

**“Driver Warning System Using Deep Learning based on Face Features Recognition”**

**Capstone Supervisor  
Dr. Taskeed Jabid**Associate Professor Department of CSE

East West University

**Submitted By**

|  |  |
| --- | --- |
| **Name** | **ID** |
| Rakibul Huda | 2018-2-60-028 |
| Al - Azim Islam Khan Iram | 2018-2-60-007 |
| Md. Asif Shahriar | 2018-2-60-060 |
| Tumpa Karmoker | 2018-2-60-012 |

**“Driver Warning System Using Deep Learning based on different Facial Features Recognition”**

**Abstract:** Many people who drive on a regular basis have fallen asleep at the wheel at some time. This may not always result in danger, but it is a significant issue with major consequences. Drowsy driving is responsible for one out of every four car accidents[1]. Drowsy driving includes not just falling asleep at the wheel, but also being unconscious or not paying attention to the road. Each year, drowsy driving causes about 71,000 injuries, 1,500 fatalities, and $12.5 billion in financial losses[2]. Drivers in Bangladesh are unwilling to obey the regulations while driving, and the majority of truck drivers are drug addicts. As a result, a system that detects drowsiness and provides continual alarms is required. The purpose of this research is to identify drowsiness, unconsciousness, and excessive excitement in drivers based on their facial expression and eye opening and closing intervals. A camera will keep recording by which we will analyze every frame to track if the driver is physically and mentally fit to drive. We have developed a comprehensive system based on deep learning that would utilize the camera to examine the driver's face and eye position in order to identify drowsiness, alert him, and keep him from dozing off. We have included facial expression as a component for detecting drowsiness. The results demonstrate that our driver drowsiness identification system works with high accuracy and that this approach is more accurate than the most recent available algorithms.

**Introduction**

One of the most important body parts is the face which holds a lot of information. Any person's mental state is revealed through their facial expression. In this article, we'll mostly be dealing with facial expressions and eye issues. So, our problem is drowsiness or a sleepy feeling during driving as none wants to sleep during driving. The facial expressions, blinking frequency, and yawning patterns of a fatigued driver differ from those of a typical driver. We call it a problem because it causes a lot of road accidents. This problem is faced mostly by truck drivers at night. They use highways regularly and it is very possible to sleep in the middle of the driving because the roads are straight and noiseless. Let us talk about some of the previous works of this problem. The goal of our capstone project is to develop a system that can ensure safe driving with great accuracy. We will be using an alarm-based system. So that it could make the driver feel when he/she is not sincere about driving. Road accidents are a major problem all over the world. Throughout the world, millions of people are dying for irresponsible driving. According to the Bangladesh Road Safety Foundation's annual report, at least 6,284 people were murdered and 7,468 others were injured in road accidents across the country between January and December 2021, compared to 5,431 people dead and 7,379 others injured in road collisions in 2020(RSF).

Reasons behind the enlarged number of accidents and casualties reckless driving, over speeding, overloading, overtaking, breaking laws, illegal and dangerous competition, long-distance driving without a break, drug and alcohol use, incompetency of the driver, hazardous road, and lack of proper design and construction are all factors. If drowsiness is identified, a warning or alarm signal is given to remind the driver to wake up and get out of the drowsy state. The system first recognizes the face, then the eyes, and finally whether the eyes are open or closed.

This paper mainly focuses on a brief review of researches in the field of Facial features recognition conducted over the past decades. First, conventional Facial features recognition approaches are described along with a summary of the representative categories of Facial features recognition systems and their main algorithms. Deep-learning-based Facial features recognition approaches using deep networks enabling “end-to-end” learning are then presented. To do such task, we employ Deep Convolutional Neural Network which has an architecture that consists of filter layers and a classification layer. We will take the image of eyes which will gives us that the eyes are open or close.

As our main goal is to prevent car accident, so it is mandatory to list the causes of the accident first. Then we are going to filter the list and select the one that we think we could solve with the help of machine learning or deep learning. Finally, we will make a web based or mobile based application for drivers. The application should detect the drowsiness and give immediate warning to the drivers so that they will awake and careful to prevent the car from accident.

**Literature Review**

In article [1], They use deep learning for nonverbal sentiment analysis on facial emotional expressions. They suggested three architectural models that focused on six emotions. The CNN algorithm is employed in this method. An input layer, a convolution layer, a max-pooling layer, a fully connected layer, and an output layer make up the CNN algorithm. For each architecture, they employed a different number of layers. Use ReLU (Rectified Linear Units) and dropout as well. For the third model, the best average accuracy of deep-learning-based FER techniques is 88.9% in this research. The convolutional layer is higher in this layout, while the fully connected layer is lower. In article [2], This article is useful for the emotional expressions of physically handicapped persons and autistic youngsters.

