed accuracy was reduced for predicting driver movements in the fanciful and prophetic nighttime.

- A low-cost solution for detecting driving force hypervigilance in order to sell these services in developing countries. Nonetheless, methodological advancements are required to achieve boom-type results.

- Experiments and tests implementations on detecting driver fatigue were performed in a controlled environment.
- The majority of the hypervigilance structures are advanced for a next-generation phone with the assistance of smart cities' technical assistance.
- There may be no global dataset available for testing and evaluation, which could be used by the research network in the event of driver sentiment detection. For which the dataset has been created by the author which can limit the accuracy of the project.
- There are numerous sensors in and outside the vehicle to predict driving force, emotions, but statistics aggregation and communication is becoming every other hassle in modern cars.
- With an ever-increasing population, shipping and communication of data have become extremely demanding, necessitating high-computation processing from massive data sets.
- Dataset collected from the users can contain various backgrounds such as daylight or nighttime, which can lead to reduced accuracy of the model.
- Compatibility of the software with different android or iOS devices can be questioned.

5. Results

In this section we will discuss the results of the three three models that were built to detect driver's sentiment or mood swings. For facial emotion recognition CNN were discovered. In the long list of neural networks, three networks were selected on the basis of ImageNet Large Scale Visual Recognition Challenge. The selected neural networks were AlexNet, VGG16 and VGG19 proposed in(Simoyan & Zisserman, 2015) and the AlexNet proposed in (Ren Wu, Shengen Yan, Yi Shan, Qingqing Dang, Gang Sun, 2015). VGG16 and VGG19 differ only in layers of convolutional base.

Given the proper driving space in real time, the proposed system should be able to get seven sensations. Performance will also affect camera quality. Thanks to a well-designed and simple interface, the proposed system can be used by drivers. Users can use the interface step by step to achieve their goals. If the user plan meets the specified requirements, the proposed system should be available for use when required. It should be able to rediscover the failure if the application crashes suddenly and is ready for use after recovery. Concept proof will be used on Software such as Jupyter Notebook, TensorFlow 2.8.0, Python 3.8.13, Android Studio 3.6.1, Snapdragon SDK for Android and Hardware like Intel Core i7 -7500U, 8 GB RAM, Intel GMA HD 2 GB.

Convolutional layers convert inputs. Allow filter applied to the input x with kernel size n m. The connecting number of each CNN neuron input is defined by n m. The listed output output looks like this:

$$C\ (xu, v)\ =\ nX2i\ =\ -n2mX2j\ = \\ -m2fk\ (i, j)\ xu - i, v - j\ - - - - - -\ (1)$$

Many fk filters with kN can be used in input to calculate rich and very different representations of inputs[2]. Fk filters are made by dividing the weights of neighboring neurons. "This has the positive effect that, unlike conventional multilayer perceptrons, lower weights should be trained because several weights are tied together."

Large Merge: High Merge reduces the input by using the amount of function in the input xi. Face detection system using convolutional neural networks ... ") If m is the filter size, the output is:

$$M\ (xi)\ =\ max\ xi\ +\ k, i\ +\ lk \leq m2, l \leq m2k, l \\ \in N\ - - - -\ (2)$$

This layer has translation variations with respect to filter size. Fixed Line Unit: Fixed Line Unit (ReLU) is a neural network cell that uses a given activation function x output calculation:

$$R\ (x)\ =\ plural\ (0, x)\ - - - - -\ (3)$$

Compared to binary units, using these cells is much more effective than Sigmoid and still transmits more information. If the weights are implemented in the same way, half of them will be negative.This helps to create a distinctive feature. Another advantage is lower calculation costs.[2]

It is not necessary to calculate the exponential function. This feature also prevents gradient blurring as linear or zero gradients are never linear. [15]. Fully Connected Layer: A fully integrated layer, also called a multilayer perceptron, connects all the neurons in the previous layer in each neuron in its layer[2]

