Real-time drowsiness detection using computer vision and deep learning techniques

MSCDADJAN22B

Olaomopo Bandele
Student ID:x21118388

School of Computing
National College of Ireland

Supervisor: Bharat Agarwal

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Olaomopo Bandele
Student ID:	x21118388
Programme:	Data Analytics
Year:	2022
Module:	MSc Research Project
Supervisor:	Bharat Agarwal
Submission Due Date:	15 December 2022
Project Title:	Real-time drowsiness detection using computer vision and deep learning techniques
Word Count:	5356
Page Count:	16

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Olaomopo Bandele
Date:	14th December 2022

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Real-time drowsiness detection using computer vision and deep learning techniques

Olaomopo Bandele
21118388

Abstract

The number of accidents caused by drowsy driving has been worrying, thus this research concentrated on developing a system that can reliably recognize a driver's face and classify it as alert (open eyes) or drowsy (closed eyes). Convolutional neural network (CNN) and support vector machine (SVM) algorithms were trained on images of open and closed eyes of both sexes (male and female) and eye structures, as well as image of drivers wearing glasses so as to address limitations previous researchers had encountered. Pre-processing was performed to ensure that the many angles at which the drivers' eyes could be positioned could be used to train the models. The SVM had an accuracy of 92% when tested on new images, whereas the CNN had an accuracy of 50%. For real-time testing, the trained CNN model was combined with open cv and a face recognition library, and detection was performed via a webcam.

1 Introduction

In 2015, 1.25 million individuals died as a result of automobile accidents, according to data from the World Health Organization (WHO) (WHO; 2018). CDC also states that about 3,700 persons are involved in accidents involving pedestrians, buses, vehicles, motorbikes, and trucks annually (for Disease Control and Prevention.; 2018). According to (Chen et al.; 2020) by 2030, it is anticipated that the cost of non-fatal and fatal crash-related injuries will rise to $1.797 trillion. Drowsy driving typically results from insufficient sleep, but it can also be caused by untreated sleep disorders, alcohol usage, long work shift, or drugs. One of the main signs of drowsy driving includes difficulty concentrating, missing road signs, an inability to measure speeds, drifting off, sluggish response times, and falling asleep.
Drowsy driving results in thousands of injuries and deaths every year. The National Highway Traffic Safety Administration estimates that in the United States, drowsy driving contributes to 15% of all fatal accidents, 24% of nighttime accidents, and 10% of daytime accidents (Higgins et al.; 2017). This indicates that in 2016, there were roughly 6000 fatal car accidents involving drowsy drivers in the United States (Higgins et al.; 2017). According to a survey, 20% of car incidents in Canada and 20% of commercial transport accidents in the EU were related to driver drowsiness (Hussein et al.; 2021). A lot of studies have gone into designing proper and efficient safety systems to track drowsy drivers and lower the number of accidents caused by them. Volvo, Ford, Mercedes, and BMW are examples of motor companies investing largely in the study (El-Barbary et al.;

2021). Volvo developed a driver alert control whose goal is to identify steadily decreasing driving abilities (Volvo; 2018). It was designed specifically for major roads because it compares the steering data of the driver to the detected side markers painted on the road. A disadvantage of this system as mentioned on Volvo's website, was that the system can generate a misleading warning because of strong uneven, bad roads and strong winds. This drawback may even cause the driver to be misled or distracted, which is another cause of auto accidents.
After listening to several people discuss how they narrowly avoided an accident because their driver was sleepy. The severity of auto accidents and the necessity of creating a system that could reliably detect drowsy driving was understood. The goal of this research is to use deep learning algorithms and computer vision techniques to design a system that can accurately identify drowsy drivers and classify them as drowsy or awake based on their eye closure.
This research also considered distinct face characteristics such as the shape of eyes and size when training the deep learning model to address a limitation (Anjali et al.; 2016a) and (Magán et al.; 2022) encountered when designing a system that can detect drowsy drivers in real-time. Considering different drivers' face features will ensure that we have enough database to train the machine language which could reduce false errors, also, this system design can reduce the number of car accidents and lives can be saved. Driver drowsiness detection can be classified under three main categories: vehicle, behavioural and physiological signal-based characteristics. Although the three categories are important, behavioural signals have shown prominent results compared to the other two as it is non-intrusive and more accurate.

1.1 Research Question and Objectives

How accurately can computer vision and deep learning algorithms be used to detect drowsiness?
The objectives of this research includes:
1. Deep learning algorithms will be used to analyze a sequence of photos that contains a person's eyes and classify them into drowsy (closed) or awake (open).
2. Real-time drowsiness detection for drivers using a computer vision algorithm.
3. Contribute to the body of knowledge by ensuring different driver's races and eye conditions are considered when detecting drowsy drivers
The remainder of the study is organized as follows: chapter 2 discusses and contrasts existing research and literature. Chapters 3 and 4 discuss the proposed research design and implementation, chapter 5 contains the evaluation and results and chapter 6 concludes the paper.

2 Related Work

Using a driver's behaviour, vehicle patterns, or physical signals, several studies have been done to assess whether a driver is drowsy or awake through computer vision and machine/deep learning algorithms. The physiological approach uses a driver's body signals such as EGG to detect and classify drowsiness, while the vehicular approach determines drowsiness using the vehicle's speed position and surroundings. The behavioural approach involves tracking numerous indicators that can be seen from a driver's face, this includes the driver's eyes, mouth and nose length. This section discusses how machine learning algorithms and computer vision have been used with behavioural, physical and vehicular signals when detecting drowsiness.