They employed classifiers such as convolutional neural networks (CNN) and long-term memory (LSTM). They first applied Haar-like features to detect faces and eyes, then used a mathematical model approach and the Lucas-Kande optical flow algorithm on their dataset. The dataset utilized in this study has the best accuracy for the CNN model, which is 99.81 percent, and they employed facial landmarks to detect emotions. For the LSTM classifier, they used EEG signals to assess emotion.

In article [3], They employed deep learning and conventional neural networks in this research, and they applied the regularization method known as "dropout" for the fully connected neural network. The average accuracy on CNN is 92.81 percent. Deep Convolutional Neural Networks are used. It is made up of two levels: filter layers and classification layers. They employed six convolutional layers, two subsampling layers, twelve convolutional layers, two subsampling layers, and an output emotion class in their research. However, for facial expression recognition, they developed a convolutional neural network design.

This study [4] presents a fully deep neural network model for facial emotion identification of six different emotions with six convolution layers and two deep residual blocks. After the second and fourth convolution layers, the deep residual blocks are implemented. There are additionally two Fully Connected layers (FC) with ReLU activation functions and a training dropout. This model can handle a variety of image sizes without the need for human intervention. To evaluate the performance, it was tested on two public datasets: JAFFE and CK+. They employed Softmax for classification and DNN for feature extraction, with regularization applied to each weight matrix. For the JAFFE dataset, their proposed model was 95.23 percent accurate, while for the CK+ dataset, it was 93.24 percent accurate. There are additionally two Fully Connected layers (FC) with ReLU activation functions and a training dropout.

This study is represented in [5] as a computer vision-based categorization model. The CK and CK+ datasets from Carnegie Mellon University were used in this study. For classification, they employed SVM, and for feature extraction, they used a Machine Learning technique. For improved speed, two types of images are used: static and real-time. The accuracy of their proposed model was 94.1 percent. Other classification algorithms performed less well.

This research [5] demonstrates a novel FERC technique based on a two-part Convolutional Neural Network. The model was built primarily to identify facial emotion. To overcome the problem of low efficiency, the author used his photographs [6]. It collaborates on datasets from CK+, CMU, and NIST. Because of the many datasets, the suggested model has varying degrees of accuracy. The accuracy obtained from the FERC model is in the range of (78-91%). Other algorithms are used, resulting in a wider variety of accuracies for the comparison problem. CK+, Caltech faces, CMU, and NIST datasets were used.

This article [7] uses a VGGNet to analyze the FER2013 dataset. To establish an optimal model for facial emotion detection, they fine-tuned all hyper parameters, set the starting learning rate to 0.0001, and ran 50 epochs. The accuracy of 73.28% was attained using various optimizers and learning rate schedulers.

This study [8] suggested a system for recognizing face expressions using fused characteristics. For feature extraction, LBP and enhanced ORB descriptors were utilized and face expression identification was done using SVM classification. Their suggested framework outperformed various frequently used approaches based on the JAFFE, CK+ and MMI dataset and 87.6% accuracy were achieved.

Authors worked on Facial Expression Recognition Using Local Direction-Based Robust Features and Deep Belief Networks in their paper [13] (DBN). They looked at two different sorts of images: RGB and Depth. They have extracted features using diverse approaches such as LDPP, PCA, and GDA and used two separate models, HMM (Hidden Markov Model) and DBN for classification. They discovered 92.50 percent accuracy in RGB based photos.

The authors of study [14] identified facial expressions and created their own CNN model from scratch. They've employed two Conv2D in their own model to deal with the incoming photos. They utilized the K-means method to recognize faces, and then used their trained model to test the results. For vast amounts of data to be trained, good quality raw data and a high-performance GPU are required.

The authors of paper [15] suggest a Neural Network that speeds up both the training and recognition processes. They utilized the Gabor filter to extract features and CNN to classify them. The approach they presented was to take photos as input, resize them, then apply Gabor filter-1, Gabor filter-2, and CNN training. They were able to obtain 91 percent accuracy in a standard CNN model after 30 epochs, and they were able to reach 94 percent accuracy by using two gabor filters and CNN.

Face recognition in a single shot has recently received a lot of attention. [16] proposes a method for anticipating facial expressions in group pictures. This study makes use of two different datasets: FER-2013 and custom datasets. Group emotion identification is conducted on both static photographs in databases and dynamic images captured in real-time using a camera. The original RGB image is converted to a grayscale image. The OpenCV Haar filter is then applied to both static and dynamic images to detect faces. Real-time photos are captured using a webcam. After the faces in a particular image have been located, the image can be cropped and facial features can be detected. The face landmarks that have been collected are then utilized.

In their paper [17], They compared five different ways for distinguishing four basic emotions (happy, sorrow, rage, and fear) from static facial pictures in real-time. They looked at three deep-learning algorithms based on convolutional neural networks and two standard approaches for categorization. In real-time, eight individuals (3 males and 5 females) were tested. The participants have all written consented to participate in the study. During daylight hours, the experiments were conducted in a room with increased brightness. Preliminary findings reveal that fine-tuned AlexNet CNN and Affdex CNN approaches outperform SVM and MLP approaches in terms of generalization power and performance in real-time applications. Commercial Affdex CNN is, on average, more accurate.