Allow x-input to be size k and l to be the number of fully connected neurons

layer. This results in a matrix:

$$W\ \times\ k.F\ x\ =\ \sigma\, W\ *\ x\ \ - - - - - - -\ (4)$$

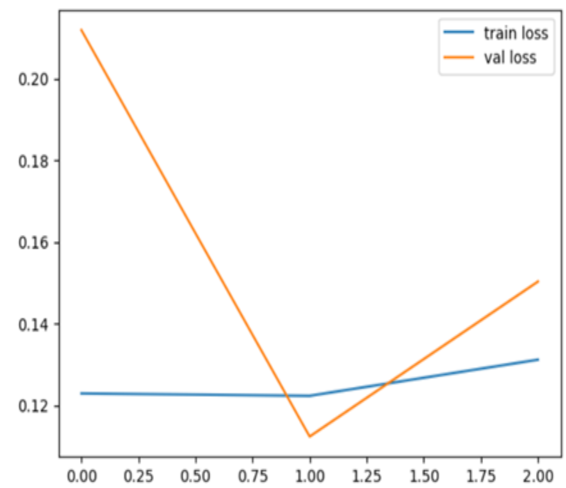
σ is the so-called activation function. In our network, σ is a copy function. Exit layer: The exit layer is a hot vector representing a given image class. SoftMax Layout: The error spreads back over the SoftMax layer. Let N be the input vector size, and SoftMax compiles the map as follows:

Sx: RN → 0.1N,

In each section 1 ≤ j ≤ N.

AlexNet and VGG19, both have about 60M parameters but there is about a 15% difference in their accuracy. But training a AlexNet requires a lot of computations (about 10 times more than that of VGG19) which means more training time and energy required. VGG16 not only has a higher number of parameters but also has an almost same accuracy as AlexNet. It takes more time to train a VGG19 with reduced accuracy. Training an AlexNet takes about the same time as training VGG19. The memory requirements are 10 times less with accuracy (about 97%)

The first model built was AlexNet, it achieved state-of-the-art recognition accuracy against all the traditional machine learning and computer vision approaches in history and so does this project as it acquired an accuracy of 96.95% and a loss of 9.90%. To understand it better the graphs of Fig 1. (train_loss,val_loss) and Fig 2. (train_accuracy, val_accuracy) were visualized.



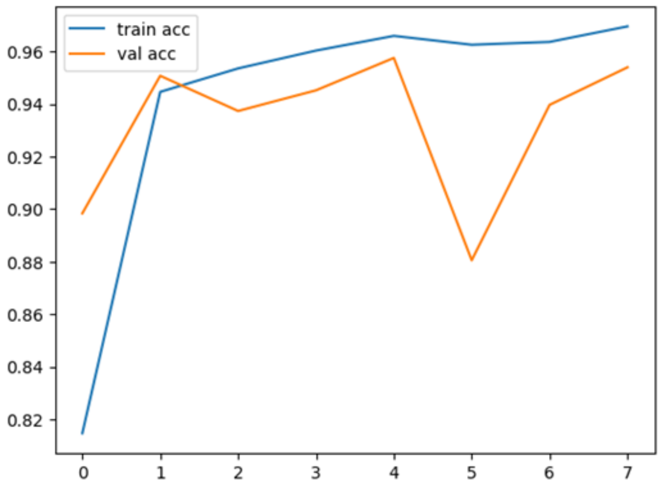
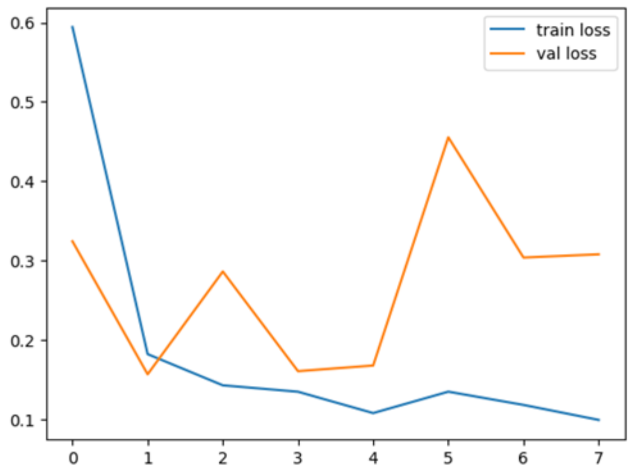


