Pre-processing is essential for any type of data to filter and remove noise due to low capturing quality. Every captured data has to undergo pre-processing stage as the image quality may be degraded due to light transmission properties such as scattering and absorption and environmental characteristics such as light changes, weather conditions, and hue are more or less prominent while vehicles move or if an unclear rigid scene, an unidentified hue, or poor light sensitivity is captured.

The real time videos are obtained from the dataset and initially, the videos are converted into frames and size of the frames are adjusted. As the acquired data is in 3D value the data should be resized and can be converted into grey scale image in order to reduce the complexity. Grayscale processing is performed on frames extracted from a video sequence. Grayscale pictures are just a reduction in complexity from 3D value to 1D value. As the videos are recorded in different

environmental conditions the lighting may be effected so to improve image contrast on the converted grey scale images CLAHE (Contrast Limited Adaptive Histogram Equalization)Method (Mathai, Balan et al. 2017) is applied. If the noise is present in an image the noise gets increased along with contrast so, contrast is restricted during the process.

The captured image contains more information than what is necessary such as background and different things in the image. For detection of drowsiness only the movements of eye and mouth are required. So, after pre-processing the data the detection of face and extraction of features is done in the next stage of process.

A person's facial expression can be used to accurately identify drowsiness. However, it is difficult to immediately identify the face in an image and then move on to identifying the eyes and mouth later when the face is mixed in with numerous other elements, such as the background and other objects.

In general, a system must be able to identify the object of interest, ascertain its state, and track its movements if it is to be able to monitor an object's state over time. Deep learning and computer vision identification and tracking procedures are based on gathering significant features from pixel data and creating models that can depict the region of interest. The face, eyes, and head are the primary targets of drowsiness detection models, which aim to determine a driver's level of drowsiness by observing the driver's facial characteristics. It can be challenging to identify the face and immediately extract the eyes, mouth, and other facial features when the face is featured in an image along with numerous other characteristics like the background and other elements. In these instances, the Dlib software is used to extract facial features from the pre-processed data.The Dlib algorithm is capable of identifying face position and can extract facial regions such as eye and mouth even though the driver wears eye-glasses or having beard and different hair styles. The Dlib analyses the faces again to find facial landmarks using 68-point technique. Mainly eyes and mouth regions are extracted.

## 

Face recognition in pictures or videos is straightforward, but more details about a person's face are needed, such as their posture, whether their mouth is open or closed, whether their eyes are open or closed, whether they are staring up, and other factors. Dual face detectors are offered by Dlib. One of them uses the Support vector machine (SVM), a linear classifier, the HOG Histogram of Oriented Gradient (HOG), and a window detection approach to find the face.. The other method is based on Maximum margin object detector (MMOD) using deep learning detection algorithms. HOG simplifies the images and removes the unnecessary information from the image such as background and different objects in the image. HOG guarantees fast detection in real-time and does not need a prediction model. MMOD has better accuracy when compared to HOG based detection but the detection process time is more and requires more information. In real-time MMOD needs to prepare prediction model. Thus, MMOD cannot be implemented in real-time usage. HOG can be implemented in real-time due to its fast detection process and use of less information. During the recording of video only single person is participated, this makes HOG to detect the faces with better accuracy because if number of faces are high in the image, then it becomes difficult to detect face accurately among all of them. Considering all the factors in this paper HOG is used a face detector from Dlib software.

The facial landmarks information is obtained once the face region has been identified using HOG. In DLIB software 68 pointson the person's face are used to extract features of the face. By creating a cascade residual regression tree, the face shape is gradually transformed from the current shape to the actual shape. Every leaf node of each GBDT stores a regression quantity, which is added to the input for regression, When the input data falls on each node. All residuals are finally combined to finish the face alignment process. 17 points just on chin region, 10 points at eyebrow regions, 9 points at the nose, 12 points near the eyes, and just on the mouth 20 points making up the 68 important facial features. The identified facial points of

the mouth, head, and eyes are prepared, and 68 face features are identified for blink, yawning, and head position recognition.

and mouth regions are used for the calculation of EAR and MAR calculation in later stage of the project. Each eye has six feature points, and the identification number for each point climbs clockwise starting from the left corner. The six key points of the eye are from P1 to P6. The points are marked from 1 as shown in and they not only describe feature points of eye but also the height and width of the eye region. The coordinates of the right eye, left eye regions and mouth are determined from 68 points. The points from 37 to 42 can be used for extraction of left eye and from

43 to 48 are useful for extraction of right eye measurements. Based on the position of key points the horizontal and vertical distance of the eye can be calculated which in turn helps in finding EAR and MOR values

In human face eye plays major role in identifying their emotions. If driver is in drowsy state significant eye changes can be observed such as blinking rate increases and slower the eye closure. Eye Aspect Ratio is the ratio of height of the eye and width of the eye. Using EAR, the blinking rate of the eye can be determined. The eye’s horizontal distance is denoted as the width of the eye and the distance between upper eye lid and the lower eye lid is the height of the eye. In equation form it can be represented.

