REAL TIME DROWSINESS DETECTION SYSTEM FOR STUDENT TRACKING

Project Report submitted to
Shri Ramdeobaba College of Engineering & Management, Nagpur in partial fulfillment of requirement for the award of degree of

BACHELOR OF ENGINEERING

In

COMPUTER SCIENCE AND ENGINEERING

By

Himani Dighorikar (03)
Shivam Gupta (76)
Shridhar Ashtikar (77)
Ishika Bajaj (104)

Guide

Prof. D. A. Borikar

RCOEM
Shri Ramdeobaba College of Engineering and Management, Nagpur

Department of Computer Science and Engineering

Shri Ramdeobaba College of Engineering and Management, Nagpur 440013

(An Autonomous Institute affiliated to Rashtrasant Tukadoji Maharaj Nagpur University, Nagpur)

November 2022

REAL TIME DROWSINESS DETECTION SYSTEM FOR STUDENT TRACKING

Project Report submitted to

Shri Ramdeobaba College of Engineering & Management, Nagpur in partial fulfillment of requirement for the award of degree of

BACHELOR OF ENGINEERING

In

COMPUTER SCIENCE AND ENGINEERING

By

Himani Dighorikar (03)

Shivam Gupta (76)

Shridhar Ashtikar (77)

Ishika Bajaj (104)

Guide

Prof. D. A. Borikar

Department of Computer Science and Engineering

Shri Ramdeobaba College of Engineering and Management, Nagpur 440013

(An Autonomous Institute affiliated to Rashtrasant Tukadoji Maharaj Nagpur University, Nagpur)

November 2022

SHRI RAMDEOBABA COLLEGE OF ENGINEERING & MANAGEMENT, NAGPUR

(An Autonomous Institute affiliated to Rashtrasant Tukadoji Maharaj Nagpur University, Nagpur)

Department of Computer Science and Engineering

CERTIFICATE

This is to certify that the Thesis on **“Real Time Drowsiness Detection System for Student Tracking”** is a bonafide work of Himani Dighorikar, Ishika Bajaj, Shivam Gupta and Shridhar Ashtikar submitted to the Rashtrasant Tukadoji Maharaj Nagpur University, Nagpur in partial fulfillment of the award of a Degree of Bachelor of Engineering. It has been carried out at the Department of Computer Science and Engineering, Shri Ramdeobaba College of Engineering and Management, Nagpur during the academic year 2022-23.

Date:

Place: Nagpur

Prof. D. A. Borikar
(Project Guide)
Department of Computer
Science and Engineering

Dr Avinash Agrawal
(H.O.D)
Department of Computer
Science and Engineering

Dr. R. S. Pande
(Principal)

DECLARATION

I, hereby declare that the thesis "REAL TIME DROWSINESS DETECTION SYSTEM FOR STUDENT TRACKING" submitted herein, has been carried out in the Department of Computer Science and Engineering of Shri Ramdeobaba College of Engineering & Management, Nagpur. The work is original and has not been submitted earlier as a whole or part for the award of any degree / diploma at this or any other Institute / University.

Date:

Place: Nagpur

Name of the Student	Roll No.	Signature
Himani Manoj Dighorikar	03	
Shivam Gupta	76	
Shridhar Ashtikar	77	
Ishika Bajaj	104	

APPROVAL SHEET

This report entitled

"REAL TIME DROWSINESS DETECTION SYSTEM FOR STUDENT TRACKING"

By

Himani Dighorikar

Ishika Bajaj

Shivam Gupta

Shridhar Ashtikar

is approved for the degree of Bachelor of Engineering.

Name & signature of Supervisor(s)   Name & signature of External Examiner(s)

Name & signature of HOD

Date:

Place: Nagpur

ACKNOWLEDGEMENT

The project is a combined effort of a group of individuals who synergize to contribute towards the desired objectives. Apart from the efforts by us, the success of the project shares an equal proportion on the engagement and guidance of many others. We take this opportunity to express our gratitude to the people who have been instrumental in the successful completion of this project.

We would like to show our greatest appreciation towards Prof. D.A Borikar for providing us with the facilities for completing this project and for his constant guidance.

We would like to express our deepest gratitude to Dr. Avinash Agrawal, Head, Department of Computer Science and Engineering, RCOEM, Nagpur for providing us the opportunity to embark on this project.

We are thankful to the Centre of Excellence in Artificial Intelligence and Deep learning (AI & DL) of Department of Computer Science and Engineering, RCOEM Nagpur for permitting us to run our model on the NVIDIA DGX DL Workstation. The resource has been of tremendous help in testing the model in lesser duration.

Finally, extend our gratitude to all the faculty members of the CSE department who have always been so supportive and providing resources needed in our project development. We are very grateful to all of them who have unconditionally supported us throughout the project.

Date:

— Projectees

ABSTRACT

Many researchers' studies on fatigue detection have taken curves towards different technologies. Machine vision based driver fatigue detection systems are used to prevent accidents and improve safety on roads. To design an alerting system for the students this will use real time video of a person to capture the drowsiness level if feeling fatigue. A chrome extension is designed which if enabled, will start the webcam and track the person. The frame per second (FPS) will be set and an alert will be generated if a continuous set of frames are detected as drowsy. The conventional methods cannot capture complex expressions, however with the origin of Deep learning models a substantial research is done on detection of states of a person in real time. The system works adequately under the natural lighting conditions and no matter the use of any accessories like cap, glasses, etc. The system is implemented using YOLOv5 models (You look only once) which is an extremely fast and accurate detection model.

The major aim of this project is to develop a drowsiness detection system by monitoring the eyes; it is believed that the symptoms of fatigue can be detected early enough to avoid sleep to achieve targets. In such a case when drowsiness is detected, a warning signal is issued to alert the person. This detection system provides a noncontact technique for judging different levels of alertness and facilitates early detection of a decline in alertness. In such a case when fatigue is detected, a warning signal is issued to alert the person.

Keywords: Deep learning, YOLO models, Convolutional Neural Network, Pytorch, OpenCV, Drowsiness detection, Online Monitoring

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	**Expansion**
CNN	Convolutional Neural Network
FPS	Frame per seconds
YOLO	You Look Only Once
AI	Artificial Intelligence
DL	Deep Learning
IOU	Intersection over Union
AOI	Area of Interest
mAP	Mean Average Precision
CSP	Cross Stage Partial

CHAPTER 1

INTRODUCTION

Online studying is gaining currency these days. People work online for hours and there may be a situation where they need to track themselves. Online courses, work has got drive after the pandemic and the drowsiness detection system would help such people to get alerts if feeling drowsy. The model uses YOLO architecture and trains itself on various states of faces and learns the features of drowsiness.

1.1 Problem Definition

Design a system based on Deep Learning approaches that detects the real time drowsiness state of a person and gives progressive alerts to ensure productivity during working hours. Drowsiness detection is a critical issue and needs to be as accurate as it can be. An approach that overpowers the traditional drowsiness system is implemented.

1.2 Motivation

The pandemic crisis has changed our lifestyles and accelerated the pace of digital transformation. How this situation has disrupted our way of life and kept everything on lockdown. Online studies are the only thing students can undertake because of the lockdown. Online courses allow anyone to access all study materials and classes taught by renowned professors from prestigious colleges while also being affordable, flexible, and providing a comfortable setting. However, online studies are less communicative as compared to classroom teaching. Whenever a teacher finds that the student is inattentive because of drowsiness, they are able to capture the students’ interest back by asking them questions relevant to the topic.