Fig 1. Graph for train_loss and val_loss data modeling

Fig 2. Graph for train_accuracy and val_accuracy modeling

The next model was VGG16 which is a part of VGGNET, a pre-trained network with the convolutional layer classifier acquired with an accuracy of 96.83% and a loss of data with 13.12%. . Visualization of loss and accuracy helps to understand the model better and also which model is better so Fig 3. Graph between train_loss and val_loss and Fig 4. Graph between train_accuracy and val_accuracy.



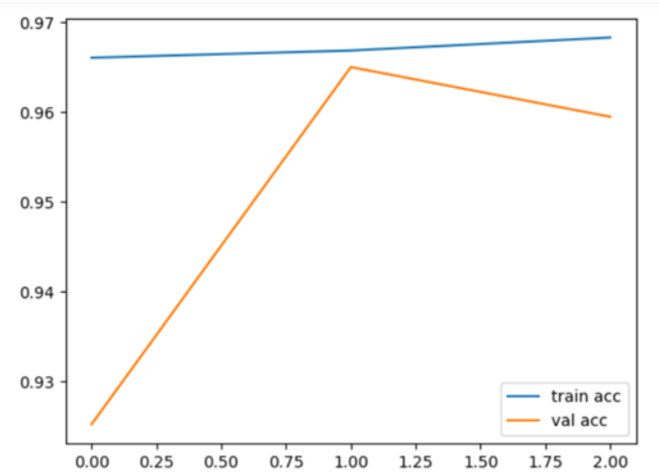
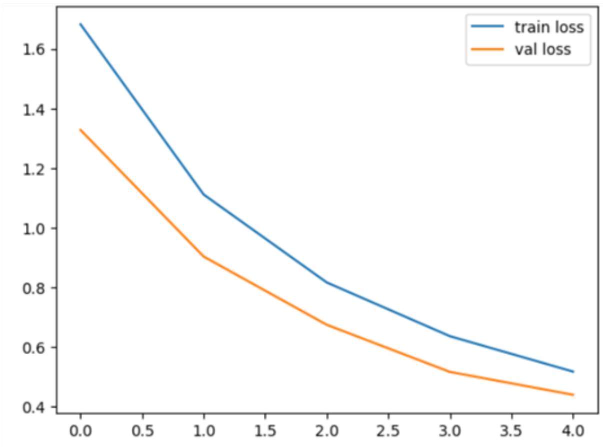


Fig 3. Graphical representation of train_loss and val_loss of VGG16 **Fig 4.** Graphical representation of train_acc and val_acc of VGG16 model.

The last model is VGG19 which is a part of VGGNET, a pre-trained network with the global average pooling layer classifier, VGG-19 had 16 convolution layers. VGG-19 is the most computationally expensive model, containing 138M weights and 15.5M MACs. This model acquired an accuracy of 82.18% with a loss of 51.67%. Fig 5.

Graphical representation of train and val loss for VGG19 model, Fig. 6 Graphical representation of train and val accuracy for VGG19 model.



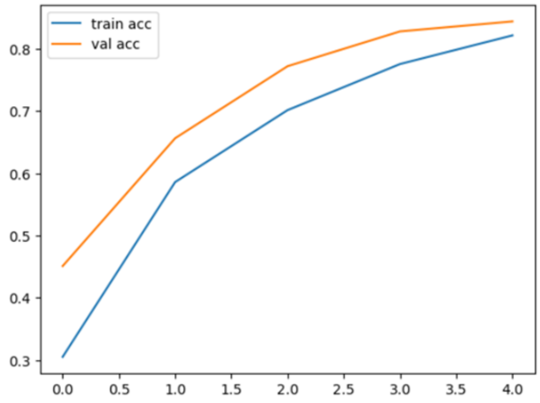


Fig. 5 Graphical representation of Train and val loss for VGG19 model. **Fig. 6** Graphical representation of Train and val accuracy for VGG19 model.