The height of the eye increases when the eyes are open thereby increasing EAR value and if the eyes are closed the height of the eye decreases and the EAR value also decreases. Usually, small variation in height of the eye can be neglected due to normal blinking of the eye. In this proposed method the EAR is monitored regularly and if there is fall in EAR value is observed and if it never increases for the next frame to the normal point then it can be concluded that the driver is feeling drowsy. If the EAR increases again then it implies that driver has blinked his eye and he is in active state.

The percentage of closed eyes (PERCLOS) is the percentage of time the eye lid is closed for more than 70% to 80% of a specific time frame. Based on threshold value of PERCLOS, person's level of sleepiness can be detected. PERCLOS is a drowsiness detection tool that estimates the proportion of time that the eyelids are closed over the pupil. It can yield good results and is employed by several real-time sleepiness detection systems. To capture the eyelid closing movement and improve the precision of the system, developers employ a variety of devices. This project makes use of a dashboard-mounted smartphone that is condiured so that the driver may be seen on the image. As a result, the face can be detected more accurately, and the PERCLOS measure can be used to determine how frequently the eyelids close. Each eye has six points that are indicated using Dlib software and Euclidean distance is determined separately. The PERCLOS of the eye can be calculated using the EAR Eye Aspect Ratio. As the requirement for the calculation of eye closure is same as the EAR. So, by using Equation1 the PERCLOS of the can be calculated.

## 

Excessive yawning is a behavioural sign of weariness or a sign that the body has become more relaxed and feeling sleepy. The driver can be warned in advance if someone notice that driver is yawning. However, it should be emphasised that yawning doesn't always happen prior to the driver being tired. It must thus be supplemented with additional drowsiness indicators because it cannot be employed as a stand-alone feature.

Yawning is a sign of tiredness that the mouth expresses. To identify the yawn state, we first apply Dlib to determine the location of the mouth's lip feature points, compute the vertical distance between those points based on their location, and analyse the mouth's state depending on the time duration of opening of mouth. The fundamental step in yawn detection is extracting the mouth feature landmarks for calculating the mouth's opening ratio (MOR), and to determine the threshold and length of MOR to assess mouth state. Yawning can be detected using facial landmarks on the face. From the extracted features of mouth, the width of the mouth is detected and distance between upper lip and lower lip is taken as height(Kazemi and Sullivan 2014). The initial four steps for blink detection and yawning detection are identical. After the detection of mouth region, the next stage is extracting the mouth landmarks ranging from 49 to 68. To reduce the impact of lip thickness from the MOR calculation 62, 68, and 64 numbered featured points and the inner lips of 66, 61, and 65 are used to calculate MOR.

The mouth condition is then assessed and compared to the established MOR threshold in conjunction with the MOR duration. The inner lips' feature points are then used to draw the convex hull profile. In the experiments, the influence of mouth shape may be avoided by using a few inner lip mouth landmarks to compute MOR. Because various persons have unique mouth shapes, some studies utilise the outer feature landmarks of the mouth to compute MOR, although the MOR values for each individual is different.

The space between the longitudinal key points in the centre of the inner lip changes significantly when the mouth opens. Therefore, points 62 and 66 and 63 and 69 are chosen as the two sets of points to measure the transverse distance and the longitudinal distance, respectively. For calculation of MOR the previously mentioned six key points are used. Eye is closed if EAR is below a specific, predetermined threshold (0.3). Y N count is increased if MOR exceeds a certain, predetermined threshold (0.6). Anytime yawning is noticed, and the frequency of yawning is determined using,The number of consecutive frames with the eye in closed state is counted. The driver is in a drowsy state if the consecutive number of frames eye is closed and the YF is larger than a fixed.

