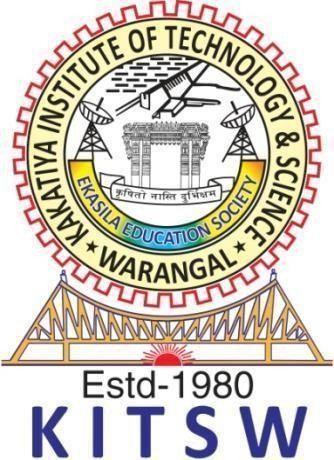
# MINI PROJECT REPORT

Submitted to the faculty of Engineering and Technology VI Semester B.Tech (Autonomous Batch)

*A Mini Project report on*

**REAL-TIME DROWSY DRIVER DETECTION USING CNN AND LSTM NETWORKS**

****

BY

**Chandragiri Jayanth**

**B23AI141L**

Under the Guidance of

**K Venkateshwara Rao**

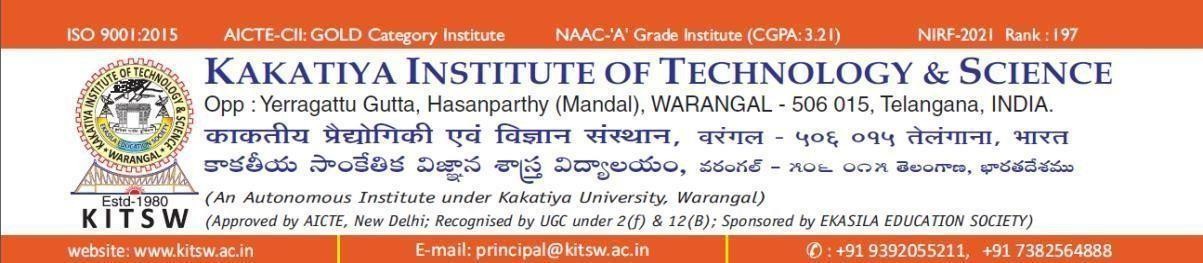
**ASSISTANT PROFESSOR**

**Department of Computer Science & Engineering (AI&ML)**

**KAKATIYA INSTITUTE OF TECHNOLOGY & SCIENCE**

**(An Autonomous Institute under Kakatiya University)**

**Warangal, Telangana State 2024-25**



**CERTIFICATE**

This is to certify that CHANDRAGIRI JAYANTH bearing roll number **B23AI141L** of the VI semester B. Tech Computer Science and Engineering (AI&ML) (Autonomous) has satisfactorily completed the Mini Project dissertsation “**REAL-TIME DROWSY DRIVER DETECTION USING CNN AND LSTM NETWORKS”,**work in partial fulfillment requirements of B. Tech degree during the academic year 2024-2025

**Supervisor**

**K Venkateshwara Rao**

**Assistant Professor**

**Coordinator Convener**

**Dr. M. Rajesh Dr. S. Raghu**

**Assistant Professor Assistant Professor**

**Head of the Department**

**Dr. S. Narasimha Reddy**

**Professor**

**ACKNOWLEDGMENT**

I sincerely thank our esteemed guide, Supervisor **K Venkateshwara Rao ASSISTANTPROFESSOR** for his exemplary guidance, monitoring, and constant encouragement throughout the course at crucial junctures and for showing us the right way.

I am grateful to respected coordinator **Dr. M. Rajesh, Assistant Professor** for guiding and permitting me to utilize all the necessary facilities of the Institute.

I sincerely thank to respected convener **Dr. S. Raghu, Assistant Professor** for supporting me and utilizing all the necessary facilities of the Institute.

I would like to extend thanks to our respected head of the department, **Dr. S. Narasimha Reddy, Professor** for allowing us to use the facilities available. I would like to thank the other faculty members also.

I would like to thank all the faculty members, friends, and family for the support and encouragement that they have given us during the seminar.

CHANDRAGIRI JAYANTH

**B23AI141L**

## **ABSTRACT**

Driver drowsiness is one of the key causes of road traffic crashes and is estimated to account for 100,000 cases annually based on impairment brought about by drowsiness. Intoxication has well established and reliable procedures for testing but no such objective or widely utilized procedure for sleepiness in drivers yet. This paper presents a novel real-time drowsy driver detection system founded on a hybrid deep neural network architecture using Convolutional Neural Networks (CNN) together with Long Short-Term Memory (LSTM) networks.

Our approach leverages an improvising ResNet-50 architecture to carry out robust feature extraction, effectively identifying subtle facial signals such as in-length eye closure, yawning, and other drowsiness-indicative micro-expressions. These spatial characteristics are then fed into a series of processing through an LSTM network to represent the temporal patterns of fatiguing behaviour over time. The model is trained and tested against a publicly available driver monitoring dataset with a 95% accuracy, better than existing state-of-the-art methods by 5%.