In online studies, there is no such provision of two-way communication between the instructor and the attendee. In other words, there is no one to keep a track of the students’ behavior during the course. Although the pandemic has ended, virtual learning has proven to be more effective in terms of proactiveness. Students can themselves learn virtually from renowned institutes and track their performance. Using a Drowsiness detection system can help them overcome unconscious sleep that has various causes like weather change, laziness, etc.

1.3 Overview

Drowsiness is a state of sleepiness which might have many reasons such as environmental conditions like rainy seasons, lack of rest, laziness, etc. Online mode has become so trending these days. After a pandemic, online mode has taken a new drift. Students are learning online and working for hours in order to complete a task. Imagine a time to get a habit of sudden unconscious sleep that results in missing out important deadlines or poor performance. To avoid such states, why not a bot is designed which detects the unconscious sleep and gives alerts. This can be useful in improving performance and being aware of unconscious sleep to achieve targets. This system detects tiredness in real time by looking at eye closure and yawning during the monitoring session and gives an alarm intensifying it at different intervals until the user stops it.

Deep learning models such as convolutional neural networks (CNN), recurrent neural networks, etc exist but CNNs are best to find patterns in the images. You Look Only Once (YOLO) models are Convolutional neural network models which are known for their fast and accurate results. In this paper, YOLO models are used to find the patterns in images to detect and classify images as drowsy or awake. In this study YOLOv5 is used which is a faster and more accurate model and various versions of YOLOv5 are trained to get best predictions. This relies on analysis of some consecutive frames and based on the common detections it predicts a state of a person.

1.4 Objectives

The following are the objectives of the project:

- To review the previous studies and efforts put on to build the predictive models.
- To provide the alerts in order to keep students attentive when feeling fatigue or drowsy while studying.
- To produce a system that detects prostration and drowsiness by closing eyes, blinking rate and yawning.

CHAPTER 2

LITERATURE SURVEY

The various work on detecting the drowsiness of the system that helps to tackle real life problems are discussed in this section.

Manikandan v. M, et al. [1] describes how Authors introduced the behavioral-based approach that analyzes the frontal features. This method involves the detection of face, eyes, and head movement. It is implemented using Matlab software. Face Detection is achieved using Viola-Jones Algorithm. The Viola-Jones Algorithm detects frontal faces from the frames that are converted from colored to gray scale images. The captured snapshot is divided into a grid-like structure. To analyze the eye movement a support vector machine (SVM) classifier is loaded. For the head movement detection, Euclidean distance between the frames is calculated continuously and is compared to a threshold value.

V B Navya Kiran, et al. [2] have identified the driver’s drowsy state using PERCLOS in order to compute the Perception of face and face pursuit, Position of eye and eye pursuit, Identification of the state of the eyes, Calculation of percentage of eyelid closure, Identification of the drowsy state. It also used a mean shift method called Cam-Shift algorithm and Viola Jones Algorithm which uses the following techniques in its algorithm HAAR based features, Integral Image Formation, AdaBoost Technology, A cascade of classifiers.

Elena Magán, et al. [3] has proposed two different solutions to approach the fatigue detection problem: The first one is focused on using deep learning to analyze a sequence of images of the driver. The second one uses a combination of AI and Deep Learning techniques to extract the important features from the image and, after that, the obtained data are introduced on a fuzzy inference system that evaluates whether the driver is drowsy or not.

Umang Lahoti, et al. [4] implemented the drowsiness detection system which uses the concept of Ear Aspect Ratio and Mouth Aspect Ratio and it is validated using the ResNet50 pre-trained model. Image-based methods are used to capture the eyes using

Haar based cascade classifier and blink detection is implemented using Histogram of Oriented Gradient (HOG) based features along with SVM classifier. Subsequently, the percentage of eye closure (PERCLOS) is calculated.

Dweepna Garg, et al.[5] focused on improving the accuracy of detecting the face using the model of deep learning. YOLO (You only look once), a popular deep learning library is used to implement the proposed work. The paper compares the accuracy of detecting the face in an efficient manner with respect to the traditional approach.

Belmekki Ghizlene Amira, et al. [6] proposed PerStat method which is used to determine the most frequent state, those states are afterwards combined with an multi-layer perceptron (MLP) that will estimate the state of the driver,it will be fed into an Elman recurrent neural network (ERNN) which will generate the most appropriate feedback agents in ways sequentially relating to these states and their evolution through time.

Feng You, et al.[7] have proposed a driver's face detection architecture based on the improved YOLOv3 and training the network with the open-source data set WIDER FACE. It used the Dlib toolkit to extract facial feature points recognized by improved YOLOv3. Then it created the FFT (Feature Vector Extraction) after analyzing the characteristics of the eye and mouth position.

Sukrit Mehta , et al,[8] have preferred Non-intrusive methods over intrusive methods to prevent the driver from being distracted due to the sensors attached on his body. The proposed approach uses Eye Aspect Ratio and Eye Closure Ratio with adaptive thresholding to detect driver's drowsiness in real-time. The facial landmarks captured by the system are stored and machine learning algorithms have been employed for classification. The system gives a best case accuracy of 84% for random forest classifier.

S. Arefnezhad et al. [9] proposed a non-interfering drowsy detection system based on vehicle steering data using a neuro fuzzy system with support vector machine and particle swarm optimization algorithm. K. Mutya et al. [10] established a system to resolve the problem of drowsiness using the steering wheel algorithm. It is basically based on image-formed or pictorial-based steering movement and the CNN algorithm for proper classification of drowsiness, which can also reduce false drowsy detection rates.

R. Jabbar et al. [11] proposed the Convolutional Neural Network (CNN) technique of the ML algorithm to detect micro sleep and drowsiness. In this paper, detection of driver's facial landmarks can be achieved through a camera that is then passed to this CNN algorithm to properly identify drowsiness.

Jamuna S , et al. [12] proposed a framework which is intended to provide real-time object detection with optimal speed and accuracy to assist the driver. This framework is achieved by implementing the state-of-the-art YOLOv5 algorithm. The whole framework is implemented in the form of three major modules, namely, extraction, detection, and visualization.

Joseph Redmon et al. [13] have proposed You Only Look Once (YOLO)—A one-time convolutional neural network for the prediction of the frame position and classification of multiple candidates is offered by YOLO. End-to-end target detection can be achieved this way. It uses a regression problem to solve object detection. A single end-to-end system completes the process of putting the output obtained from the original image to the category and position. Bounding box prediction and feature extraction of YOLO architecture in our work was inspired by the technique discussed in this paper.

CHAPTER 3

METHODOLOGY

3.1 Dataset

The dataset comprises of more than 3000 subjects is collect which includes Closed Eyes in the wild (CEW) face images, DSM (Driver status monitor) HD videos where frames are generated at random intervals from the videos, human faces including different expressions are taken out from different sources on Google and some videos are recorded by the team members capturing different gestures with different cases like dim light, bright light, different angles, etc. to make the model learn as much gestures and situations as possible. Under drowsy state, yawning and eye closeness is considered [15][16][17]. Figure 3.1 describes the process of custom dataset generation.