2.1 Physiological Signals based Drowsiness Detection

The electrooculogram (EOG) and electrocardiogram (ECG), are examples of physiological indicators. These are signals, which indicate the amount of activity in the autonomic nervous system, it can be obtained from a driver with the aid of an electrode and utilized to detect if the driver is drowsy or awake. Utilizing these signals is possible because theta and delta waves increase visibility as a person transitions from an awake to a drowsy state and finally a sleep state (Wang et al.; 2020).
Three wearable electrodes were used by (Hayawi and Waleed; 2019) to track drivers' vertical eye movements and heart rates. The electrodes were placed at the back, below and on top of the ear to extract, and examine their EOG data and subsequently used a KNN algorithm to classify the signals. The adrino board was used to handle filtering, shifting and amplification at a filter band selected between 5 to 10 Hz using the signals received from the electrodes as input. The KNN classifier was applied once the signals had been normalized, and when tested, it had an accuracy of 84%. (Ma et al.; 2016) collected physiological signals of the driver's eye movement using two electrodes, one below the eyes and another on top of the driver's eye. Two electrodes were chosen to avoid situations where the electrodes become invasive and uncomfortable for the drivers. Similar to (Hayawi and Waleed; 2019), the adrino board received these signals for analysis, the researcher analysed the signals collected and forecasted the periods the driver would start feeling drowsy. A feature extraction technique that incorporated a sliding window was also used to ensure the accuracy of the ARIMA prediction model does not decline as the number of steps to be predicted increases. The ARIMA model was tested and evaluated, and the results showed that it accurately predicted when a driver would become drowsy 0.5 seconds ahead of time. According to (Wang et al.; 2015), electroencephalogram (EEG) signals and human vigilance levels are closely related. As a result, single-channel EEG signals were gathered, and the non-stationary and stationary signals were processed using the Hilber-Huang transform (HHT) to produce more precise and dependable results. They tested their proposed pre-processing method (HHT) against the continuous wavelet transform (CWT) and fast Fourier transform (FFT), and found that it performed better at differentiating between awake and sleepy states.
Although the methods discussed above are highly accurate, drivers found them to be very intrusive when driving.

2.2 Vehicular Signals based Drowsiness Detection

(Zhenhai et al.; 2017) assessed the state of awareness of a driver in a simulated environment by measuring a car's velocity angle. The angular change of the steering wheel was monitored over time using a sliding window. To avoid a misclassification due to the driver readjusting the steering wheel to balance, a threshold of 5 was defined. Two video streams to show the driver's face when drowsy and awake were also captured for the study while monitoring the velocity angle of the car. To confirm if the angular pattern can tell if a driver is drowsy, they compared the data from the car's sensors to the driver's facial expression in both sleepy and awake states and based on their results when the driver began to feel drowsy their cars velocity angle began to change quickly. (Wei et al.; 2013) gathered information on a car's vehicular pattern using two wireless sensors and combined the information using a multi-source fusion approach. One wireless sensor was installed in the driver's car to monitor lane tracking, steering wheel motion, and brake pattern. Another sensor was installed on the road to measure magnetic fields and weather. The driver state (drowsy/non-drowsy) was determined using a decision tree classifier after the signals had been pre-processed. The angle of a car's steering wheel and tracking of a car's lane were the two top key predictors and the decision tree had an accuracy rate of 80%. In a situation where we have a straight road and no curves or bumps exist, (Ashouri et al.; 2018) suggested that features asides from the angle of the velocity of a car can be used to detect drowsy drivers. The airflow direction between the car's route, car position, steering wheel angle, and yaw rate of the car, were other factors the researcher suggested that could be used. The features were collected in a simulation environment which showed that the yaw rate of the vehicle and steering wheel angle were stronger predictors over other factors as they spotted drowsy situations a minute and 10s before they occurred. Using a statistical method, (Forsman et al.; 2013) investigated the different vehicle features suggested in 87 different literature to find the most important vehicle feature in identifying drowsiness among drivers. The vehicular features taken into consideration were rate of acceleration, yaw angle, speed, the variance of the lanes, steering motions and angle. They used a PCA (principal component analysis) and varimax rotation to group the different features which showed that the major indicator of drowsiness using a vehicle can be gotten based on the drivers steering wheel pattern and the lane variance. Although using vehicle features and characteristics has shown good results when detecting drowsy drivers, there is a drawback of false drowsiness detection as different factors such as weather, wind, driving at night time and a driver having bad driving skills might impact the driver's driving style.

2.3 Behavioural Signs based Drowsiness Detection using Computer Vision

The use of computer vision and machine/deep learning algorithms to detect driver drowsiness using behavioural signs is explained here. These behavioural actions are primarily focused on the driver's facial characteristics (mouth and eyes).
(Suryawanshi and Agrawal; 2020) used the driver's head position and eye blinking as features to determine whether or not the car driver was drowsy. A Local Binary Pattern was utilized to detect the face from a live video feed and after, a Haar cascade was used to detect eyes which was the feature used to detect sleepiness in real-time. To concentrate on eye motions (Blinks), a unique eye-blinking model was created using the AdaBoost algorithm. The researcher chose to use Adaboost because when combined with Haar Cascade the accuracy of detecting blinks improved. One significant flaw in their technology was their inability to detect the driver's tiredness in low-light conditions. With the use of the IBUG 300-W dataset, (Sidaq et al.; 2021) suggested computing the eye aspect ratio (EAR) and detecting the driver's eye portion to detect drowsiness in real-time. Using the dlib package, they created an eye landmark detection method for determining the EAR. The EAR was calculated when the eyes were detected, and a threshold of 0.29 was set; if the EAR score is above the threshold, it indicates that the driver is drowsy, and an alert will ring to wake him or her. They reported that the system fails to accurately detect tiredness if the driver's facial angle is at 60 degrees during real-time testing, but they still managed to reach an accuracy of 92.57%. Similar to (Sidaq et al.; 2021), (Daengsi et al.;

2021) used EAR to identify sleepy drivers, although their architecture was designed in a way that allowed for the addition of other drowsiness indicators (Yawning, EGG). Only three participants were used to test their method, and it achieved an accuracy of 98.15%. In contrast to (Sidaq et al.; 2021), their suggested approach worked well even when the camera was positioned at specific angles.
Although a driver's EAR is effective at detecting fatigue, (Al-madani et al.; 2021) devised a highly accurate and quick detection technique that uses dlib and facial cues like the eyes and mouth to identify drowsiness in drivers. The algorithm calculates a driver's EAR to know how long a driver's eyes have been closed, and while a yawn was being detected in real-time, no calculation or threshold was used. Their suggested method was not evaluated, although possible potential detections were considered and discussed (mouth open and eyes closed, mouth closed and eyes opened and mouth closed and eyes closed). Using the dlib facial landmark extractor with fixed threshold values of 0.15 and 0.83 for the eyes and mouth, respectively, (Mohanty et al.; 2019) identified driver drowsiness in real time. On pre-recorded movies, they independently assessed the EAR and yawn detection, and the average real-time test accuracy for eyes and yawns were 82.02% and 85.44% and 93.25% and 96.71%, respectively. Additionally, they stated that their model worked better on pre-recorded footage rather than in real-time, which they attributed to various lighting circumstances.