Now, a comparison between the three models was visualized to get a better understanding as to which model holds the most accuracy and has the best chances to give the project the exact model it needs.

Sr No.	Model	Loss	Accuracy
1	AlexNet	9.90	96.95
2	VGG16	13.12	96.63
3	VGG19	51.67	82.18

Analyzing the performances obtained, clearly AlexNet has more accuracy than VGG16 and VGG16 model has an overall accuracy of 96.63% and has 16 layers whereas AlexNet has only 8 layers and the only main difference between these two models is that VGG16 is much deeper than AlexNet network which means AlexNet is simpler to implement but VGG16 has more details to its modelling which means more complexity and also we can notice from Table YYY. that the accuracy of VGG16 is 96.63% and AlexNet with an accuracy of 96.65% that is almost near to VGG16 and also the time-complexity while performing AlexNet was much more than VGG16 and VGG19. For VGG19 the results were not that great as one can expect because the number of layers it contains, the reason we came across was that maybe the ImageNet recognition power was not that much with VGG19 as it was with AlexNet and VGG16 as it is quite visible in the graphs above.

While CNN modelling we faced issues like in AlexNet, the model with highest accuracy, the time complexity was way too much than compare to the other two models but it is the model with simpler architecture compared with the other two because it took approximately 5 hours to run AlexNet model with an epochs of 15 whereas the VGG16 model took 3 hrs with the same number of epochs and also the results with VGG16 were more fine-tuned than with VGG19 or AlexNet. VGG19 with the highest number of layers in architecture did not perform that well because it generates irregular results and takes a lot of time to run. So, in conclusion for this project the observations indicate that VGG16 is the appropriate model to work on the provided dataset due it's time complexity being less w.r.t other two model,that is, AlexNet and VGG19 and also it has a great accuracy of 96.63% with less loss validation value of 13.12%.

6. Conclusion and Future Scope

This Real-time working model which was created for vehicles will overcome the problems and limitations of various similar projects which have been published or proposed so far.This project will help prevent numerous accidents which happen due to mood swings which occur in drivers. Mood Swings are such a thing which cannot be controlled by people other than themselves who are going through it so to notify the drivers through the project will help the driver realise what emotion they are going through and they themselves can alter it and help prevent accidents.By developing/publishing this system, road security against drivers can be established with low cost, advanced technology and in a effective way.

In this work the lead Authors have used AlexNet, VGG16 and VGG19 image processing algorithms to obtain information about the status of the expressions captured by the face. During monitoring, the system can determine the different expressions of the face which the authors have mentioned in this paper. Surprised, Sad, Angry, Happy, Afraid, Neutral and Disgust. In the event of an error, the system can recover and correctly localise the face. The system's main components are a hardware system, a smartphone, and software implementations for real-time video images of the driver's face tracking.

The proposed system allows the developers to achieve the following: real-time face tracking, minimal external illumination interference, and use smartphone capabilities such as camera and small hardware which can be fixed in the car which make the driver feel comfortable.Their research thesis is motivated by two research objectives, which are primarily concerned with modelling drivers' sentiment analysis and the mapping of real-life equivalence is of particular interest in this study.

The Authors aim to provide additional features which will help the driver alter their emotions more efficiently by providing them with different tunes according to the emotions. Just like how meditation has different tunes for various emotions the developers will aim at providing the drivers the same.These tunes which will be deployed in the system which will directly trigger alter signals in the drivers mind and help the driver alter his emotion more conveniently.