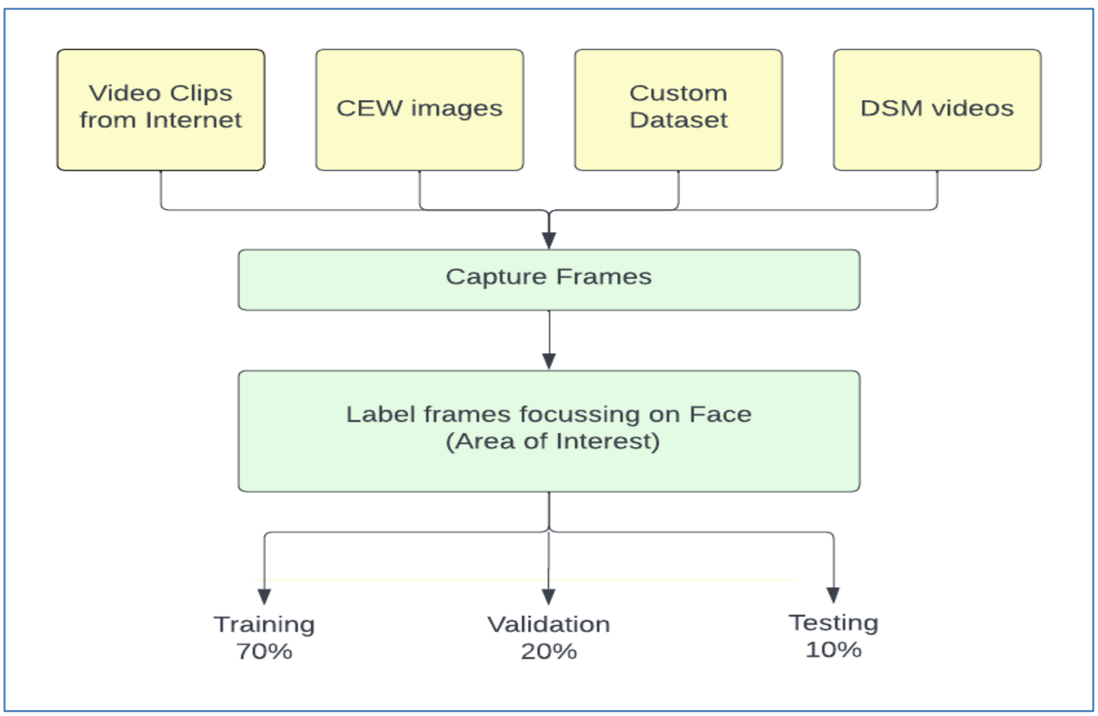


Fig 3.1 Overview of Dataset collection

The states include whether the person is drowsy or awake or no person is detected. So two states are included in the dataset: drowsy and awake. The data collected is then labeled with either drowsy or awakened states. There are a total of 6066 images out of which 2734 are in drowsy state and 3332 are awakened. The data is divided into three parts: training, validation and testing in the ratio of 7:3:1. The training set has 4k images, the validation set has 1.2k images and the testing part includes 606 images.

3.2 Data Labeling

The images are labeled using labeling software to focus on the area of interest (capturing face portion). To train a model on the face region and extract features like closed eyes, normal state, yawning as shown in figure 3.2.

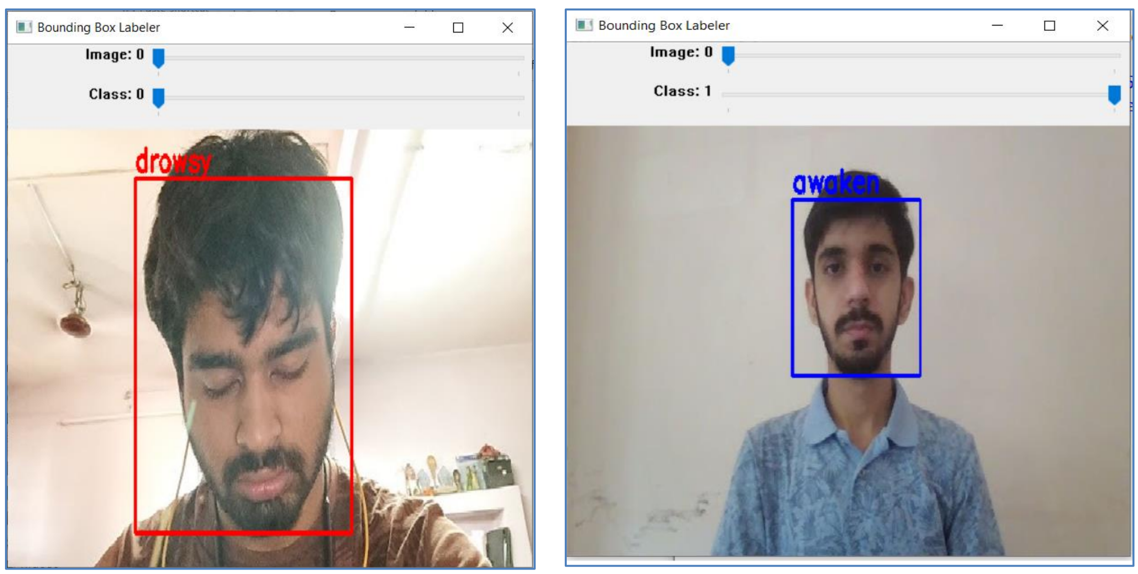


Fig. 3.2 Data labeling cases: (a) Drowsy, (b) Awaken

3.3 Proposed Methodology

The Fig. 3.3 describes an overview of the proposed methodology. The process starts from the data collection of various images showing different expressions mainly at drowsy state and normal state. Model detects the faces and classifies them into two classes: drowsy and awakened. The recognition is based on eye status and yawning state. Check the number of consecutive frames and then based on the states recognized in consecutive frames the alert will be given which will be progressive in nature i.e., increasing alarm intensity if not responded on time.

Fig. 3.3 describes the process, in this the person in front of the webcam will be continuously monitored. Real time frames will be generated using OpenCV. The trained model will identify the state of each frame and recognize the drowsy state. It will check for 20-50 consecutive frames. If all consecutive frames are drowsy, then an alert will be generated. However, the system continues to monitor the person and gives progressive alarm if not responded on time.