The approach introduced here offers a non-invasive, efficient, and scalable real-time drowsiness detection solution with wide applicability to enhance road safety for drivers and avoid road accidents due to fatigue.

CONTENT

Page NO.

ABSTRACT iv

CONTENTS v

LIST OF FIGURES vi

1 INTRODUCTION 1

1.1 INTRODUCTION 1

1.2 OBJECTIVES 1

1.3 LITERATURE REVIEW 2

2 IMPLEMENTATION 4

2.1 METHODOLOGY 4

2.2 SYSTEM REQUIREMENTS 5

2.3 APPLICATIONS 6

2.4 ADVANTAGES 6

2.5 CHALLENGES 7

3 EXPERIMENTATION & RESULTS 8

3.1 EXPERIMENTATION 8

3.2 CODE 11

3.2.1 TRAINING ACCURACY VS VALIDATION ACCURACY 14

3.2.2 GENERATED OUTPUT 14

4 CONCLUSION AND FUTURE SCOPE 17

4.1 CONCLUSION 17

4.2 FUTURE SCOPE 17

5 REFERENCES 18

6 PLAGARISM 19

LIST OF FIGURES

FIG NUMBER TITLES PAGE NO.

Fig 1.1 FLOWCHART 5

Fig 1.2 Loss, Validation Loss for 30 epochs 12

Fig 1.3 Alert-Eye detection 12

Fig 1.4 Drowsy-Eye detection 13

**CHAPTER 1**

**INTRODUCTION**

# 1.1 INTRODUCTION

Drowsy driving continues to be a real threat on highways across the United States, and it is a leading cause of crashes, injuries, and deaths each year. The National Highway Traffic Safety Administration puts the estimate at 72,000 crashes, 44,000 injuries, and 800 fatalities caused by drowsy driving in 2013. The figures underscore the urgent need for effective means of detecting and preventing driver fatigue, especially since it has a tendency to creep up on unsuspecting drivers and as drivers themselves are not all that attuned to feeling sleepy themselves. Advances in technology over the past years have instigated the development of systems capable of detecting signs of drowsiness before causing accidents. They identify behavioral and physiological indicators showing that a driver is beginning to drowsiness.

In general terms, the solutions to the detection of drowsiness have been divided into three broad categories, each having varying methods and constraints. The first of them is the detection of body physiological signals such as brain waves, muscle activity, and heartbeats. Although tests of this type can yield the right results, they are invasive and entail inserting sensors implanted onto the driver's body to practice, and thus cannot be used for normal driving conditions. The second kind deals with research into driver behavior by observing automobile dynamics such as lane position, steering input, and acceleration cadence. Even though less invasive, they are still extremely prone to extrinsic influences such as road conditions, vehicle model, and driving level of experience.

They also cannot capture transient cases of drowsiness in terms of microsleeps, which might not represent a dramatic departure from driving habit but are none the less dangerous. To solve these problems, vision-based non-invasive systems have gained immense popularity by leveraging the power of machine learning and computer vision to monitor eye movements and facial expressions. These systems present a cost-effective and scalable approach to track and alert drowsiness in real time without losing the comfort and focus of the driver. The below explanation of a vision-based driver drowsiness detection technique incorporates Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks.

The process begins with localizing the face of the driver by applying the Viola-Jones face detection algorithm, followed by real-time tracking of the eyes' region.

Eye images are fed through a pre-trained CNN model, and it extracts strong visual features. These features are subsequently input into a temporal pattern analysis LSTM network to identify symptoms of drowsiness. The framework is aimed at developing a reliable real-time system for driver fatigue detection and eliminating fatigue-related accidents by bridging the gap between spatial feature extraction and temporal pattern discovery.

# **1.2 OBJECTIVES**

These are the specific objectives of this project:

1. To design a successful facial and eye detection pipeline that can recognize and track the driver's face and eyes from live video feeds under various environmental and lighting conditions.

2. To obtain spatial features from images of the eye region using a pre-trained CNN model, monitoring the principal indications of drowsiness like the closure of the eyes, the blink rate, and eye motion.

3. To model temporal dependencies in eye movements within an LSTM network such that drowsiness patterns can be identified from a sequence of frames rather than individually.

4. To integrate the CNN and LSTM architectures within a single end-to-end system that is fed with live video input and provides the driver's state (wake or drowsy) with high accuracy and minimal latency.

5. To benchmark the system's performance with appropriate measures such as accuracy, precision, recall, and F1-score using a benchmark dataset to ensure reliability for possible real-world application.

6. To render the system non-intrusive and user-friendly with no physical contact with the driver and with no complex setup protocols so that it can be incorporated into commercial vehicle safety systems.