In [18], Tensor Flow was used to implement basic emotion recognition. To categorize photographs into 1,000 different groups, the TensorFlow model was taught. The ML.NET model uses a component of the Tensor Flow model into its workflow to train the image separation model. They divided the data into two groups: train and test. The train data was then used to create the model, and the test data was used to evaluate the performance. The log loss was 0.33 on average.

According to [19], If humanoid robots can understand human emotion via face emotion recognition, they can better adjust the conversation ahead of time. They employed CNN to process picture datasets and extract features in the form of edges, which could then be used to perform classification tasks instead of depending on human-predicted features. The real-time robust emotion classifier achieved a lower computational cost and a nearly accurate output. To improve accuracy, variable filter sizes, network convolution layers, and late-in-the-network downsampling of the picture size are utilized, allowing the network to extract the best characteristics from deep layers. The network's resilience is increased by using many datasets since it experiences all possible changes, such as non-face images, illumination, age, and occlusion.

In [51], this paper presents a research on the identification of dryness of a person while driving using deep learning. CNNs and LSTMs are used for detecting drowsiness. The face detection is done by Viola Jones face detector and eyes are detected by eye detection module. The accuracy they got is about 88.5%.

In [52], this paper focuses on the detection of micro sleep and drowsiness by using CNN.Drowsiness detection system can be easily developed in embedded device or mobile phones.CNN is used for the detection.The whole process is the application used to capture the image data and then sends data for processing.This android mobile application is able to take pictures of the drivers for detecting micro sleeps.

In [53], this paper represents drowsiness detection technology using deep neural networks. Multilayer Perceptron Classifier has been used to detect the facial landmarks of the drivers. An android application has been developed for this system. This application got around 81% accuracy.

In [54], this shows the approach of driver drowsiness detection using deep CNN model. The proposed research work achieved result by using four CNN models. It has gained around 85% accuracy.

In [55], this paper shows the detection of real-time drive-drowsiness systems using face-tracking algorithms, CNN and DriCare.Owing to the shortcomings of the previous algorithm, the paper was introduced a new face-tracking algorithm to improve the tracking accuracy. The research paper has achieved accuracy 92%.

**Problem Analysis**

Bangladesh, with a population of 162.1 million people and a land area of 1, 48,610 square kilometers, is one of the world's most heavily populated countries. Almost a quarter of the population lives in cities, and the population in urban areas has really been between 7 to 8 percent during the last decade [9]. Road accidents are a typical occurrence in a country like Bangladesh with such a huge population, and the cause is generally due to the vehicle driver's lack of attention [10]. Sometimes, the person who drives the vehicle gets sleepy and his eyes get drowsy. This drowsiness can lead to a serious problem like accidents. In this study, we are going to propose a method that ensures safe driving considering the driver's drowsiness problem.

Sleeping less can make a drowsy driver far more hazardous on the road than a speeding driver. So, our main focus will be the detection of the level of drowsiness with the help of some machine learning algorithms. Firstly, we are going to collect some images of people whose eyes in those images are open eyes and closed eyes. Then we will extract only the eye part from the images and fit them in various models and find the accuracy. Again from the face dataset, Imagine if the eyes look open then the output will be level 1, if the eyes are completely closed then level 2. For level 1, there will be a safety alarm so that the driver will wake up and prevent the vehicle from accident.

Additionally, we will also recognize some of the facial emotions for example: angry, yawn, happy, neutral etc to detect the mental condition of the driver. Too much joy can raise the excitement, on the other hand too much anger can manipulate the brain. There will be warning level alarms for them. There will be one danger alarm for a yawn face also. So, we have analyzed the problem initially like that but it can be changed or optimized according to the implementation.

**Materials and Method**

We needed some technical specifications to implement our concept. Hardware and software specs are the two types of specifications accessible. We needed to fulfill the following hardware requirements in order to implement our code: Depending on the Collaboratory assignment, the Graphic Processor Unit (GPU) was utilized. In all, 12.7GB of RAM was accessible in all cases. We worked on our local system, which was running Windows 10 X64 Bit and had an Intel Core I5 8500 CPU with 16 GB of 2666MHz RAM, when Colaboratory displayed inadequate memory for processing. For data retrieval and storage, we additionally connected a 1TB hard drive (HDD) and Dropbox. We had to fulfill the following software requirements in order to develop our application. The most crucial tool for us to continue forward with our implementation was Google Colaboratory.This project was built with Python 3.7. In TensorFlow 2.3.0, the Keras library was utilized to generate all of the models. We utilized Google Colaboratory for the most part. On occasion, we also used Anaconda, JupyterNotebook, and PyCharm. We took data and information from Kaggle and saved it to Google Drive and Dropbox.