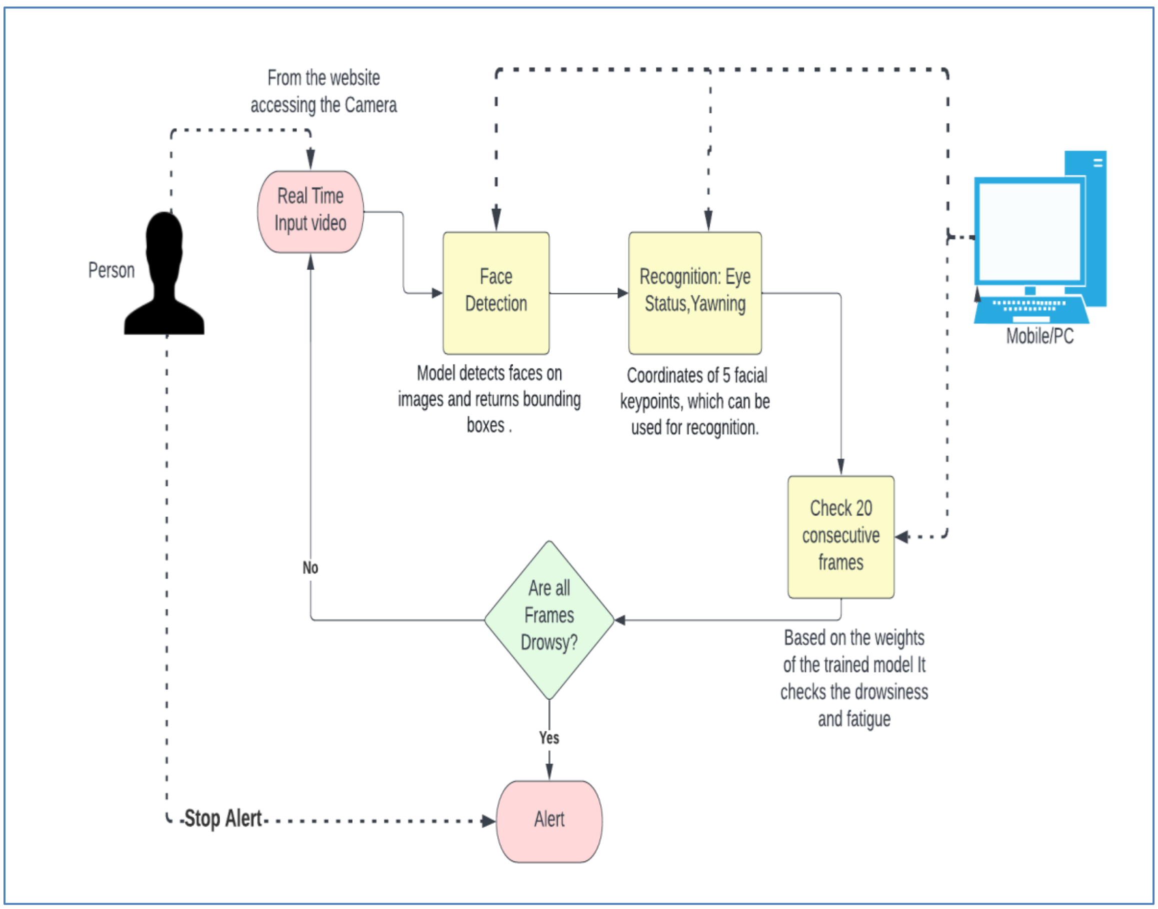


Fig. 3.3 System Overview for Drowsiness Alert

3.4 Modeling Approach

Fig. 3.4.1 shows the use case of the system that depicts steps to detect the state of a person and the scope of the whole system. The process starts with the application activation. Once the webcam is enabled real time image processing is done. Features will be extracted, for drowsiness two states yawn and eyes closure will be monitored. Analyzing these parameters and FPS the alert will be generated

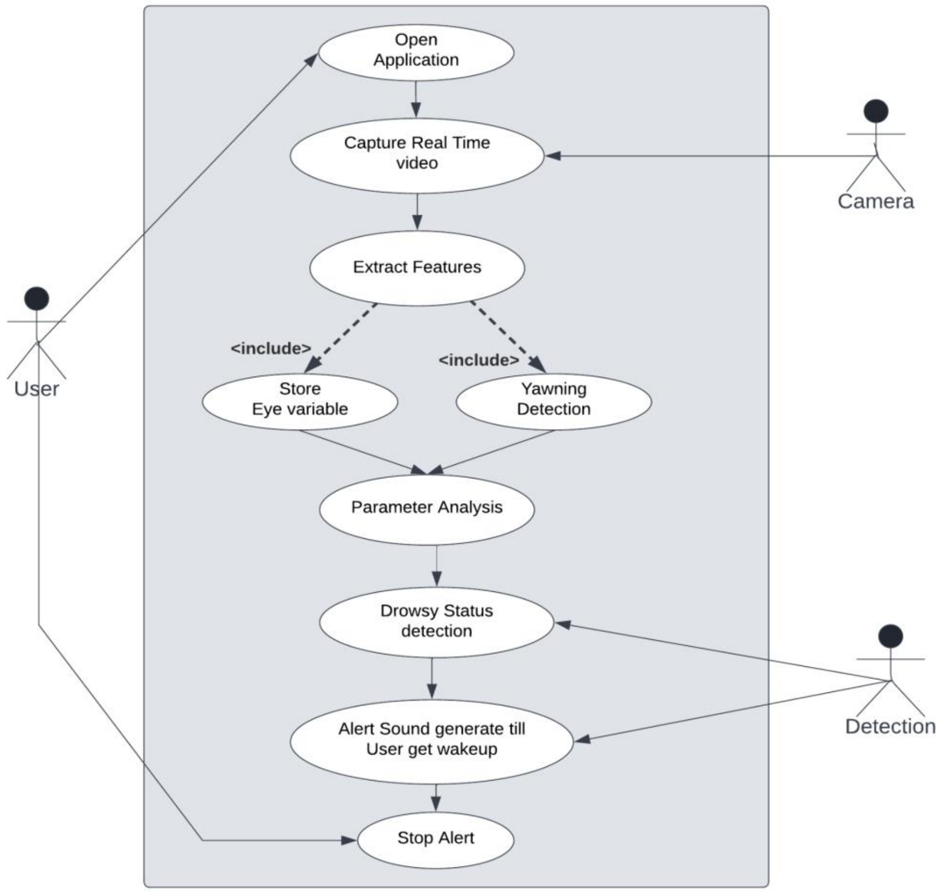


Fig. 3.4.1 Use case diagram for alerting system

3.4.1 YOLO v5 architecture

YOLO v5 is a single-stage object detector; it has three important parts like any other single-stage object detector.

- Model Backbone
- Model Neck
- Model Head

The architecture of YOLO v5 comprises of 3 parts: (1) Backbone: CSPDarknet, (2) Neck: PANet, and (3) Head: Yolo Layer, as shown in Fig. 3.4.2

Model Backbone is mainly used to extract important features from the given input image. In YOLO v5 the CSP—Cross Stage Partial Networks are used as a backbone to extract rich in informative features from an input image.

CSPNet has shown significant improvement in processing time with deeper networks. Model Neck is mainly used to generate feature pyramids. Feature pyramids help models to generalize well on object scaling. It helps to identify the same object with different sizes and scales.

PANet is used as a neck to get feature pyramids. Feature pyramids are very useful and help models to perform well on unseen data. There are other models that use different types of feature pyramid techniques like FPN, BiFPN, PANet, etc.

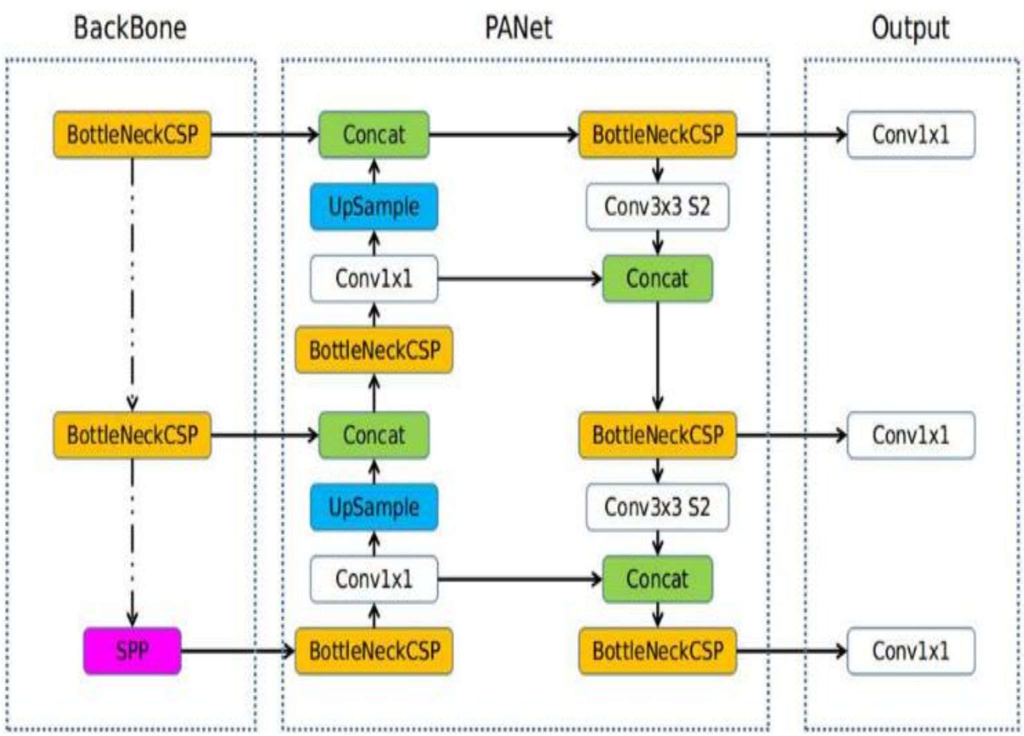


Fig. 3.4.2 YOLO v5 architecture

Activation Function

In YOLO v5 the Leaky ReLU activation function is used in middle/hidden layers and the sigmoid activation function is used in the final detection layer.