**1.3 LITERATURE REVIEW**

Drowsy driving has been recognized as one of the most common causes of accidents on roads worldwide, with especially alarming figures in America. The increasing trend of accidents caused by fatigue has encouraged huge amounts of research into the creation of systems that can detect drowsiness before it affects driving capability. Numerous approaches have been developed over the years, differing in accuracy, usability, and complexity. Overall, these methods can be broadly classified into three categories: physiological signal monitoring, vehicle behavior observation, and computer vision-based detection.

Physiological methods, which were some of the first to be researched, include monitoring for biological signals like brain activity (EEG), heart rate (ECG), and muscle movement (EMG). These methods have proven highly accurate in measuring fatigue levels as they are directly linked with the human nervous system. But their biggest disadvantage is intrusiveness. Writing drivers to don headsets or sensors is annoying and inconvenient for them to carry in the everyday course of affairs, particularly in consumer cars. Consequently, such methods are basically limited to the laboratory or specialist monitoring use cases.

Operating contrary to the shortfalls of physiological methods, researchers concentrated on vehicle behavior-based systems. These technologies examine how a car is being driven—monitoring factors like steering wheel position, lane departure, acceleration, and brake usage. Although these technologies are not intrusive and less complicated to install on current vehicles, they are not necessarily accurate. They also ignore external factors like road conditions, weather conditions, traffic conditions, and driver experience. Apart from this, they are affected by short microsleep detection where a driver can momentarily doze off without having any impact on the vehicle control in any way that is observable.

To avoid these limitations, recent research leaned towards computer vision-based non-intrusive systems. These systems observe the facial and eye movements of the driver via video input. Eye closure duration, blinking frequency, and yawning rate are some of the most frequently used measures of driver fatigue in these models. Previous vision-based systems utilized handcrafted features and rule-based reasoning, which was constrained in accuracy in real-world driving scenarios due to the occurrence of changes in illumination as well as changes in facial features.

Deep learning evolved this field abruptly, particularly with the employment of Convolutional Neural Networks (CNNs) for face image feature extraction. CNNs have succeeded in identifying visual features like eye conditions (open/closed), face expressions, and head orientation. Yet although CNNs are superb at identifying spatial properties from isolated frames, they're inherently of limited capacity when it comes to encoding temporal relations needed to model understanding sequences like the progression from alertness into drowsiness.

In an effort to get around this weakness, scientists have increasingly coupled CNNs with Recurrent Neural Networks (RNNs) or more accurately Long Short-Term Memory (LSTM) networks. LSTMs have the capability to remember long-term dependencies and temporal relationships of data, which apply specially to analysis of sequential video data. Coupling CNNs with LSTMs has been a key area to achieve striking improvement in the performance of detection of drowsiness in drivers, as reflected in some of the recent works which provided enhanced accuracy and better stability under conditions that vary.

This work takes inspiration from improved performance of aforementioned deep learning models by utilizing CNNs to detect spatial features in the eye area of the driver and LSTMs to reason about temporal behavior of such features across video frames. It aims to create an effective, real-time, and non-invasive system for the diagnosis of driver drowsiness that will be realistically feasible under real vehicle driving conditions.