Data has been taken from FER-2013 for faces and Drowsiness-dataset for eyes. The FER-2013 had 6 types of data and we reduced it to 4 types. The classes of data were not also accurate so we needed to rearrange them for better accuracy. In terms of research ethics we wanted to make this model work for people from any ethnicity. The Viola-Jones algorithm has been used to detect the face and eyes from the window and later the pre-trained Xception model has been use to predict the face emotion and drowsiness. Later the whole project has been implemented using Django framework. Below we have shown how the alarm will work in different face and eye position.

|  |
| --- |
| **For Faces** |
| |  |  |  | | --- | --- | --- | | **Face Emotion** | **Sound Frequency** | **Sound Time** | | Neutral | No Sound | - | | Happy | 300 Hz | For 0.3 second | | Anger | 700 Hz | For 0.5 seconds | | Yawn | 2000 Hz | For 1 seconds | |
| **For Eyes** |
| |  |  |  | | --- | --- | --- | | **Eye Position** | **Sound Frequency** | **Sound Time** | | Open Eye | No Sound | - | | Close Eye | 3000 Hz | Until Eyes open | |

**Fig**: How alarm will work

The Viola-Jones algorithm keep detecting the face and eye to import them into the model. Later the model gives the prediction differently for eyes and for faces. If the driver closes his eyes for more than 2 seconds than an alarm will make sound of certain frequency which will be able to make him alert. Likewise there will be different kind of sounds of different frequency.

Environmental Impact Assessment (EIA) [22] is a method of assessing the possible environmental consequences of a proposed project or development, taking into account interconnected socioeconomic, cultural, and human-health consequences, both positive and negative. Environmental and social impact assessment is a multidisciplinary technique that combines a cost-benefit analysis of a project with a consideration of the project's environmental consequences.

The capstone project we are working on focuses on safe driving. Our focus is to guarantee the safety of passengers and drivers by anticipating drowsiness and various emotions that drivers may experience. Many modern technologies, such as data processing, machine learning, and deep learning, would be used to make these predictions. For ensuring safe driving we will detect the face of the driver. If they feel neutral, drowsy, extremely happy, angry, fear the system will alert them because if the driver is not neutral the passengers might be in danger. When the system detects the face of the driver without neutral it will alert them by giving a loud alarm. So, the impact created by this capstone project on society and the environment can be a matter of discussion. To do so we are going to talk about a modern framework called ESIA. According to the most frequently accepted definition, an Environmental and Social Impact Assessment (ESIA) is a complete document examining a Project's prospective environmental and social risks and repercussions. In the case of our capstone project, we have to discuss two impacts. One is Environmental impact which is also called EIA and another one is Social impact which is also known as SIA.

The major goals of an Environmental Impact Assessment ("EIA") are to inform decision-makers about the Scheme's environmental consequences on people and the environment, as well as to minimize the project's negative effects within engineering and other limits. The social impact assessment (SIA) method is used to evaluate the social consequences of infrastructure projects and other development measures [23].

The capstone project we are doing on safe driving does impact the environment but it does not do any harm or imbalance to nature. In our project, we will design a complete system that will utilize artificial intelligence to identify the driver's drowsiness while driving and reduce the hazards to the passengers. For this reason, there was no use of raw material or a new type of hardware that would impact our environment negatively. As there is no use of raw material there will be no sign of soil damage, water damage, or air pollution. It keeps the environment as it is and helps people in need of that project's purpose. Our capstone project can save people's lives. Because every life is valuable, our system will be built to reduce the risk of each passenger traveling. As a result, we will only consider the welfare of the people in our society, and we will ensure that our project will not play any dishonest or immoral role in society.

After evaluating everything, we can confidently state that our capstone project has no negative environmental or social consequences and that it has a positive beneficial influence on our environment and society.