Optimization function

For optimization function in YOLO v5, we have two options

- SGD
- Adam

In YOLO v5, the default optimization function for training is SGD.

Cost Function or Loss Function

In the YOLO family, a compound loss is calculated based on object score, class probability score, and bounding box regression score.

Ultralytics have used Binary Cross-Entropy with Logits Loss function from PyTorch for loss calculation of class probability and object score.

3.4.2 System Specification

The requirements to train the model are mentioned in Table 3.4.

Table 3.4 System Specification

Specification	Systems Configuration
Operating system	Windows 10 Pro
CPU	Intel(R) Core(TM) i5-8250U CPU @ 1.60GHz
RAM	14 GB (usable)
GPU	NVIDIA GPU server
Frameworks	PyTorch, OpenCV

3.5 Technology Stack

The project is divided into two parts, one is training a YOLO v5 model to predict words and another one is to develop a deliverable for the predictive system. For the user's sake, a chrome extension is designed that will be enabled first and it starts tracking users through a webcam.

Table 3.5 Detection and Recognition Tools

python™	**Python:** We used python as a programming language. Python is an interpreted high-level general-purpose programming language. Python's design philosophy emphasizes code readability with its notable use of significant indentation.
PyTorch	**PyTorch :** PyTorch is a machine learning framework based on the Torch library, used for applications such as computer vision and natural language processing,
	OpenCV: OpenCV is the huge open-source library for computer vision, machine learning, and image processing and now it plays a major role in real-time operation which is very important in today's systems. OpenCV is used to capture images in real time.
	TensorFlow: TensorFlow is an end-to-end open-source platform for machine learning. It has a comprehensive, flexible ecosystem of tools, libraries and community resources that lets researchers push the state-of-the-art in ML and developers easily build and deploy ML-powered applications.
NVIDIA.	**GPU:** Server used for faster processing.

CHAPTER 4

IMPLEMENTATION

4.1 Training YOLO model on the dataset

After the data is pre-processed, it is feed to the model on the custom dataset shown in Fig. 4.1. The training frames are given to the model and features are recognized and classified based on labels. Different YOLOv5 models are trained on the same dataset and performance of models are then evaluated.

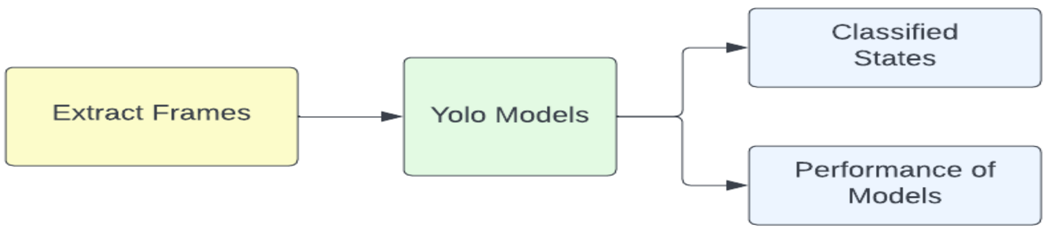


Fig. 4.1 Overview of the Implementation steps

4.2 Face Detection and recognition based in YOLOv5 models

The YOLO v5 architecture, which is a single stage object detector, is used to extract important features from the input images using regions of interest (ROI). It needs labeled images for training on the features. So, the custom dataset that is collected is annotated accordingly, focusing on the face region to achieve the weights. This works on a single face for detecting the states whether drowsy or awakened. Hence frames extracted from videos and given to the framework.

The proposed architecture is a fusion of DarkNet and Cross Stage partial network (CSP). The grids do the feature extraction and target information annotation from the input images. During the face identification the image is divided into N x N grids. If the target image center lies in the grid, that grid is accountable for the identification. The confidence that the grid contains the target is calculated as:

$$\text{Confidence } (C_{ij}) = \text{Probability of object } (P_{ij}) \times \text{ IOU score}$$

C_{ij} denotes the confidence score that the j^{th} bounding box is present in the i^{th} grid. P_{ij} is the probability that there exists a target in the grid, Intersection over union (IOU) between the true and the estimated box. The IOU score must be high for the confidence to be more precise [5]. The illustration of the complete drowsiness system is shown in figure 3.2.

4.3 The Detection and Recognition process

When the detection is enabled, it will ask for webcam permission. Automatically it starts capturing images using OpenCV and a trained model will detect the real time images. Rate of frame creation is 30 (FPS). So, if ceaseless 30-60 frames detected as a drowsy system will invoke the alarm. The alarm would last a couple of seconds and again the detection is monitored. If the drowsy state is continued the alarm intensifies its effect and turns on. The utility for this detection system is described in the following Fig. 4.2.