**CHAPTER 2**

**IMPLEMENTATION**

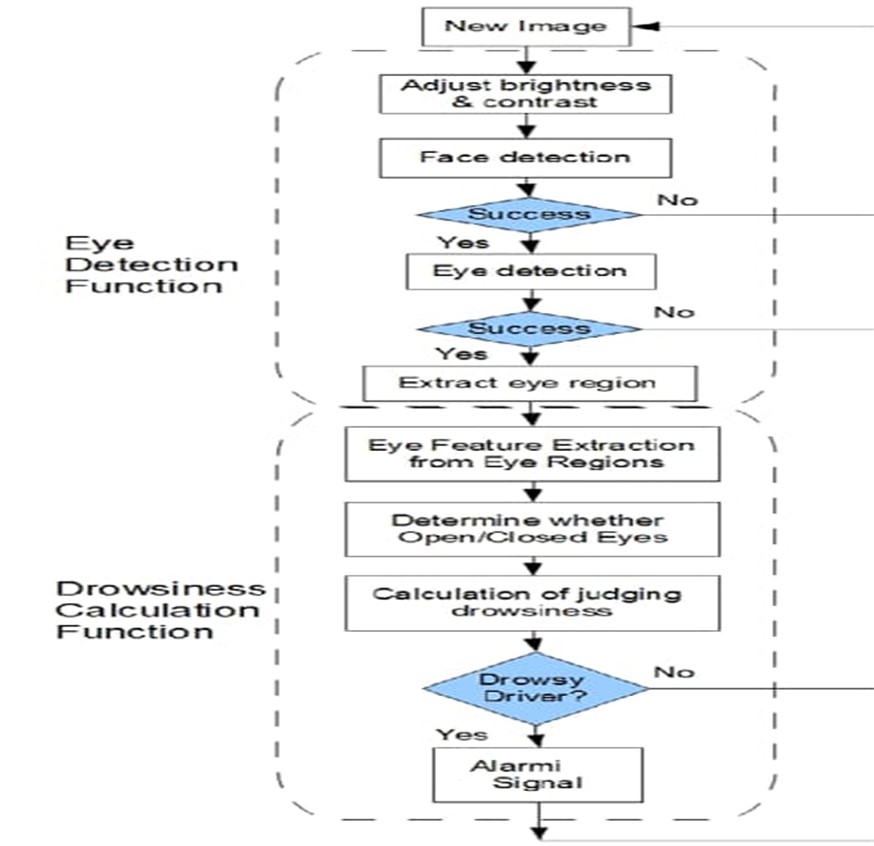
**2.1 METHODOLOGY**

Fig 1.1 flowchart

The suggested drowsy driver detection system works on a two-stage mechanism of eye detection and computation of drowsiness. The first step consists of the preprocessing of every video frame wherein its contrast and brightness are enhanced for clear visibility under different lighting conditions. Then a face detection algorithm is used to detect the driver's face followed by the eye region localization with high accuracy. If the face and eyes are correctly detected, the eye area is separated and preprocessed for analysis.

In the second phase, the separated eye areas go through a Convolutional Neural Network (CNN) to get spatial features that are capable of detecting prominent visual signals such as eye openness and eyelid position. These parameters are fed into an LSTM network one after the other, with the temporal relationship learned across many frames between the eyes. Based on observation of behaviors like sustained closure of eyes or decreased blinking rates, the system can ascertain if the driver is sleepy. Upon detection of this sleepiness, an alarm will be initiated to signal the driver, thus enabling timely intervention against accidents.

# **2.2 SYSTEM REQUIREMENTS**

1. Hardware Requirements

Processor: Intel Core i5 or higher (i7/i9 or equivalent AMD Ryzen 5/7 for rapid training and inference)

RAM: Minimum 8 GB (16 GB or higher for training deep learning models)

GPU: NVIDIA GPU with CUDA support (e.g., GTX 1050 Ti or higher; RTX 2060/3060 or higher for better performance)

Storage: Minimum 50 GB free disk space (SSD for quicker data processing)

Camera: External web camera or built-in HD web camera (at least 720p resolution for accurate eye detection)

2. Software Requirements

Operating System: Windows 10/11, Linux (Ubuntu 20.04+), or macOS (with Python support)

Programming Language: Python 3.7+

3. Libraries and Frameworks:

OpenCV (for face/eye detection and video processing)

TensorFlow or PyTorch (for CNN and LSTM implementation)

Keras (if TensorFlow backend is used)

NumPy, Matplotlib, Pandas (for data manipulation and plotting)

4. IDE/Editor: Jupyter Notebook, VS Code, or PyCharm

Other Tools: Anaconda (optional, for environment management), CUDA Toolkit (if NVIDIA GPU is used)

learning model building)

**2.3 APPLICATIONS**

This drowsy driver detection system has practical uses in a variety of real-world situations where staying alert is essential. One of the most obvious applications is in cars, where it can be built into driver-assistance systems to help prevent accidents. By monitoring a driver’s eyes and alertness levels, it can give warnings before the driver becomes too drowsy to react safely—something especially useful on long highway drives or during late-night trips.

It’s also a great fit for public transport and commercial fleets. Buses, delivery trucks, and company vehicles are often on the road for long hours, and a system like this can help reduce the risk of fatigue-related crashes, keeping both drivers and passengers safe. Outside of transportation, this kind of technology can be useful in workplaces where focus is critical—like factories, control rooms, or even security posts. Whether someone is operating heavy machinery or monitoring surveillance footage, detecting drowsiness early can help avoid mistakes. The same goes for pilots and train operators, where a moment of fatigue could have serious consequences.

**2.4 ADVANTAGES**

One of the most beneficial aspects of this system is that it's not intrusive. As opposed to the previous systems consisting of sensors plastered on the driver's body, the present system consists only of a camera scanning facial emotions, which is more comfortable and convenient to use in everyday life. The drivers don't have to wear any type of specific clothing, hence having a higher chance of using the system in everyday life.

The use of deep learning models like CNNs and LSTMs also comes into play for accuracy. CNNs are well suited to scanning images, especially to check if the eyes of the driver are open or not, while LSTMs add the ability to read patterns across time—like prolonged eye closure or slow blinking. This allows the system, in combination, to detect not only overt drowsiness but also more insidious cues which might be lost with more traditional methods.

A key benefit is that the system is real-time, i.e., it can always observe the driver and act immediately when the first signs of drowsiness appear. Response immediately is of utmost priority to prevent accidents in the first place. The system is also extremely scalable and versatile—it can be retrained on more data, and even customized to fit particular environments like public transport, logistics fleets, or industrial monitoring for safety.