2.4 Behavioural Signs based Drowsiness Detection using Deep learning and Machine learning Algorithms

(Lin et al.; 2018) pointed out the importance of taking into consideration the different eye problems drivers could have, thus they narrowed their study to focus on drivers who wore and did not wear glasses and used a face detector to identify tired drivers. Greyscaled images of the driver obtained from video recordings were the only data used to train the Haar ad boost algorithm. The classifier correctly recognized drowsy or non-drowsy drivers 91% when the driver wasn't wearing glasses and 95% when they wore glasses. To detect whether a driver is awake or tired, (Dwivedi et al.; 2014) looked at a range of machine/deep learning methods, which consisted of a deep CNN neural network, decision trees, LSTM (long-short-term-memory) and logistic regression. The behaviour signs observed were pupillary dilation (PD), mouth aspect ratio (MAR), and eye aspect ratio (EAR). Using photos of awake and sleepy drivers as training data, each of the machine-learning models were designed and evaluated. Based on an accuracy of 70%, the best model was the decision tree as it performed better than the other models and the most important feature was EAR. A Viola-Jones facial landmark detector and a multilayer perceptron algorithm were used by (Anjali et al.; 2016b) to detect drowsiness. The model was designed to be deployed on an embedded system therefore a multi-layer perception was used as its simple to install. The model achieved an 80.9% accuracy on the NTHU dataset although when sunglasses are worn by the driver, the accuracy of the model diminishes.
A convolutional neural network that used a feature learning representation method was proposed and investigated by (Dwivedi et al.; 2014) in the identification of drowsiness. A quick facial detection algorithm called "the Viola-Jones algorithm", was used to get the eye features from each image. When put to the test on never before seen images, their model's accuracy was 92% this is due to the fact that their dataset was limited.

3 Methodology

The goal of this project is to develop a real-time drowsiness detection system that can classify faces into drowsy or awake states and lessen traffic accidents. The development of a drowsiness detection model based on a driver's physical signals such as their eyes has shown promising results because according to previous research, detecting drowsiness via a psychological or behavioural method can be overly intrusive or result in a high number of errors. Additionally, models based on deep learning and machine learning are much more effective when used to identify sleepy drivers.
When solving a problem that requires an analytical approach like ours, the CRISP-DM (Cross Industry Standard Process for Data Mining) methodology is proven to be extremely helpful. Therefore, this study classifies drowsy drivers using the CRISP-DM data analysis method. Figure 1 illustrates the CRISP-DM process for the research methodology, which includes similar procedures taken during the study.

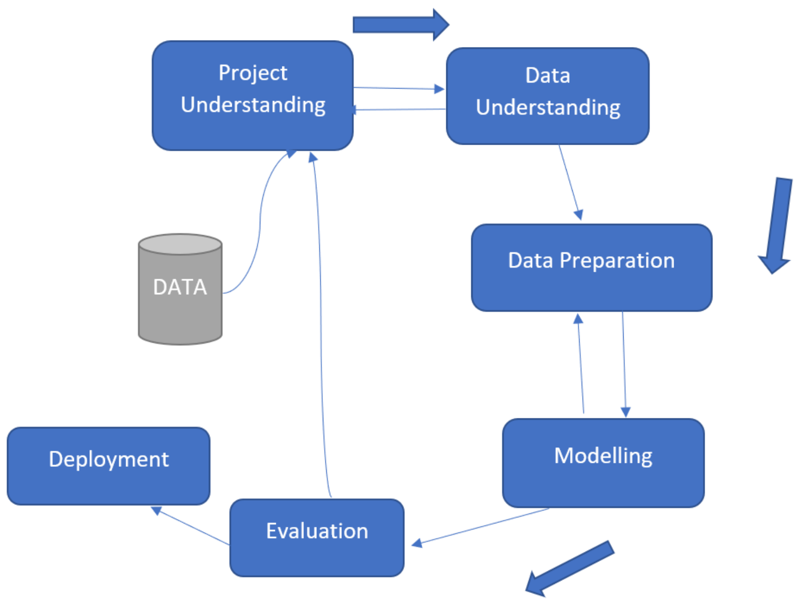

Figure 1: CRISP-DM Methodology

3.1 Project Understanding

The business goal is to develop a deep/machine learning model that uses images to determine if a driver is drowsy or not as related to the research question in chapter 1.1. Since the driver's eyes are one of the earliest markers of tiredness according to (Sidaq et al.; 2021), only images of the driver's eyes will be utilized to train the models. After training the models, the best model will then be integrated and tested in real-time using a computer vision algorithm (Open Cv) and a face recognition algorithm (dlib). By using this approach, drowsy driving will be detected earlier, helping to avert serious road accidents and catastrophizes and saving the lives of other road users.

3.2 Data Understanding

The two datasets used in this study are (Dataset1)[1] and (Dataset2)[2]. Both datasets were gotten from a publicly accessible source called "Kaggle". The obtained data was completely comprehended for the next processes, including data pre-processing, and model training and assessed to see if they satisfied project criteria. While dataset 2 featured a total of 1452 coloured eye photos (RGB images) of closed and opened eyes, dataset 1 contained a total of 48000 grey-scaled images of closed and opened people's eyes. Both datasets were merged to obtain over 49,458 images of the eyes of different people, ranging in size from 84 by 86 to 277 by 277. Although dataset 2 includes pictures of yawns and pictures without yawns, it was not used when training the model because only the driver's eyes are important and considered for this study. Images used for the study can be seen in figure 2.