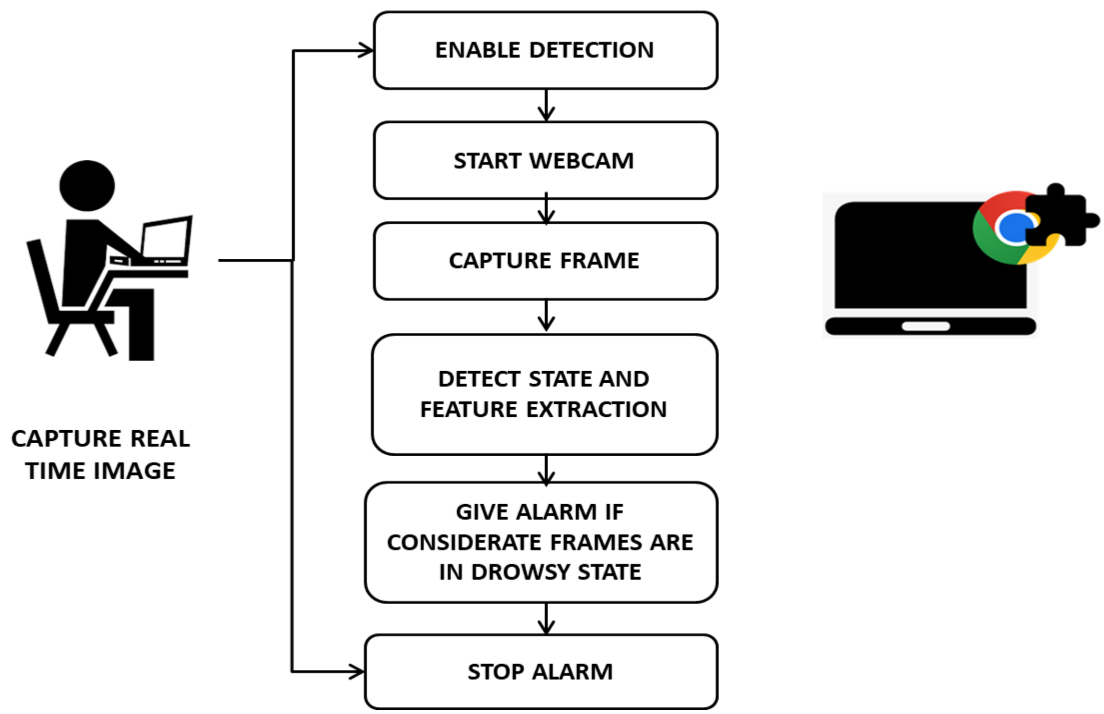


Fig.4.2 Real time Detection System's working

CHAPTER 5

RESULTS AND EVALUATION

The datasets contain images of varied scenarios like talking, screaming, gazing, etc and cases like drowsy, opening and closing of eyes, yawning, nodding are taken into account in order to train the system for all possible instances. The resolution of frames where different for different clips however to keep the size common to capture images during real time, the resolution of frames is kept 640 x 480. The faces in the dataset have significant transformation and perceptions. Hence the dataset is appropriate to the performance in practical scenarios.

Since the utilization of this model will be for real time capturing using webcam, the custom videos were taken using the webcams with the resolution of 1280 x 640. The model is trained on different versions of the YOLO v5 model.

Three model architectures are trained on the dataset.

1. YOLO v5 small model that has the speed of 98ms and is trained on the whole dataset. When the model results are evaluated the model training accuracy was seen to be good but detection was less accurate as compared to other models trained later. Graphs for box_loss, obj_loss, cls_loss, precision, recall, mAP in training and validation are given to interpret results for YOLO v5s model in Fig. 5.1. The precision, recall and mAP are stated in Table 5.1 for comparison with other models.

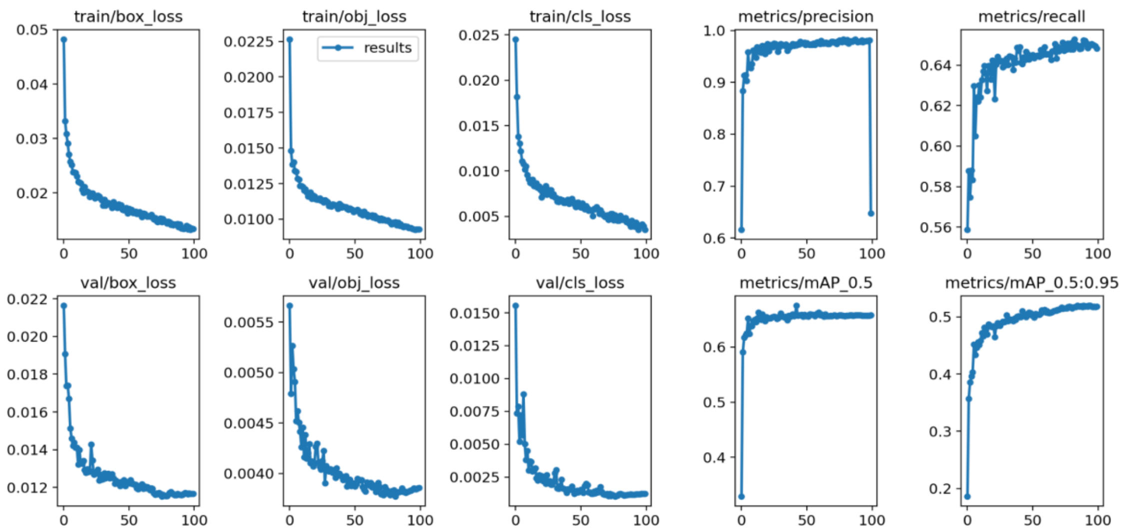


Fig. 5.1 YOLO v5s results

2. YOLO v5 m (medium) model that has the speed of 224ms and is trained on the whole dataset. When the model results are evaluated the model training accuracy was seen to be good but detection was less accurate as compared to other models trained later. Graphs for box_loss, obj_loss, cls_loss, precision, recall, mAP in training and validation are given to interpret results for YOLO v5m model in Fig. 5.2. The precision, recall and mAP are stated in Table 5.1 for comparison with other models.

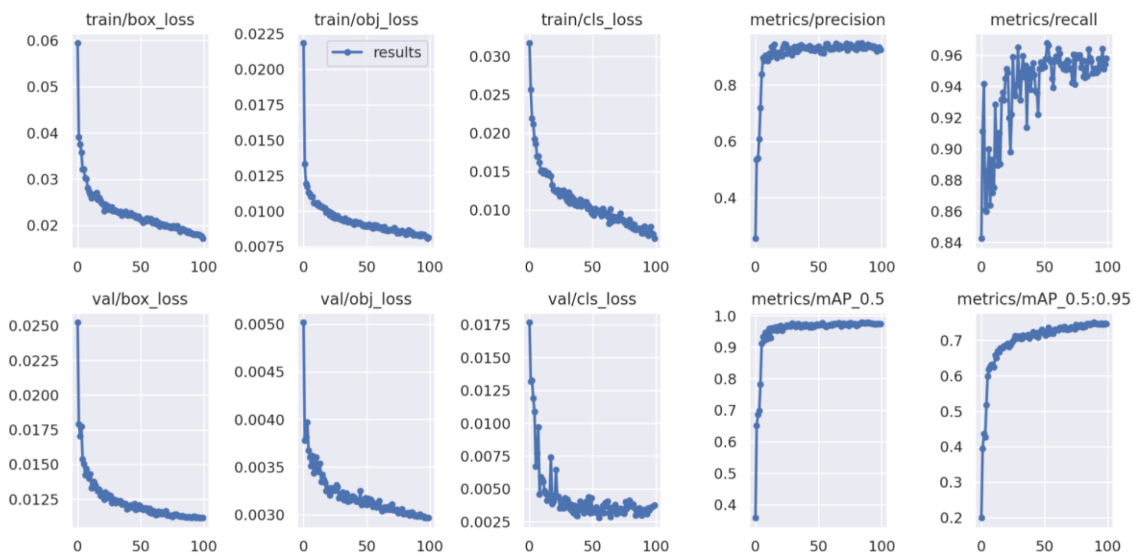


Fig. 5.2 YOLO v5m results

3. YOLO v5 x model that has the speed of 998ms and is trained on the whole dataset. When the model results are evaluated the model training accuracy was seen to be good and it is showing exceptionally better results. Graphs for box_loss, obj_loss, cls_loss, precision, recall, mAP in training and validation are given to interpret results for YOLO v5x model in Fig. 5.3. The precision, recall and mAP are stated in Table 5.1 for comparison with other models.