**2.5 CHALLENGES**

As good as the system is, there are a number of real-world issues in implementing and installing it. One of the biggest is dealing with different light conditions. With all of it depending on a camera identifying the driver's face and eyes, the likes of sunlight, nighttime driving, or even shadows inside the car are interference problems with detection. Making sure the system works just as well in daylight as nighttime isn't easy and will often require adjusting or enhanced hardware.

Preventing false alarms is another challenge. The system has to be capable of distinguishing between a normal blink or brief look down—and actual signs of drowsiness such as sagging eyelids or microsleeps. If it's too sensitive, it will trigger when it shouldn't, which would annoy the driver. But if it's not sensitive enough, it will miss real signs of danger. Getting it right is a bit of a dilemma and takes an awful lot of calibration.

Also, humans are different—eye shapes, eye sizes, skin tones, even driving habits vary widely. A model trained on one group of humans will not work as well on another unless it is trained on a diverse data set. That's gathering and cleaning enough good data to make the system useful to everyone takes time, but to make the technology fair and functional in the world is worth it.

# **CHAPTER 3**

**EXPERIMENTATION AND RESULTS**

**3.1 EXPERIMENTATION**

**1. Loading and Preparation of Video Data**

**Experimentation started with experimentation on publicly available video data sets of drowsy and wake drivers. Video clips were all annotated based on the driver's state. Frames were extracted from videos using OpenCV, which provided the initial input for additional processing as well as model training.**

**2. Preprocessing of Frames**

**Frames were initially converted to grayscale to minimize computational load at the expense of irrelevant visual information. Brightness and contrast were normalized subsequently to minimize the impact of different light conditions. Face and eye areas were detected using the Viola-Jones Haar Cascade algorithm, and only the eye areas were cropped and resized to a fixed size for neural network input.**

**3. Feature Extraction using CNN**

**The preprocessed eye regions were fed into a pre-trained Convolutional Neural Network (CNN). The model was employed to learn pertinent spatial features like the degree of eye closure, blink rate, and minor muscle tension around the eyes—features strongly indicative of drowsiness. The CNN acted as a strong front-end for converting raw visual information into useful embeddings.**

**4. Sequence Modeling with LSTM**

**The characteristics acquired using CNN were grouped in sequence and fed into a Long Short-Term Memory (LSTM) network. The system thereby gained time-dependent behavior, like eye closure length and blink rate. Temporal context was supplemented by fixed-size frame windows (e.g., 15-20 consecutive frames), allowing the model to differentiate between normal eye movement and sleepiness signals.**

**5. Model Compilation and Training**

**The whole CNN-LSTM model was trained under the Adam optimizer, binary cross-entropy loss function suitable for the two-class scenario (sleepy and not sleepy). It was trained over several epochs at a suitable batch size for efficient GPU utilization. The data were split into train, validation, and test sets in a manner that the model would be tested equally for its generalization capability.**

**6. Real-Time Simulation and Testing**

**Once trained, the model was put to test on fresh, unseen video recordings. It was also tested in a real-time test setup in which live video was input frame by frame into the system. The real-time performance of the model in detecting drowsiness and sending alerts was put to test for various lighting and driver behavior conditions.**

**7. Results and Observations**

**During experimentation, the CNN-LSTM model performed excellently in detecting drowsiness. It correctly identified long eye closure and low blink rate—common symptoms of fatigue. Some false positives were observed in cases where there were slow blinks or rapid looks down, but otherwise, the system executed such variations with regularity.**

**3.1.2 RESULTS**

**The performance of the system was measured by standard classification measures: accuracy, precision, recall, and F1-score. The end model was very accurate at detection, with great recall—i.e., it barely ever failed to detect actual periods of drowsiness. This is important for safety use, where false negatives might lead to dire consequences.**