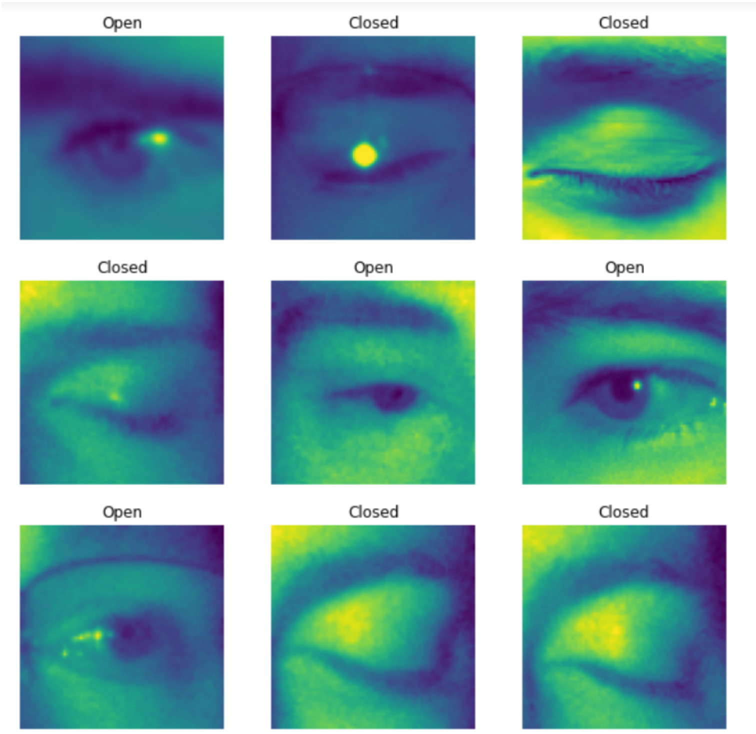

Figure 2: Eye Image

3.3 Data Preparation

Before using any of the models, the data needs to be prepared and processed properly. All images were scaled to a size of 75 by 75 because the images from the two datasets were of varying sizes, with the lowest being 86 by 86. This size was selected because, in each shot, the driver's eyes were clearly visible and would not be as affected by resizing. The RGB colored channeled images was then changed into greyscaled images because the color of the images is not crucial for drowsiness detection. Since CNN models needed a

[1]www.kaggle.com/datasets/kutaykutlu/drowsiness-detection
[2]www.kaggle.com/datasets/dheerajperumandla/drowsiness-dataset

lot of images for training, additional pre-processing methods were also carried out using the image data augmentation. Pre-processing steps included re-scaling to guarantee that every image pixel is between 0 and 1, rotating the image by 20 degrees, zooming in by 0.2, and flipping it horizontally. Afterwards, Python libraries were used to visualize the data to understand it and the pre-processing techniques as seen in figure 3.

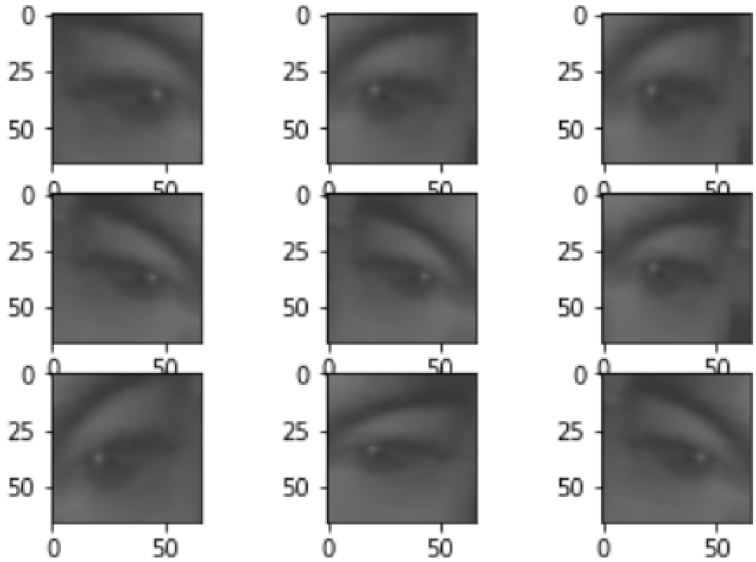

Figure 3: Pre-processed Image

3.4 Modelling

3.4.1 Support Vector Machine (SVM)

For image classification and other classification issues, the Support Vector Machine, or SVM, is a reliable linear technique. It has a wide range of applications and can handle both nonlinear and linear classifications. The SVM algorithm technique uses a hyperplane or dot matrix to divide and categorize the data. SVM only accepts 1 dimension array of images, so each image originally in 2 dimensions had its pixels extracted into a NumPy array and the NumPy array was then flattened to a 1 dimension array before being passed to the SVM model. A collection of hyperparameters can be tuned to accurately identify the hyper plane between the two classes of images; for this research, the parameters C=1, gamma = 0.001, and kernel = "rbf" were selected.

3.4.2 Convolutional Neural Networks (CNN)

Convolutional Neural Networks have a wide range of applications in the field of image and video processing. Many researchers have made attempts to create a model that can more accurately categorize sleepy drivers. The goal of this study was to create a model that can accurately identify drowsy drivers using 2D photographs of a driver's eyes. Two main blocks make up the CNN architecture, one block is for classifying the image and the other is for extraction of features and understanding the images. 4 convolutional layers with 32,64,128 and 128 and a stride of 3 by 3 were employed to build the CNN model, with the first three convolutional layers having padding set to "same". A 3 by 3 stride makes it easier to recognize and extract the distinctive and important aspects in each eye image, whether it is closed or open. It also lowers computing expenses and weight sharing. Because the eye image pixels utilized for training the model take up the entire height and width of the picture frame, padding was included to ensure that those parts of the images are captured and learned by the model. Then, for each of the three convolutional layers, the activation function was set to RELU (Rectifier Unit) a max pooling layer of 2 by 2 alongside a drop-out layer of 0.3 was added to lower the chances of the model overfitting. The second block of the CNN consists of a flattening layer, a dense and activation layer, a dropout layer of 0.3 and another dense layer of size 124, respectively. The incoming 2-D image is converted into a 1-D long vector of features by the flattened layer. The dense layer is the last layer, which acts as an output layer. The number of nodes in this layer (1), to represents the number of classes the model is classifying. The model makes use of the sigmoid activation function for the dense layer to decide which class has the highest probability and an "Adam" optimizer with a learning rate of 0.001 for classification.