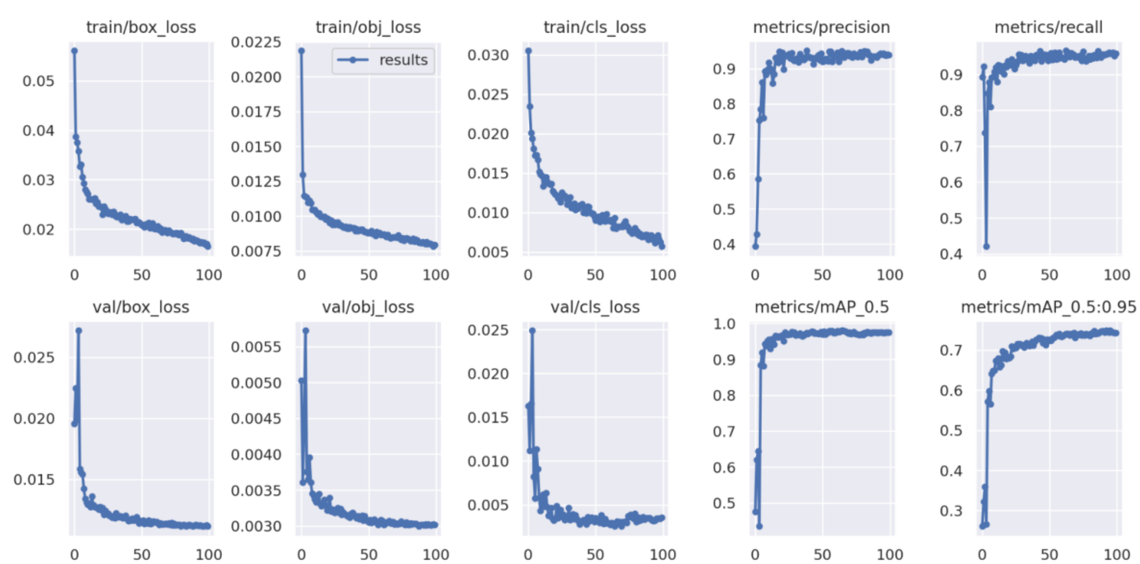


Fig. 5.3 YOLO v5x results

For real time state detection the YOLOv5 x model is chosen. The results of YOLOv5 when real time detection is performed were exceptionally well. The confusion matrix is shown in the following Fig. 5.4.

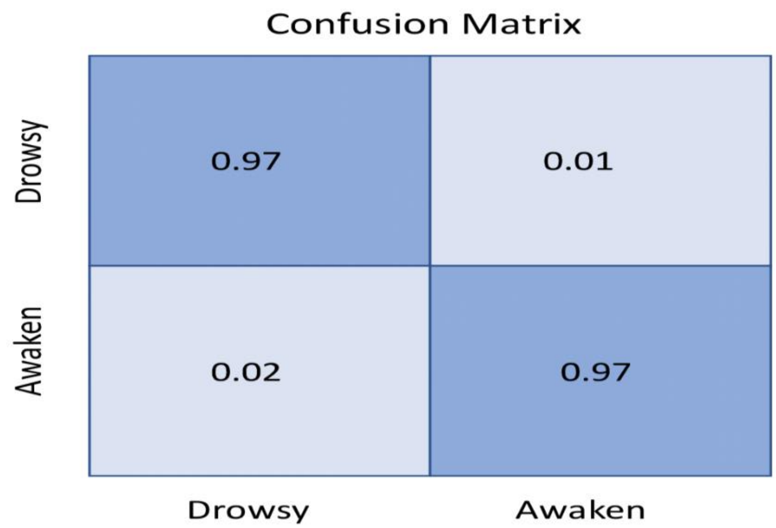


Fig. 5.4 YOLOv5 x confusion matrix

We use the confusion matrix from the Fig. 5.4. to get various measures which are observed for the trained model shown in Table 5.2 and comparison between all models trained is shown in Table 5.1.

Table 5.1 Scores from various YOLO models on custom drowsiness datasets

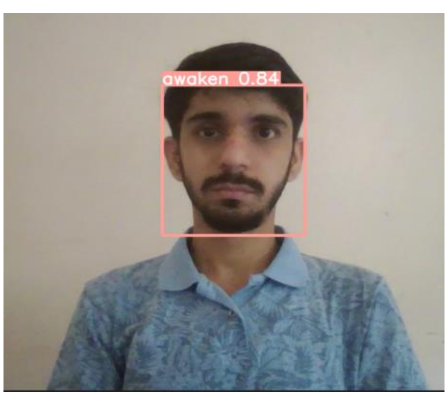
Sr. No.	Models	Precision	Recall	mAP
1	YOLO v5s	0.981	0.984	0.98
2	YOLO v5m	0.941	0.957	0.975
3	YOLO v5x	0.944	0.948	0.979

TABLE 5.2 Scores of YOLO v5x model on trained dataset

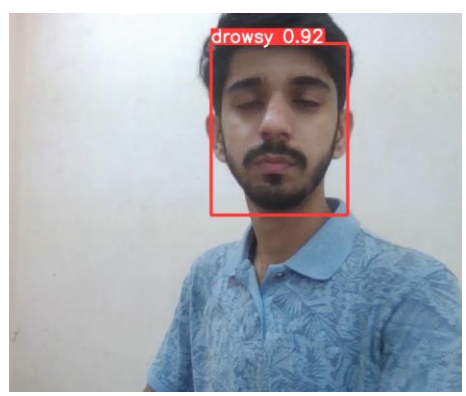
Measure	Value
Sensitivity	0.9898
Specificity	0.9798
Precision	0.9798
Negative Predictive value	0.9898
False positive Rate	0.0202
False Discovery Rate	0.0202
False Negative Rate	0.0102
Accuracy	0.9848
F1 Score	0.9848

State Detection Evaluation

The YOLO v5 model is implemented as core architecture with different configurations to make the model learn desired states. The model is trained for 100 epochs. The precision and mean average precision mAP at 0.5 IOU thresholds is shown in the images. After several tests the speed of the model on the real time images is observed to be 30 FPS. The results are evaluated on videos and images. The confidence score with the predicted class is shown in following results, Fig. 5.5.

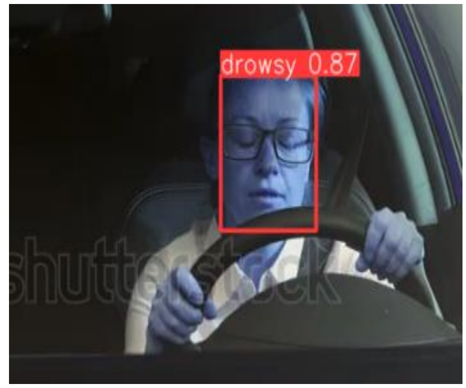


a. (Awaken state with 84% confidence)

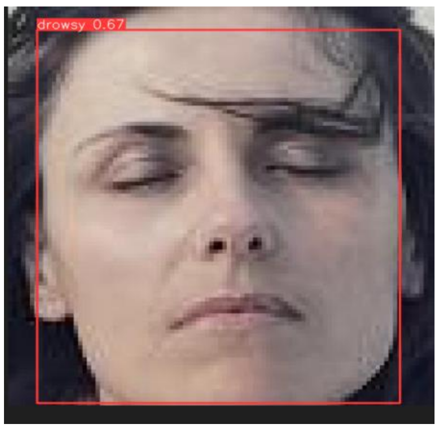


b. (Drowsy state with 92% confidence)

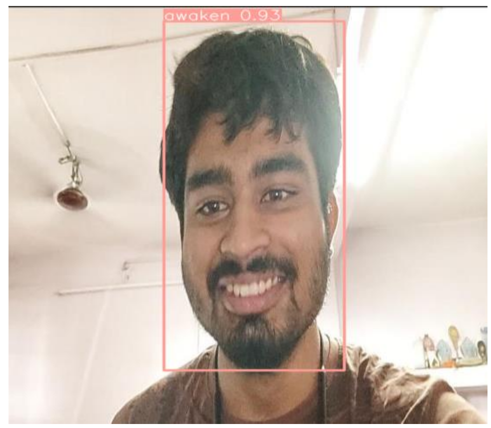
c. (Drowsy state with 66% confidence)



d. (Drowsy state with 87% confidence)



e. (Drowsy state with 67% confidence)



f. (Awaken state with 93% confidence)

Fig. 5.5: Class predictions

The monitoring is continued and even after alarming, the state of the person is predicted continuously. The alarm intensifies its frequency, if not received response on time and it continues until the person is detected as awakened for next consecutive frames.