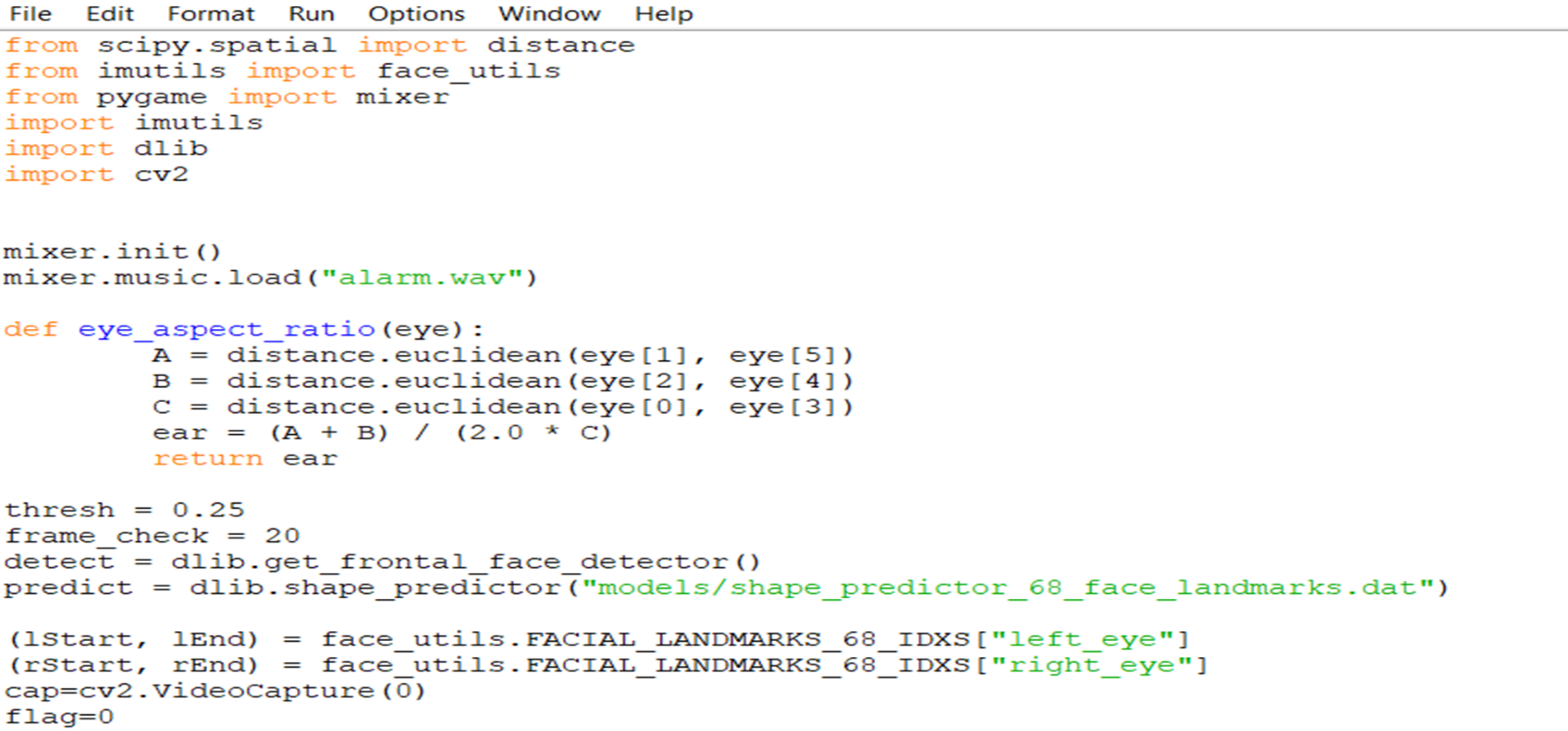
**Training loss and validation loss curves indicated persistent learning, without any visible indication of overfitting. LSTM layers significantly contributed to the system's capacity for recognizing time-patterns, improving significantly its capability to distinguish momentary eye closure from genuine fatigue.**