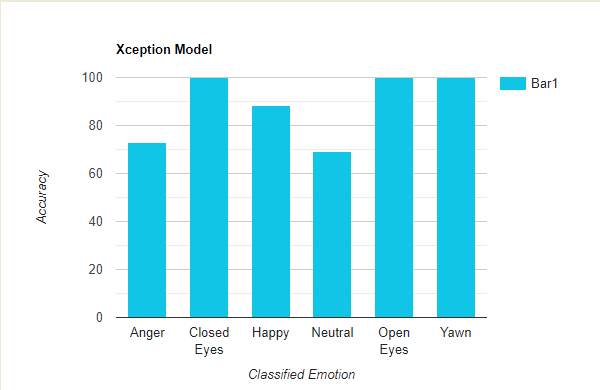
**Results and Discussion**

First we collected the data from two different datasets. One is FER-2013 and another one is drowsiness dataset for open eyes and closed eyes. Images of eyes need to be converted to grayscale image as they were RGB image. Only static images in the dataset are used to train all of our models. The images of our dataset contains only the face which are enough cropped then they had been turned into grayscale image because it is well-known for extracting descriptors rather than working directly with RGB images, and grayscale significantly simplifies the process and minimizes CPU needs. Like the face images, eyes are also in grayscale image form. As we wanted to warn the driver for any kind of emotional imbalance we had to combine same sort of data like combining angry, disgusted or disturbed faces for anger. The surprised face has been used as yawning face. Feature extraction and classification are included in emotion identification models. After data collection and we need to process them accordingly like resizing, rescaling, zooming etc. The size of the input is significant because the larger the input, the more factors the network must deal with. More parameters may cause several issues because they demand more computer power and data to train on. Overfitting can happen if there are a lot of parameters but not enough data, which is a concerned issue with CNNs. Our target input size was 224\*224, rescaled as 1/255 and zoom range was 0.2. After data collection and pre-processing, we had run 9 different models so far which are Inception V3, Vgg16, Vgg19, CNN, Mobile Net V2, Densenet, Efficient B7, Alex Net, Xception. Below we attached our all model accuracy and our chosen model’s accuracy of each classification.

|  |  |
| --- | --- |
| **Model Name** | **Accuracy** |
| Xception | 96.47 % |
| Inception V3 | 94.02 % |
| Mobile Net V2 | 87.07 % |
| VGG19 | 84.62 % |
| CNN | 80.77 % |
| Dense Net | 79.06 % |
| VGG16 | 78.85 % |
| Efficient B7 | 72.22 % |
| Alex Net | 45.51 % |

**Fig**: Model Accuracy

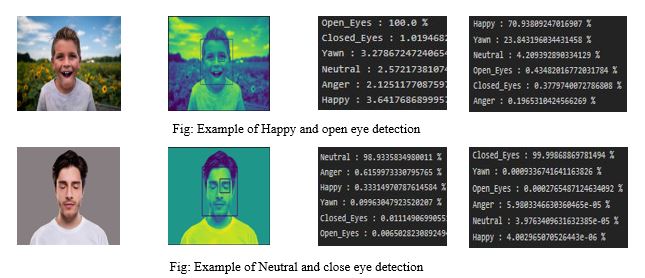
In every model we had gone through 20 epochs and 6 for Dense as we have 6 category which are close eye, open eye, happy, angry, neutral faces and yawning face. The best three models for us are Xception, Inception V3 and CNN respectively. Our second best model Inception V3 has also good accuracy but we had to avoid that because of accuracy per classification.



**Fig**: Accuracy for each classification

Like for happy and anger the accuracy was not good enough which is the main reason for choosing Xception over Inception V3. The training accuracy is also greater than the validation accuracy which is practically good. We also could manage to make validation loss as low as possible. The difference between validation loss and training loss is not very much which is also practically good.

When we were ready with our data and model with proper accuracy, we went for stream data and more specifically uncropped face images. Standard Viola-Jones algorithm is used to detect the face from the stream data and also extracts the eyes from the detected face. This face identification and eye extraction approach is quick and accurate, and it is regarded as the gold standard in the image processing world. Viola-Jones algorithm is based on haar-cascade and it utilizes grayscale image to achieve real time performances. This algorithm was giving almost 99% accuracy from static images. Below examples are attached to show how perfectly the detection method was working and their predicted result was also perfect.



The issue arises when the images were coming from stream data and we have to predict the category on the go. First issue was detecting the face constantly and the extraction of the eyes. The prediction was also working almost as real life usable. The face and eye detection was on point but it was lagging in terms of prediction. Below we attached some screenshots of real time frames.

We are still working on these issues and hopefully we can be able to approach with a good solution next time. The fundamental idea is to give highest security in highway driving.



**Fig**: Detection and Prediction from Webcam

It will keep tracking the face of the driver and detecting his emotion let’s say at least 5 times in a second. If we find that his eyes are closed for last 2 seconds or he is feeling angry, anxious or sleepy for last 5 seconds, we will notify to stop or take some rest. As the highway driving has many cases of accident due to sleepy driving, this system will be able to reduce this kind of accident.

|  |  |  |
| --- | --- | --- |
| **Approach** | **Accuracy** | **Reference** |
| Xception | 96.47 % |  |
| MC-KCF | 93.6 % | [19] |
| Logistic Regression using HRV | 90 % | [25] |
| CNN and LSTM | 88.5 % | [15] |
| CNN | 88 % | [24] |
| Ensemble (Alexnet, VGG-FaceNet, FlowImageNet, ResNet) | 85 % | [18] |
| D2CNN-FLD CNN | 83.33 % | [16] |
| MLP | 81 % | [17] |

**Fig**: Comparison of different approaches

There are several works have been done with drowsiness detection but we could not find any work where face emotion has been added. We tried to keep tracking the emotions of the driver. According to our findings the MC-KCF was the best accuracy holder. Below we attached a table containing the accuracy of other approaches.