CHAPTER 6

CONCLUSION AND FUTURE SCOPE

The proposed methodology, drowsiness detection has two components, detects the state whether drowsy or awake and based on the confidence and time alert users. The architecture used for the detection and recognition uses YOLO v5 pre-trained CNN architecture. After several checks and analyses, the model using s, m and x YOLO versions and achieved the mAP score of respective models are 0.98, 0.975, 0.979.

However the accuracy when tested on YOLOv5 x is observed to be more and hence considered for further study. The precision and recall score was observed to be 0.944 and 0.948. The trained model is then fed in the extension to ease the utilization of the system.

Future Scope

This technology is still a hot topic to research and can be used in various fields. Based on the work completed, following modifications can be done:

- The architecture of YOLOv5 can be used as Driver's drowsiness detection. Driver's drowsiness is a sensitive utility. The current work can be extended to extract more features focusing on eyes and lips and recognition algorithms can be applied.
- The dataset can be trained on other object detection models or other YOLO models as to compare results and find optimal model for student tracking.
- The model can be integrated in hardware to use the model in the real world. The future work can include integration of the proposed system with globally used applications like Uber and Ola. The system, if integrated, can reduce the number of casualties and injuries that happen regularly due to these drowsy states of the drivers. This experiment can run as a part of pilot plan i.e. for a few days/months in different regions of the world where such incidents occur regularly. Thus, our proposed approach also gives the same accuracy for the people wearing spectacles. Accuracy of our proposed system improves with the increase in brightness of the surrounding environment. The work can be extended for different types of users such as bike riders or in different domains like railways, airlines etc.

REFERENCES

[1] Manikandan v. M. **An Efficient Drowsiness Detection Scheme using Video Analysis**, International Journal of Computing and Digital System, January 2022, Andhra Pradesh India.

[2] V B Navya Kiran, Raksha R, Anisoor Rahman, Varsha K N, Dr. Nagamani N P, **Driver Drowsiness Detection**, International Journal of Engineering Research & Technology (IJERT), NCAIT - 2020 Conference Proceedings , Bangalore India

[3] Elena Magán , M. Paz Sesmero , Juan Manuel Alonso-Weber, Araceli Sanchis ,” **Driver Drowsiness Detection by Applying Deep Learning Techniques to Sequences of Images** ,” Applied Science 2022.

[4] Umang Lahoti, Rakshita Joshi, Nupur Vyas, Kashmira Deshpande , “**Drowsiness Detection System for Online Courses**” , International Journal of Advanced Trends in Computer Science and Engineering · May 2020.

[5] Dweepna Garg ,Parth Goel ,Sharnil Pandya ,Amit Ganatra ,Ketan Kotecha ,” **A Deep Learning Approach for Face Detection using YOLO,** ” 20XX IEEE.

[6] Belmekki Ghizlene Amira, Mekkakia Maaza Zoulikha, Pomares Hector, “**Driver Drowsiness Detection and Tracking Based on Yolo with Haar Cascades and ERNN**” , International Journal of Safety and Security Engineering, December 2020.

[7] Feng You, Yunbo Gong, Haiqing Tu, Jianzhong Liang, Haiwei Wang ,” **A Fatigue Driving Detection Algorithm Based on Facial Motion Information Entropy,** ” Journal of Advanced Transportation, June 2022.

[8] Sukrit Mehta, Sharad Dadhich, Sahil Gumber, Arpita Jadhav Bhatt, “**Real-Time Driver Drowsiness Detection System Using Eye Aspect Ratio and Eye Closure Ratio**,” International Conference on Sustainable Computing in Science, Technology and Management. , Hosting by Elsevier SSRN , February 26 - 28, 2019.

[9] S.Arefnezhad, S. Samiee, A. Eichberger, and A. Nahvi, "**Driver drowsiness detection based on steering wheel data applying adaptive neuro-fuzzy feature selection**," Sensors, vol. 19, no. 4, p. 943, 2019

[10] K. Mutya, J. Shah, A. D. McDonald, and J. Jefferson, "**What steering pictures are worth? using image-based steering features to detect drowsiness on rural roads**," Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 63, no. 1, pp. 2041–2045, 2019.

[11] R. Jabbar, M. Shinoy, M. Kharbeche, K. Al-Khalifa, M. Krichen, and K. Barkaoui, "**Driver drowsiness detection model using convolutional neural networks techniques for android application**," in 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT), IEEE, 2020.

[12] Jamuna S. Murthy, G. M. Siddesh, Wen-Cheng Lai , B. D. Parameshachari , Sujata N. Patil, K. L. Hemalatha, "**ObjectDetect: A Real-Time Object Detection Framework for Advanced Driver Assistant Systems Using YOLOv5,** " Wireless Communications and Mobile Computing Volume 2022

[13] Redmon J, Divvala S, Girshick R, Farhadi A. "**You only look once: Unified, real-time object detection,** " In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016, pp. 779–788.

[14] F.Song, X.Tan, X.Liu and S.Chen, **Eyes Closeness Detection from Still Images with Multi-scale Histograms of Principal Oriented Gradients,** Pattern Recognition, 2014.

[15] N. Petrellis, P. Christakos, S. Zogas, P. Mousouliotis, G. Keramidas, N. Voros, C. Antonopoulos, "**Challenges Towards Hardware Acceleration of the Deformable Shape Tracking Application**", In the Proceedings of the IEEE VLSI SoC Virtual Conference, 4-8 October 2021, Singapore

[16] N. Petrellis, S. Zogas, P. Christakos, P. Mousouliotis, G. Keramidas, , N. Voros, C. Antonopoulos, "**Software Acceleration of the Deformable Shape Tracking Application**", In Proceedings of the ACM, 2nd Symposium on Pattern

Recognition and Applications as workshop of ESSE 2021, Larissa, Nov 6-8, 2021.

[17] Iason Katsamenis, Eleni Eirini Karolou, Agapi Davradou, Eftychios Protopapadakis, Anastasios Doulamis, Nikolaos Doulamis, Dimitris Kalogeras (2022). TraCon: **A novel dataset for real-time traffic cones detection using deep learning**. 10.48550/arXiv.2205.11830.

[18] https://towardsdatascience.com/yolo-object-detection-with-opencv-and-python-21e50ac599e9

[19] https://docs.ultralytics.com/tutorials/torchscript-onnx-coreml-export/

[20] http://parnec.nuaa.edu.cn/_upload/tpl/02/db/731/template731/pages/xtan/ClosedEyeDatabases.html - Dataset

ANNEXURE – A

NVIDIA DGX DL WORKSTATION

The Centre of Excellence has been helpful for faster processing in the project. RCOEM with association NVIDIA has set-up the 1st Centre of Excellence in Artificial Intelligence and Deep learning (AI & DL) in central India.

Throughout the project, the facility provided by NVIDIA DGX DL Workstation of Centre of Excellence for AI and DL of RCOEM has been useful and the development process was faster and effective. This helped us to train models faster and focus on many other aspects of training a ML/DL model. The system has increased the training where our normal system used to take only few samples from the input samples seeded; the NVIDIA server helps to train on large data samples.