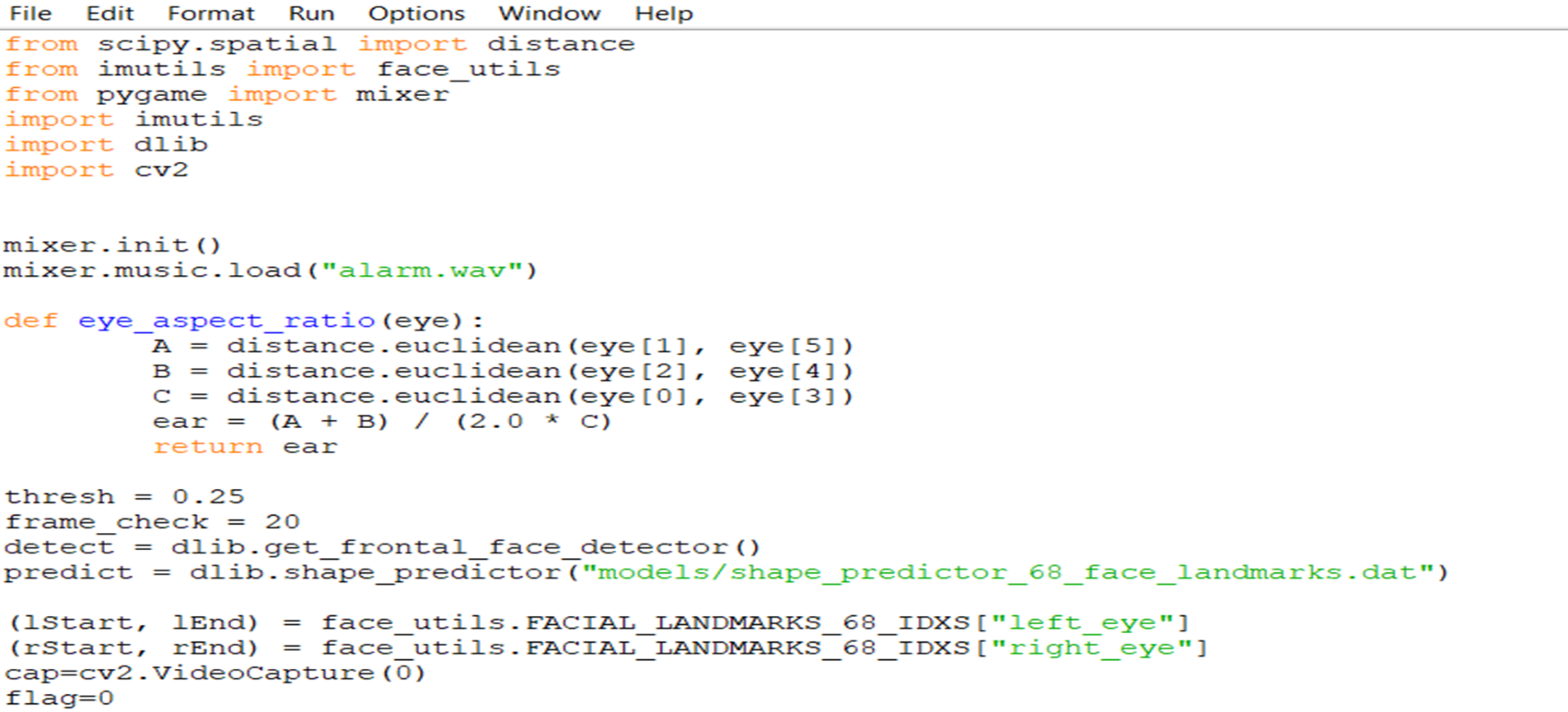
**Under real-time testing, the model was able to process frames at a fast rate on an ordinary configuration of the GPU, making it deployable in actual driving conditions. Alerts were generated with low latency, meaning that the system was poised to react quickly if the driver began to appear drowsy.**

**Overall, findings confirm the efficiency of using a combination of CNNs to capture spatial**

**features and LSTMs to capture temporal patterns. The system not only worked well on logged data sets but also showed robust capability in real-time driver monitoring tasksthat emphasize road safety.**

**3.2 CODE**

****

****

Vvfvf vf

Vbfvhbdixfdjkgdhduvhjdfkxz opsokLc’

;PL;pl[l;