3.5 Evaluation

In this stage of the CRISP-DM, the findings from the evaluation of the models including loss graphs, accuracy, precision and recall are addressed. Chapter 4 of this paper will continue the discussion of the findings in detail.

3.6 Deployment

Since it is a research project, it can be utilized as detecting software in the automobile and aviation industries, helping to avoid accidents in the future.

4 Design Specification

To carry out the tasks outlined in the objectives section in chapter 1, an efficient, economical architecture is designed. The system will function as shown in figure 4 by first loading the two different datasets from each source and then merging them to create a folder that has 49452 pictures of people with open and closed eyes. The data will then go through pre-processing, including resizing and normalization because it was obtained from several sources, which means it may not be uniform and contain several errors.

Image pre-processing involves resizing all the images to a uniform size of the same width and height and conversion of all the images into grey scale. Additionally, data augmentation techniques were also applied to the images, in particular images that are going to be used with the convolutional neural network. The augmentation includes rotating, flipping and zooming. After preparation and pre-processing of the data, the data is then split into a standard splitting ratio for image data which is 70% for training, 20% for validation and 10% for testing. Numerous research has been conducted in this field to evaluate the use of classification models for detecting drowsiness, and after a thorough assessment of the literature, CNN, and SVM models were taken into consideration for this study. Lastly, the performance of each algorithm must next be evaluated using the performance metrics listed in the evaluation stage in chapter 5.

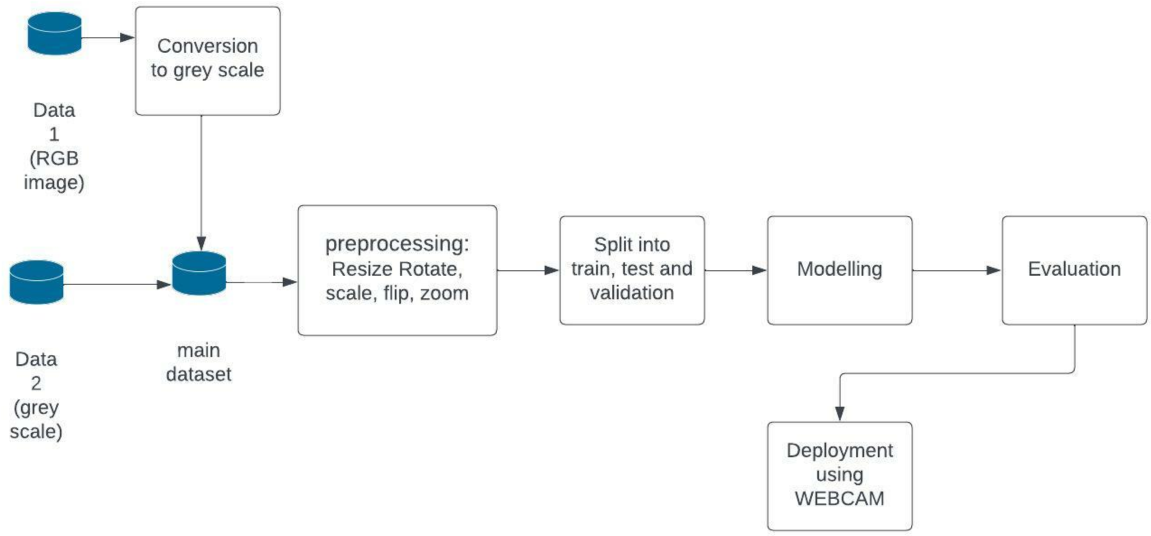

Figure 4: Project Design

4.1 Implementation

This section describes how drowsy driving design was accurately implemented using machine learning and deep learning models based on photographs of people with their eyes open and closed.
The aforementioned architecture in figure 4 was implemented in an Anaconda virtual environment utilizing a Jupiter notebook. The design is a binary classification and just requires the prediction of an open eye or a closed eye, Jupiter notebook was selected as the training time for both the CNN and SVM models was short. For image normalization, rescaling, and augmentation, Python's Numpy, Open cv, and Keras libraries were employed. The implementation of all three models makes use of the Python libraries TensorFlow and Keras.
For dataset 1 and dataset 2 respectively, the folders closed and open eyes were present in each dataset. The closed eye folder comprises photographs of closed eyes of people wearing and not wearing glasses, which was then merged with the dataset 2 closed eye folder, and similarly, the same merging of images was done with the open eye folders. The photos were split into test and train folders using a split folder library in Python. 34616 photos of both closed and opened eyes were in the train folder, 5946 were in the test folder, and 9890 were in the validation folder.
Python scripts and Opencv modules were used to carry out the necessary data processing and transformation. The same size for all the images was determined using both the Open cv library and Image augmentation library in Keras and Tensorflow. The only pre-processing done on the photos used to train the SVM models was to convert them to greyscale, resize them to a consistent size, and standardize them as one advantage of SVM is that it does not require a large number of data to train. The CNN model was trained using images that had been rotated, flipped horizontally, and zoomed because, unlike the SVM, it is easily susceptible to overfitting. Although the pre-processing of the pictures used to train the CNN models was augmented, the data used to validate wasn't and was pre-processed similar to that used to train the SVM model.

Following training and evaluation, the CNN model is saved and integrated with open cv, which provides access to the computer's webcam. A person's face is recognized through the camera using dlib's facial recognition method, and the eyes are retrieved and passed into the CNN model for classification. Eyes closed for less than the threshold of 0.5 seconds will be classified as open, but anything lingering longer will be classified as closed.

5 Evaluation

49452 photographs of individuals were used in this research, including 17308 images with closed eyes and 17308 images with open eyes. After effective pre-processing of the images, splitting data into training, validation, and testing sets, and implementation of classification algorithms (SVM and CNN), it is important to assess each of the deep learning and machine learning models performance on the validation and test image dataset using a few metrics explained below and compare them after the training of each of the models. The performance of the SVM model is assessed using Accuracy, Precision, Recall, and F1 Score, along with the length of time required to train the model. For the CNN model, evaluation is carried out using accuracy, and loss per epoch of the model on test data.