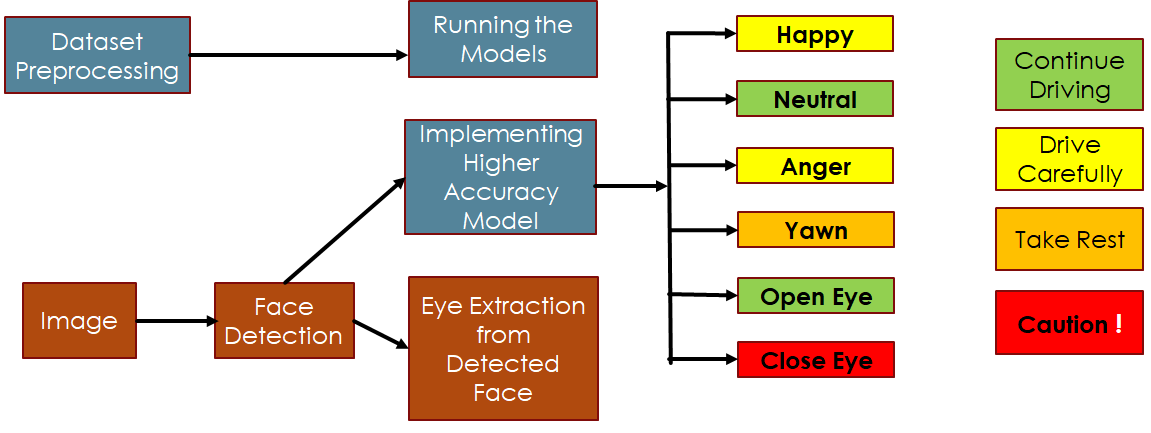


Fig: Workflow diagram

**Conclusion**

This study suggests an improved pretrained model based approach for detecting driver’s inattention. In most of the drowsiness detection system their main focus is to detect yawning or close eyes but we had been added face emotion as extra safety feature which will keep track if the driver is mentally stable to drive. The main objective of this extra feature is to detect driver’s mental mood so that the system can alert him if he gets angry or too excited for speeding. The above model alerts instable drivers with a sound when it detects yawning face, close eyes and inappropriate mentality from the recording using deep learning model (Xception) with an accuracy of 96.47%. It can farther be improved by using night vision technology so that when the driver need to drive at night the system can still function properly. To enable cutting-edge driving assistance technologies, this may simply be implemented into dashboards in the upcoming generation of automobiles.

**References:**

[1] N. Meeki, A. Amine, M. A. Boudia, and N. Meeki, “Deep Learning for Non Verbal Sentiment Analysis : Facial Emotional Expressions,” *GeCoDe Lab. Dep. Comput. Sci. Tahar Moulay Univ. Saida.*, vol. 3014, pp. 1–11, 2020.

[2] A. Hassouneh, A. M. Mutawa, and M. Murugappan, “Development of a Real-Time Emotion Recognition System Using Facial Expressions and EEG based on machine learning and deep neural network methods,” *Informatics Med. Unlocked*, vol. 20, p. 100372, 2020, doi: 10.1016/j.imu.2020.100372.

[3] D. Y. Liliana, “Emotion recognition from facial expression using deep convolutional neural network,” *J. Phys. Conf. Ser.*, vol. 1193, no. 1, 2019, doi: 10.1088/1742-6596/1193/1/012004.

[4] D. K. Jain, P. Shamsolmoali, and P. Sehdev, “Extended deep neural network for facial emotion recognition,” *Pattern Recognit. Lett.*, vol. 120, pp. 69–74, 2019, doi: 10.1016/j.patrec.2019.01.008.

[5] S. Gupta, “Facial emotion recognition in real-time and static images,” *Proc. 2nd Int. Conf. Inven. Syst. Control. ICISC 2018*, no. Icisc, pp. 553–560, 2018, doi: 10.1109/ICISC.2018.8398861.

[6] N. Mehendale, “Facial emotion recognition using convolutional neural networks (FERC),” *SN Appl. Sci.*, vol. 2, no. 3, pp. 1–8, 2020, doi: 10.1007/s42452-020-2234-1.

[7] Y. Khaireddin and Z. Chen, “Facial Emotion Recognition: State of the Art Performance on FER2013,” 2021, [Online]. Available: http://arxiv.org/abs/2105.03588.

[8] B. Niu, Z. Gao, and B. Guo, “Facial Expression Recognition with LBP and ORB Features,” *Comput. Intell. Neurosci.*, vol. 2021, 2021, doi: 10.1155/2021/8828245.

[9] R. Jadhav, J. Bhuke, N. Patil, and A. Navi Mumbai ---------------------------------------------------------------------, “Facial Emotion Detection using Convolutional Neural Network,” *Int. Res. J. Eng. Technol.*, vol. 11, no. 8, p. 1077, 2008, [Online]. Available: www.irjet.net.

[10] M. M. Taghi Zadeh, M. Imani, and B. Majidi, “Fast Facial emotion recognition Using Convolutional Neural Networks and Gabor Filters,” *2019 IEEE 5th Conf. Knowl. Based Eng. Innov. KBEI 2019*, pp. 577–581, 2019, doi: 10.1109/KBEI.2019.8734943.