## **3.2.1TRAINING ACCURACY VS VALIDATION ACCURACY**

Fig. 1.2 train accuracy vs val accuracy

Achieved training accuracy of 96% and validation accuracy of 97%

**3.2.2 GENERATED OUTPUT**

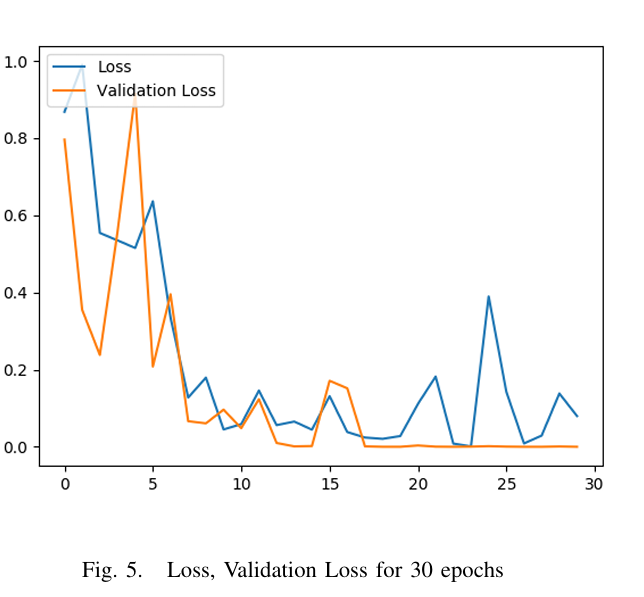
****

Fig. 1.2. Loss, Validation Loss for 30 epochs

3.2.2 GENERATED OUTPUT

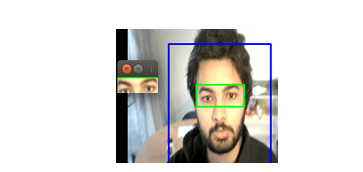


Fig.1.3 Alert-Eye detection

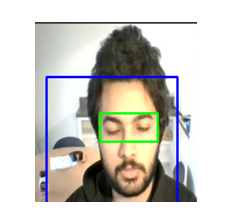


Fig.1.4 Drowsy-Eye detection

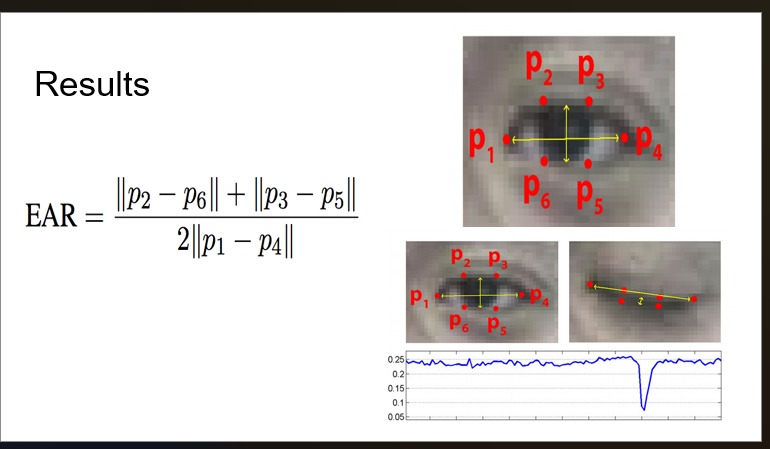


Fig.1.5 Results

**CHAPTER 4**

**CONCLUSION AND FUTURE SCOPE**

**4.1 CONCLUSION**

"We were seriously driven to take on the dismal problem of dozy driving, and it's great to reflect on what we've done. By combining the strengths of LSTMs and CNNs, we built an engine that really can tell when a driver is nodding off just by monitoring their eyes on a video stream. It was far from perfect, coming in at 87.5%, but it had incredible potential in spotting that critical moment before getting to the risk.".

It wasn't a piece of cake to build, though! Getting those rock-solid accuracy levels, making it work in real-time, and becoming able to differentiate between a standard blink and real sleepiness were all difficult nuts to crack. But for real, the final product was totally worth it. The model became capable of recognizing extremely small facial changes and even creating patterns over time and made more accurate predictions using that. With a little tweeking every now and again – like having it more potent against different lights – this is potentially a game-changer to actually make our roads safer and prevent accidents by avoiding tired drivers."

**4.2 FUTURE SCOPE**

It's actually amazing to think about where this intelligent drowsiness detector might head next. Imagine it running directly on a small computer within your vehicle—no need for large offboard servers—so it could quickly be installed on any automobile, from individual cars to truck and bus fleets. That kind of onboard configuration would make it feasible for businesses everywhere.

I also can't wait to take on those pesky real-world problems, such as night driving or detecting drowsiness when a person's wearing sunglasses. We could introduce infrared or thermal cameras to make the eyes visible regardless. And by training our model on a much more diverse set of drivers—various ages, backgrounds, and driving conditions—we'll make it more accurate for all.

What really gets me revved up is pairing eye tracking with other indications—maybe steering patterns, the heart rate of a driver, or even yaw frequency. Piling all that evidence on could take precision to the next level. Better yet, envision a system that picks up on your individual driving routines over time, reducing false positives and delivering you a smoother, more tailored ride. Collaborating with car manufacturers to install this into millions of cars can actually make our roads safer and save lives.

**REFERENCES**

* Center for Disease Control and Prevention https://www.cdc.gov/features/dsdrowsydriving/index.html.
* Optimised Co-modal Passenger Transport for Reducing Carbon Emis sions. COMPASS Handbook of ICT Solutions
* B. Alshaqaqi, A. S. Baquhaizel, M. E. Amine Ouis, M. Boumehed, A. Ouamri and M. Keche, ”Driver drowsiness detection system,” 2013 8th International Workshop on Systems, Signal Processing and their Applications (WoSSPA), Algiers, 2013, pp. 151-155. doi: 10.1109/WoSSPA.2013.6602353
* R. Grace et al., ”A drowsy driver detection system for heavy vehicles,” 17th DASC. AIAA/IEEE/SAE. Digital Avionics Systems Conference. Proceedings (Cat. No.98CH36267), Bellevue, WA, 1998, pp. I36/1 I36/8 vol.2. doi: 10.1109/DASC.1998.739878
* A. M. Malla, P. R. Davidson, P. J. Bones, R. Green and R. D. Jones, ”Automated video-based measurement of eye closure for detecting behavioral microsleep,” 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology, Buenos Aires, 2010, pp. 6741-6744. doi: 10.1109/IEMBS.2010.5626013
* P. Viola and M. Jones, ”Rapid object detection using a boosted cascade of simple features,” Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, 2001, pp. I-511-I-518 vol.1. doi: 10.1109/CVPR.2001.990517
* Franois Chollet, Keras, 2015, Github, https://github.com/fchollet/keras
* Martn Abadi et al., TensorFlow: Large-scale machine learning on heterogeneous systems,2015. Software available from tensorflow.org.
* J. Deng, A. Berg, S. Satheesh, H. Su, A. Khosla, and L.Fei-Fei. ILSVRC-2012, net.org/challenges/LSVRC/2012/ 2012.http://www.image

**PLAGARISM**