5.1 Evaluation Metrics -Accuracy of Drowsiness Detection System

The accuracy of the drowsiness detection model indicates how well it classifies the two groups (closed and open eyes). The equation for calculating accuracy in this study is given below.

$$Accuracy = \frac{Correctlypredictedclasses}{Correctlypredictedclasses + Incorrectlypredictedclasses} \tag{1}$$

5.1.1 Precision of Drowsiness Detection System

Precision describes how correctly the models classified each of the classes. To reduce the likelihood of accidents, it is preferred and important that the number of times the model mis-classified an open eye as a closed eye be reduced. The equation for calculating precision is:

$$Precision = \frac{Correctlypredictedclass(openeye)}{Correctlypredictedclass(openeye) + Incorrectlypredictedclass} \tag{2}$$

5.1.2 Recall of Drowsiness Detection System

The incidences of false negatives are measured by recall. It's crucial to reduce the number of times that the model incorrectly classified a closed eyes as open when it is indeed closed.

$$Recall = \frac{Correctlypredictedclass(closedeye)}{Correctlypredictedclass(closedeye) + Incorrectlypredictedclass} \tag{3}$$

5.2 Results

The SVM model used in this study had its hyperparameter tuned according to (Shadeed et al.; 2020), and afterwards, the model was trained. The SVM model trained for a total time of 1062s and had an accuracy of 0.9254 on the training data and 0.9213 on the testing data.

Model	Accuracy	Precision	Recall
SVM	92.3%	92.3%	92.3%

Table 1: Results of SVM model.

It is crucial to consider the number of times the model classifies a drowsy driver as awake and when an awake (opened eyes) is misclassified drowsy (closed) in order to prevent false warnings or mistakes in classification. According to Table 1, the SVM model had a precision of 0.92 and a recall score of 0.92, which is an excellent score and indicates that the model will occasionally misclassify drowsiness (less than 8% of the time).
The CNN model was trained using a large number of training and validation data for an epoch of 15 and a batch size of 100 it is evident from figure 5 that CNN achieves 99% training accuracy and 99% validation accuracy.

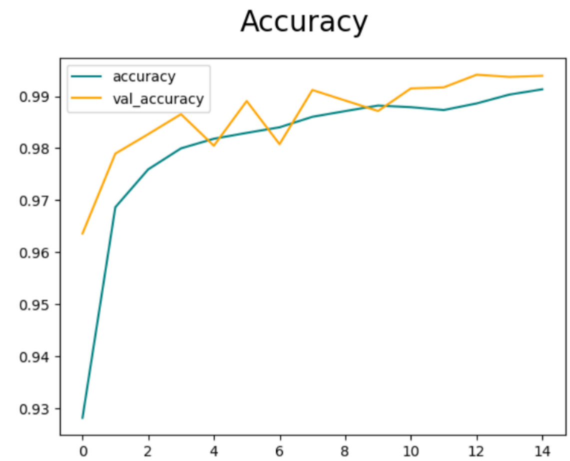

Figure 5: Accuracy

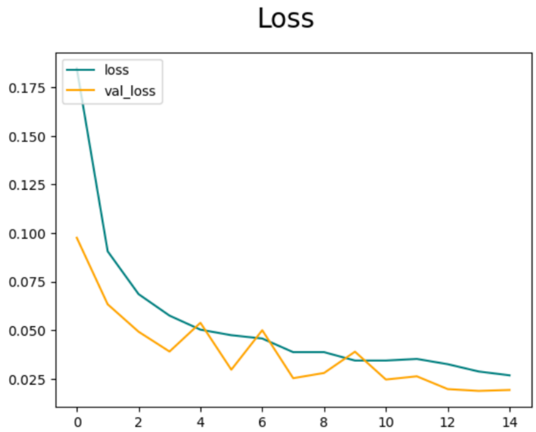

Figure 6: Loss

Although the accuracy of the validation and training sets is close, it can be seen that the model tends to overfit on validation data from the 1st epoch until progressively returning to the training data's accuracy at the 4th epoch and continuously overfitting for the remaining epoch. Furthermore, the model's validation loss shows the same pattern of overfitting as the loss is way less than that of the testing data. The cause of the constant overfitting by the model could be associated with the presence of noise in the images, as small only eye images were used to train the data and some of the images could have ended up blurry.

Model	Accuracy	Loss
CNN	50%	0.9063

Table 2: Results of CNN model.on training data

As shown in table 2, the trained CNN model had an accuracy of 50% when used on unseen photographs of closed and opened eyes. This score is significantly lower than the scores achieved on the training and validation sets, indicating that the model has been overfitting. Overfitting occurred despite the fact that the model was trained on augmented photos and the layers of the models had drop-out layers.