From the standpoint of knowledge discovery the goal is to analyse user-generated data collected from video which will be then converted to frames.These frames will help get to better understanding the effects of traffic accidents on roads to comprehend the information on how to address this problem without causing annoyance to the driver. From a decision-making standpoint, the goal is to make the application Android,iOS compatible and begin collaborating with various drivers-providing services such as - Rapido,OLA and Uber. Following that work, they intend to move forward with it and bring it to the international market to reduce the number of accidents caused by driver sentiments across the world.

References

[1] N.Sebe, M.S.Lew, Y.Sun, I.Cohen, T.Gevers, T.S.Huang, "Authentic Facial Expression Analysis",Volume. 25,Issue. 12, pp. 1856-1863,2007.

[2] P.Burket, F.Trier, M.Z.Afzal, A.Dengel, M.Liwicki, "DeXpression:Deep Convolutional neural Network for Expression Recognition", arXiv:1509.05371v2[cs.CV].

[3] E.A.Clark, J.Kessinger, S.E.Duncan, M.A.Bell, J.Lahne, D.L.Gallagher, S.F.O'Keefe, "The Facial Action Coding System for Characterization of Human Affective Response to Consumer Product-Based Stimuli:A Systematic Review", Sec.Emotion Science, 2020.

[4] W.Mellouk, W.handouzi, "Recognizing Facial Expressions Using Deep Learning", Volume 175, pp. 689-694, 2020.

[5] W.Wei, Q.Jia, Y.Feng, G.Chen, "Emotion recognition Based on Weighted Fusion Strategy of Multichannel Physiological Signals", Volume 2018, Article ID 5296523.

[6] N.Saffaryazdi, S.T.Wasim, K.Dileep, A.F.Nia, S.Nanayakkara, E.Broadbent, M.Billinghurst, "Using Facial Micro-Expressions in Combination With EEG and Physiological Signals for Emotion Recognition", PMCID: PMC9275379, PMID: 35837650.

[7] S-Woo.Byun, S-Peok.Lee, "Human emotion recognition based on the weighted integration method using image sequences and acoustic features", *Multimed Tools Appl* **80**, 35871-35885(2021).

[8] Q.Ji, "Probabilistic Graphic Models for Computer Vision ", 5.4.1 Facial expression recognition, 2020.

[9] H.Xu, L.Dai, J.Fu, X.Wang, Q.Wang, "High-Quality Real Time Facial Capture Based on Single Camera", arXiv:2111.07556[cs.CV].

[10] G.M.Smith, "What is ADAS(Advanced Driver Assistance System)?", 30.Aug 2021.

[11] Iglesias, Iñaki. (2013). "Vehicle modelling for real time systems application. The virtual rolling chassis." Dyna (Bilbao). 88. 206-215. 10.6036/5177.

[12] Thimbleby, Harold. "Technology and the future of healthcare." *Journal of public health research* vol. 2,3 e28. 1 Dec. 2013, doi:10.4081/jphr.2013.e28.

[13] S. Deshmukh, M. Patwardhan, and A. Mahajan, "Survey on Real- Time Facial Expression

Recognition Techniques," IET Biom., pp. 1- 9, 2015.

[14] Z. Kowalczuk ∗ M. Czubenko ∗ T. Merta ∗, Emotion monitoring system for drivers, ∗ Gdańsk University of Technology, Narutowicza ,11/12 ,80-233 Gdańsk.

[15] Y. Gao and K.H. Leung,"Face recognition using line edge map," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 6, June 2002.

[16] N. Peppes, T.Alexakis, E.Adamopoulou, K.Demestichas,"Driving Behaviour Analysis Using Machine and Deep Learning Methods for Continuous Streams of Vehicular Data",*Sensors*. 2021; 21(14):4704. https://doi.org/10.3390/s21144704.