[11] Dshahid380, “Convolutional Neural Network. Learn Convolutional Neural Network,” no. Icoei, pp. 1–68, 2019, [Online]. Available: https://towardsdatascience.com/covolutional-neural-network-cb0883dd6529.

[12] A. Kartali, M. Roglic, M. Barjaktarovic, M. Duric-Jovicic, and M. M. Jankovic, “Real-time Algorithms for Facial Emotion Recognition: A Comparison of Different Approaches,” *2018 14th Symp. Neural Networks Appl. NEUREL 2018*, pp. 1–4, 2018, doi: 10.1109/NEUREL.2018.8587011.

[13] G. K. Saini, H. Chouhan, S. Kori, A. Gupta, and ..., “Recognition of human sentiment from image using machine learning,” *Ann. …*, vol. 25, no. 5, pp. 1802–1808, 2021, [Online]. Available: https://www.annalsofrscb.ro/index.php/journal/article/view/4703%0Ahttps://www.annalsofrscb.ro/index.php/journal/article/download/4703/3767.

[14] S. Jaiswal and G. C. Nandi, “Robust real-time emotion detection system using CNN architecture,” *Neural Comput. Appl.*, vol. 32, no. 15, pp. 11253–11262, 2020, doi: 10.1007/s00521-019-04564-4.

[15] H. B. Vanjani and U. Varyani, “Identify dozyness of person using deep learning,” *Int. J. Appl. Eng. Res.*, vol. 14, no. 4, pp. 845–848, 2019.

[16] 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,” *2020 IEEE Int. Conf. Informatics, IoT, Enabling Technol. ICIoT 2020*, pp. 237–242, 2020, doi: 10.1109/ICIoT48696.2020.9089484.

[17] R. Jabbar, K. Al-Khalifa, M. Kharbeche, W. Alhajyaseen, M. Jafari, and S. Jiang, “Real-time Driver Drowsiness Detection for Android Application Using Deep Neural Networks Techniques,” *Procedia Comput. Sci.*, vol. 130, pp. 400–407, 2018, doi: 10.1016/j.procs.2018.04.060.

[18] M. Dua, Shakshi, R. Singla, S. Raj, and A. Jangra, “Deep CNN models-based ensemble approach to driver drowsiness detection,” *Neural Comput. Appl.*, vol. 33, no. 8, pp. 3155–3168, 2021, doi: 10.1007/s00521-020-05209-7.

[19] W. Deng and R. Wu, “Real-Time Driver-Drowsiness Detection System Using Facial Features,” *IEEE Access*, vol. 7, pp. 118727–118738, 2019, doi: 10.1109/ACCESS.2019.2936663.

[20] SidhyarthaMukherjee3 DeepakSharma, Uday Shankar Singh2, “A study on road traffic accidents in Anand-Gujarat,” *Heal. Line*, vol. 2, no. 2, pp. 2–5, 2011, doi: 10.13140/RG.2.2.18110.08007.

[21] K. M. Hasib, M. I. H. Showrov, and A. Das, “Accidental Prone Area Detection in Bangladesh using Machine Learning Model,” *2020 3rd Int. Conf. Comput. Informatics Eng. IC2IE 2020*, pp. 58–62, 2020, doi: 10.1109/IC2IE50715.2020.9274581.

[22] P. Martinez, “Environmental and Social Impact Assessment ( ESIA ) Key elements of an ESIA and an ESIA report,” no. July, pp. 1–7, 2016.

[23] R. F. M. Ameen, M. Mourshed, and H. Li, “A critical review of environmental assessment tools for sustainable urban design,” *Environ. Impact Assess. Rev.*, vol. 55, pp. 110–125, 2015, doi: 10.1016/j.eiar.2015.07.006.

[24] S. Jadhav, R. Ghangale, N. Mundhe, and P. S. Bhosale, “Drowsy Driver Detection using Representation Learning and Face Detection Technique,” *Imp. J. Interdiscip. Res.*, vol. 2, no. 11, pp. 2012–2015, 2016.

[25] M. Babaeian, N. Bhardwaj, B. Esquivel, and M. Mozumdar, “Real time driver drowsiness detection using a logistic-regression-based machine learning algorithm,” *2016 IEEE Green Energy Syst. Conf. IGSEC 2016*, 2016, doi: 10.1109/IGESC.2016.7790075.

[26] Y. Albadawi, M. Takruri, and M. Awad, “A Review of Recent Developments in Driver Drowsiness Detection Systems,” *Sensors*, vol. 22, no. 5, pp. 1–41, 2022, doi: 10.3390/s22052069.