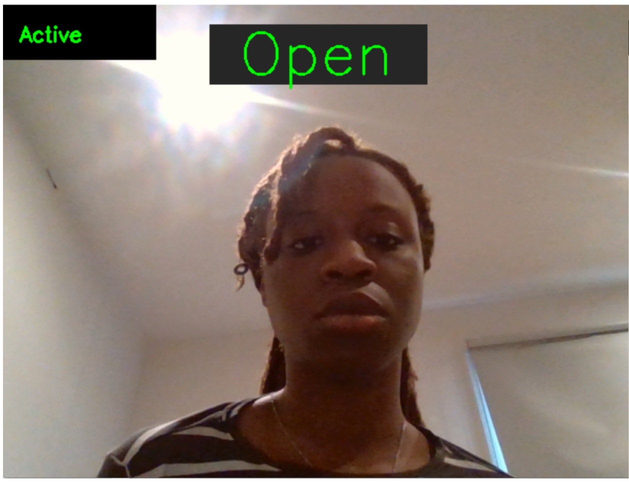

Figure 7: Opened eyes

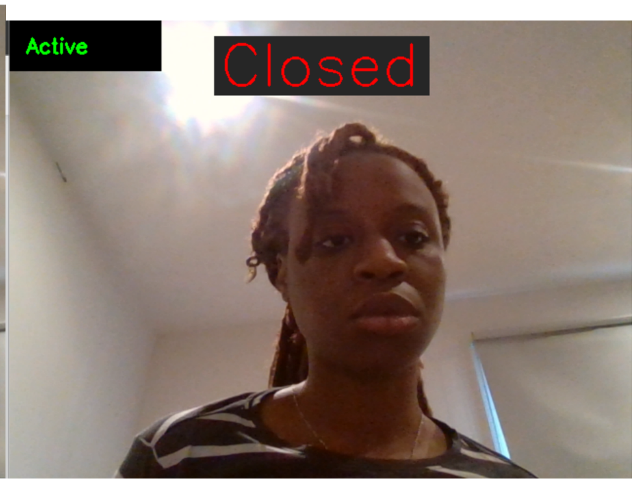

Figure 8: Mis-classified open eyes

When tested with a webcam, the model can reliably identify open eyes, but when a person isn't looking straight into the camera or a light source is blocking the camera, the model frequently mis-classified it as closed, even though the person's eyes are actually open (see figure 7 and 8)

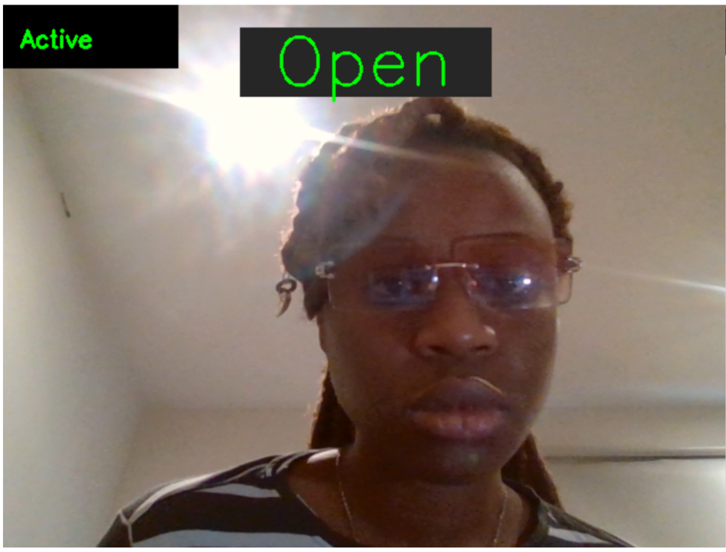

Figure 9: Opened eyes with glasses

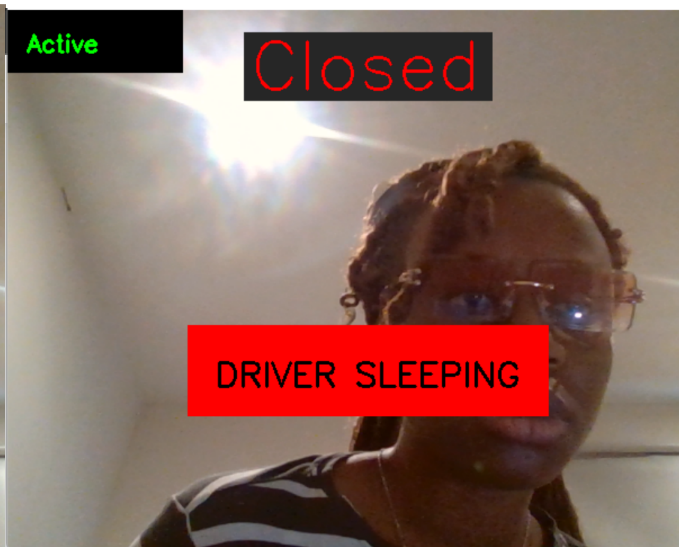

Figure 10: Mis-classified open eyes with glasse

As described in previous sections, this research sought to solve a limitation in accurate detection caused by a person wearing eyeglasses. Although the model can determine whether a person's eye is open when wearing spectacles as seen in figure 9, it also labels an open eye as closed (figure 10).

6 Conclusion and Future Work

Real-time drowsiness monitoring is crucial in today's society, not just for drivers but also for other sectors like manufacturing and aviation. This study examined the various methods that computer vision, machine learning, and deep learning algorithms can identify drowsy driving and classify them properly. One drawback shared by other studies was the inability to detect tiredness precisely if the driver was wearing glasses and if the driver's race differed. Therefore, this study trained a CNN and SVM algorithm using photos of eyes wearing glasses in low light and different races of eyes that were taken from two different Kaggle datasets.

With an accuracy of 92% from the SVM model and an accuracy of 50% for the CNN, the SVM model accurately detected driver drowsiness. Although, the SVM outperformed the CNN algorithm, the CNN model was then saved and integrated with Open cv, dlib facial recognition algorithm and used for the detection of drowsiness. The CNN model was used regardless of its low accuracy because it is easier to integrate the CNN model with open cv and detect faces as CNN accepts 2-dimensional images and SVM does not. In real-time, the CNN algorithm still did not perform accurately as the real-time detection kept classifying closed eyes and open eyes.

With an accuracy of 50% in this study, the CNN model performed poorly when compared to (Dwivedi et al.; 2014) model, which had an accuracy of 92%. The model's failure to distinguish between sleepy (eyes closed) and awake (eyes open) states in real-time is a shortcoming of this research, therefore this model can not be deployed in real-time as it can lead to inaccuracies and even distract the drivers. Regardless that the model could not differentiate closed eyes from open eyes, the model can still accurately detect a person's eyes even if they are wearing glasses or not.

In the future, a well-constructed dataset that contains a person's entire face rather than just eyes in both their awake and asleep phases can be created and utilized to train the models and detection can be done using an external camera and not a webcam since webcams are not so clear and bright. Additionally, rather than creating a CNN model from scratch, pre-trained models like Google Net can be used for training and detection in real-time.

References

Al-madani, A. M., Gaikwad, A. T., Mahale, V., Ahmed, Z. A. and Shareef, A. A. A. (2021). Real-time driver drowsiness detection based on eye movement and yawning using facial landmark, *2021 International Conference on Computer Communication and Informatics (ICCCI)*, pp. 1–4.