[17] R. Kala,"4-Advance Driver Assistance Systems",*On-Road Intelligent Vehicles,Pages 59-82,2016.*

[18] Huafei Xiao,Wenbo Li,Guanzhong Zeng,Yingzhang Wu,Jiyong Xue,Juncheng Zhang, Chengmou Li,Gang Guo, On-Road Driver Emotion Recognition Using Facial Expression, Topic Intelligent Transportation Systems,13 January 2022.

[19] D. A. Pitaloka, A. Wulandari, T. Basaruddin, and D. Y. Liliana, "Enhancing CNN with Preprocessing Stage in Automatic Emotion Recognition," Procedia Comput. Sci., vol. 116, pp. 523–529, 2017.

[20] P. Hespanha, D. J. Kriegman, and P. N. Belhumeur, "Eigenfaces vs. Fisherfaces : Recognition Using Class Specific Linear Projection," vol. 19, no. 7, pp. 711–720, 1997.

[21] Chandra Bhushan Singh, Babu Sarkar,Pushpendra Yadav, Facial Expression Recognition,SSRN(May 25,2021).

[22] Z. Kowalczuk ∗ M. Czubenko ∗ T. Merta ∗, Emotion monitoring system for drivers, ∗ Gdańsk University of Technology, Narutowicza ,11/12 ,80-233 Gdańsk.

[23] Bindu Verma and Ayesha Choudhary,A Framework for Driver Emotion Recognition using Deep Learning and Grassmann Manifolds, IEEE International Conference on Intelligent Transportation Systems (ITSC), November 2018. ("Identification System Based on Resolution Adjusted 2D Spectrogram of ...")

[24] Huafei Xiao,Wenbo Li,Guanzhong Zeng,Yingzhang Wu,Jiyong Xue,Juncheng Zhang, Chengmou Li,Gang Guo, On-Road Driver Emotion Recognition Using Facial Expression, Topic Intelligent Transportation Systems,13 January 2022.

[25] Tamanani, R.; Muresan, R.; Al-Dweik, A. Estimation of Driver Vigilance Status Using Real-Time Facial Expression and Deep Learning. IEEE Sens. Lett. 2021, 5, 1–4. [CrossRef]

[26] Abtahi, S.; Omidyeganeh, M.; Shirmohammadi, S.; Hariri, B. YawDD: Yawning Detection Dataset. IEEE DataPort 2020. [CrossRef]

[27] Lorente, M.P.S.; Lopez, E.M.; Florez, L.A.; Espino, A.L.; Martínez, J.A.I.; de Mi-guel, A.S. Explaining Deep Learning-Based Driver Models. Appl. Sci. 2021, 11, 3321. [CrossRef]

[28] Kerinab Beenu, G., Pavithra, J., A study on the prospective consumer's perception towards utility cars in Chennai city, International Journal of Applied Engineering Research, V-9, I-22, PP:7526-7531, 2014.

[29] Mane, D., Bidwe, R., Zope, B., Ranjan, N. (2022). Traffic Density Classification for Multiclass Vehicles Using Customized Convolutional Neural Network for Smart City. In: Sharma, H., Shrivastava, V., Kumari Bharti, K., Wang, L. (eds) Communication and Intelligent Systems . Lecture Notes in Networks and Systems, vol 461. Springer, Singapore. https://doi.org/10.1007/978-981-19-2130-8_78

[30] Bhujbal, Avinash, and Deepak Mane. "A survey on deep learning approaches for vehicle and number plate detection." *Int. J. Sci. Technol. Res* 8.12 (2019): 1378-1383.

[31] Khetani, Vinit, et al. "Cross-Domain Analysis of ML and DL: Evaluating their Impact in Diverse Domains." *International Journal of Intelligent Systems and Applications in Engineering* 11.7s (2023): 253-262.

[32] Raghavendra, S., Dhabliya, D., Mondal, D., Omarov, B., Sankaran, K.S., Dhablia, A., Chaudhury, S., Shabaz, M. Retracted: Development of intrusion detection system using machine learning for the analytics of Internet of Things enabled enterprises (2023) IET Communications, 17 (13), pp. 1619-1625.