Top of Form

**Appendix:**

|  |  |  |
| --- | --- | --- |
| Outcome Description (CO) | Knowledge Profile (KP) | Engineering Problem (EP) |
| Analyzed various aspects of the objectives for designing a solution for the capstone project. (CO3) | **K1**: We tried to study about road accidents and what causes the most destructive and underrated accidents and how to solve them.  **K2:** We mostly did the numerical analysis to analyze our best models and their accuracy of different classifications.  **K3**: The standard Viola-Jones and Deep Learning model have been used to structure our framework.  **K4**: The experience of machine learning and image processing has been enough to build a framework which is built by both of them. | **EP1**: To solve the problem, the knowledge we had, is applying machine learning and the basic idea of image processing to feed it any ML model.  **EP2**: The conflicting issues were among the models as they have different accuracy for different classification. The detecting methods also had many issues as we had to work in real time.  **EP3**: There has not been any work done on facial features to predict drowsiness or other emotional problems during driving.  **EP6**: Police, cargo drivers, bus drivers can be our stakeholders to acknowledge many unknown factors to solve our conflicting factors.  **EP7**: The interdependence issues are like if someone stays neutral and sleepy at the same time, it gets cumbersome to predict if it is dangerous or safe to drive though the machine will predict it as dangerous. |
| Designed and developed solutions for the capstone project that meet public health and safety, cultural, social, and environmental considerations. (CO4) | **K5**: Before putting our detecting drowsiness idea into action, we learned what we needed to know to make it work. We looked at many sources and our academic courses to gather technical knowledge, design methods, and applicable tools and resources for developing components, systems, and processes that met particular requirements. We looked at many methodologies such as CNN and machine learning models, and after exhausting all possibilities, we designed our product in a way that will be extremely beneficial for society. Importing various types of tensorflow packages, data pre-processing, data augmentation, statistical analysis of datasets, CNN model deployment, machine learning model deployment, and finally detecting drowsiness in any photos are all part of our detecting drowsiness using face feature project. | **EP1**: The required indepth ideas are training the dataset, building models, face and eye detection and predicting them accordingly.  **EP2**: The conflicting areas were detecting the eye from the extracted face. putting the right resize factor for the input image.  **EP4**: The infrequent issues are detecting the mouth as eye or not predicting the angry face as anger.  **EP5**: We concentrated on making our system comprehensible and maintaining the code standard so that outside concerns are addressed by professional engineering standards and norms of practice and the user can easily recognize the face and eye and forecast the outcome.  **EP6**: We can volunteer our user as a stakeholder to collect data for further improvement.  **EP7**: One major problem is constantly predicting the frames from webcam. |
| Identified and applied modern engineering and IT tools for the design and development of the capstone project. (CO5) | **K6**: We had to investigate various machine learning and deep learning algorithms in order to use the various models for drowsiness detection and risk level evaluation. We were aided by our knowledge of machine learning and digital image processing courses. Many scholars have developed numerous algorithms specifically for detecting tiredness because drowsiness while driving has been a threat to human life. Those algorithms have also been examined. Finally, we employed a variety of deep learning techniques such as Convolutional Neural Networks and other deep learning algorithms. We used the most up-to-date tools and features to carry out these operations and analyses. We concentrated on working as a team because we required a high-performance computer with a GPU system to work with photos. | **EP1**: Knowledge of image processing technologies such as OpenCV, Keras, and TensorFlow is required to appropriately apply the tools and features as well as analyze the results. We must have comprehensive fundamental engineering knowledge at the K2, K3, K4, K5, and K6 levels.  **EP2**: In order to train the dataset and acquire suitable results, we ran into a number of unusual challenges throughout the implementation phase. At the K3, K6, K7, and K8 levels, we use theory-based as well as fundamental science and engineering reasoning to address the situation.  **EP4**: Multiple people cannot be recognized in streaming video or photographs since the driver is always a single person. Google Colabatory and Kaggle can't interpret video or detect tiredness as inputs.  **EP5**: The project was developed using modern tools while following a consistent design and development procedure. We also worked on reducing flaws in our detection stage by applying logical, and ethical problem-solving methodology at the K4, K5, K6, and K7 levels while preserving professional standards. |
| Assessed and addressed societal, health, safety, legal, and cultural aspects related to the implementation of the capstone project considering the relevant professional and engineering practices and solutions. (CO6) | **K7**: The capstone project we have done detects drowsiness which helps people who are driving. Our project does not have any negative impact on the environment and poses no such threat.  Again, our capstone project has been built with a mobile app which will not cost much energy. With this view our capstone project is impacting the society for its betterment. So we can say it undoubtedly that our project does not have any negative societal impact. | **EP2**: The conflicting issue will be the using mobile phone for the whole time while driving. Another issue is capturing photos at night. But the important problem is making people aware of using this tool while driving.  **EP3**:  In the case of our capstone project the program is efficient enough, maintaining all the rules and regulations of professional engineering.  **EP4**: One infrequent problem can be the asking for ratings in an app which will stop the app immediately. A wide range of stakeholders in the R&D of the automobile industry can be included here. The appropriate feedback of the people in the automobile sector might be useful. |