Anjali, K. U., Thampi, A. K., Vijayaraman, A., Francis, M. F., James, N. J. and Rajan, B. K. (2016a). Real-time nonintrusive monitoring and detection of eye blinking in view of accident prevention due to drowsiness, Institute of Electrical and Electronics Engineers Inc.

Anjali, K. U., Thampi, A. K., Vijayaraman, A., Francis, M. F., James, N. J. and Rajan, B. K. (2016b). Real-time nonintrusive monitoring and detection of eye blinking in view of accident prevention due to drowsiness, *2016 International Conference on Circuit, Power and Computing Technologies (ICCPCT)*, Institute of Electrical and Electronics Engineers Inc.

Ashouri, M. R., Nahvi, A. and Azadi, S. (2018). Time delay analysis of vehicle handling variables for near-crash detection of drowsy driving using a bus driving simulator, *2018 6th RSI International Conference on Robotics and Mechatronics (IcRoM)*, pp. 243–249.

Chen, S., Kuhn, M., Prettner, K. and Bloom, D. E. (2020). The global macroeconomic burden of road injuries: estimates and projections for 166 countries, *The Lancet Planetary Health* **3**: e390–e398.
URL: *https://www.cdc.gov/injury/features/global-road-safety/index.html*

Daengsi, T., Poonwichein, T. and Wuttidittachotti, P. (2021). Drowsiness/sleep detection and alert system: Development of a prototype using raspberry pi, *2021 2nd International Conference for Emerging Technology (INCET)*, pp. 1–4.

Dwivedi, K., Biswaranjan, K. and Sethi, A. (2014). Drowsy driver detection using representation learning, *2014 IEEE International Advance Computing Conference (IACC)*, IEEE Computer Society, pp. 995–999.

El-Barbary, M. M., Maximous, G. S., Tarek, S. and Bastawrous, H. A. (2021). Validation of driver drowsiness detection based on humantenna effect using facial features, Institute of Electrical and Electronics Engineers Inc., pp. 216–219.

for Disease Control, C. and Prevention. (2018). Global road safety — cdc.
URL: *https://www.cdc.gov/injury/features/global-road-safety/index.html*

Forsman, P. M., Vila, B. J., Short, R. A., Mott, C. G. and Dongen, H. P. V. (2013). Efficient driver drowsiness detection at moderate levels of drowsiness, *Accident Analysis and Prevention* **50**: 341–350.

Hayawi, A. A. and Waleed, J. (2019). Driver's drowsiness monitoring and alarming auto-system based on eog signals, *2019 2nd International Conference on Engineering Technology and its Applications (IICETA)*, Institute of Electrical and Electronics Engineers Inc., pp. 214–218.

Higgins, J. S., Michael, J., Austin, R., Åkerstedt, T., Dongen, H., Watson, N., Czeisler, C., Pack, A. and Rosekind, M. (2017). Asleep at the wheel: A national compendium of efforts to eliminate drowsy driving, *undefined* .

Hussein, M. K., Salman, T. M., Miry, A. H. and Subhi, M. A. (2021). Driver drowsiness detection techniques: A survey, *2021 1st Babylon International Conference on Information Technology and Science (BICITS)*, pp. 45–51.

Lin, C. Y., Chang, P., Wang, A. and Fan, C. P. (2018). Machine learning and gradient statistics based real-time driver drowsiness detection, *2018 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW)*, Institute of Electrical and Electronics Engineers Inc.

Ma, Z., Li, B. C. and Yan, Z. (2016). Wearable driver drowsiness detection using electrooculography signal, *2016 IEEE Topical Conference on Wireless Sensors and Sensor Networks (WiSNet)*, Institute of Electrical and Electronics Engineers Inc., pp. 41–43.

Magán, E., Sesmero, M. P., Alonso-Weber, J. M. and Sanchis, A. (2022). Driver drowsiness detection by applying deep learning techniques to sequences of images, *Applied Sciences (Switzerland)* **12**.

Mohanty, S., Hegde, S. V., Prasad, S. and Manikandan, J. (2019). Design of real-time drowsiness detection system using dlib, *2019 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE)*, pp. 1–4.

Shadeed, I., Alwan, J. and Abd, D. (2020). The effect of gamma value on support vector machine performance with different kernels, *International Journal of Electrical and Computer Engineering (IJECE)* **10**: 5497.

Sidaq, H., Bukhari, F. and Iqbal, W. (2021). Automatic driver alertness detection using machine learning, *2021 International Conference on Robotics and Automation in Industry (ICRAI)*, pp. 1–6.

Suryawanshi, Y. and Agrawal, S. (2020). Driver drowsiness detection system based on lbp and haar algorithm, *2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, pp. 778–783.

Volvo (2018). Driver alert control (dac)* — driver alert system — driver support — v40 2018 — volvo support.
URL: *https://www.volvocars.com/en-th/support/manuals/v40/2017w17/driver-support/driver-alert-system/driver-alert-control-dac*

Wang, C., Guragain, B., Verma, A. K., Archer, L., Majumder, S., Mohamud, A., Flaherty-Woods, E., Shapiro, G., Almashor, M., Lenne, M., Myers, R., Kuo, J., Yang, S., Wilson, N. and Tavakolian, K. (2020). Spectral analysis of eeg during microsleep events annotated via driver monitoring system to characterize drowsiness, *IEEE Transactions on Aerospace and Electronic Systems* **56**: 1346–1356.

Wang, R., Wang, Y. and Luo, C. (2015). Eeg-based real-time drowsiness detection using hilbert-huang transform, *2015 7th International Conference on Intelligent Human-Machine Systems and Cybernetics*, Vol. 1, pp. 195–198.

Wei, L., Mukhopadhyay, S. C., Jidin, R. and Chen, C. P. (2013). Multi-source information fusion for drowsy driving detection based on wireless sensor networks, *2013 Seventh International Conference on Sensing Technology (ICST)*, pp. 850–857.

WHO (2018). Road traffic injuries.
URL: *https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries*

Zhenhai, G., Dinhdat, L., Hongyu, H., Ziwen, Y. and Xinyu, W. (2017). Driver drowsiness detection based on time series analysis of steering wheel angular velocity, pp. 99–101.