Apart from facial expressions, this extraction concentrates on the driver's physiological status, which is the second most important element for detecting sleepiness. In terms of drowsiness detection, the study in physiological measures has more accurate and proven research findings. Here, the major focus is on heart rate of driver. The heart rate of an individual might change depending on a variety of physiological and psychological factors. The heart rate and pulse rate are extracted using a smart watch worn by the subject. The watch is connected with mobile phone through wireless Bluetooth connection. According to a resting heart has 60-100beats per minute for an adult. The heartbeat of the individual due to several factors such as temperature, air, different emotions and any food or liquid intake. Normally the Heart Rate(HR) of a person decreases when they are sleeping. So, based on this fact the drowsiness of the subject can be predicted. A threshold limit of 65bpm is selected so that the if the HR of the person goes below this point, then the subject is stated as drowsy. This classification is done after receiving the signals from smartwatch. The smart watch sends the extracted signals to mobile phone and later they are fused with behavioural features for the detection.

It is challenging to get a good performance of drowsy driving detection systems built using data from drivers' heart rates or drivers' facial expressions because of the individual variances amongst drivers and the different driving environments. In this paper the behaviour-physiological

characteristics are jointly modelled by combining the drivers' facial expression data with the pulse rate features collected from the driver's smart phone and smart watch using concatenation process in order to increase the accuracy and stability of the drowsy driving detection. The extracted drowsiness related features and the fusion of multiple features a drowsiness detection model is established, Transformers use the self-attention method to identify input and output global dependencies. Transformers are frequently employed in natural language processing (NLP) activities. Transformers were developed for NLP tasks after great success in that field they were then expanded to include vision tasks for computer vision. A multi head attention mechanism of transformer introduced in 2017 has outperformed many CNN based models and have achieved better accuracy. When comparing with LSTM and RNN, Transformer can be parallelized. But the disadvantage of transformer is it losses it’s position information and have to recompute entire history in context.,LSTM (long short-term memory) is an artificial recurrent neural network (RNN) architecture which can take in a sequence of input and learn dependencies over a lengthy period of time. In other words, while learning, LSTM can remember data series over extended periods of time. For monitoring driver continuously high amount of data is required and should be stored.

In contrast to standard neural networks, which are primarily used to analyse single data points, LSTMs are a special sort of recurrent neural network (RNN) that specialise in learning long-range dependencies. The existing methods analysed each single frame regularly using tradition machine learning approaches and detected the drowsiness of driver but accuracy is decreased for long time intervals. Since LSTMs have been found to be particularly good at tracking long-range dependencies without causing concerns about the fading gradients problem, LSTM is the best neural network can be applied for this model.

Although CNN is a better classifier and detect tiredness there are still issues with it, such as the fact that even though a driver's eyes are often closed, they are still thought to be drowsy. even though driver real world is awake. The accuracy can be increased by using LSTM model. The LSTM integrates the data similar to long term memory. For storing the memory and analysing the data for long dependencies LSTM won’t be sufficient to get good accuracy. For this reason, vision transformer is employed along with LSTM to solve this issue. The gradient descent problem is avoided with LSTM-transformer networks, and long dependencies are well managed.

In this paper , the LSTM-Transformer model is proposed to solve emotion classification difficulties. We propose a hybrid LSTM-Vision Transformer architecture for improving emotion classification performance. Furthermore, the hybrid model performance is evaluated by long term dependency learning. In the proposed LSTM-Transformer model the transformer architecture's positional encoding is replaced with the LSTM recurrent process. If the positional encoding is used in transformer architecture, then it was believed that the computation cost will be higher as the re- computing is needed regularly. By using LSTM in place of positional encoding the input states are preserved.

LSTM can solve short term memory problems and alone LSTM is not sufficient to solve the long- term dependencies. So, in transformer encoder layer instead of employing an Attention layer on the LSTM, we combined both the attention layer with the Multi-Head Attention Mechanism. The transformer encoder contains ReLu layer and normalization layer. Parallelization is possible using transformer layer. This allows the model to learn the long dependencies faster. This combination helps in understanding long term dependencies and improves the accuracy in prediction. A linear layer is applied in the last to so that overfitting can be avoided.

The top six ethical issues with face recognition technology include racial prejudice and disinformation, racial discrimination , data privacy, lack of explicit consent and transparent, data breaches, mass surveillance and ineffective legal support.

One of the major concerns of public in the drowsiness detection models is privacy, as there is no transparency in storing the data and how data is managed. Facial recognition violates citizens' absolute right to be constantly monitored by the government and stored their images without their permission, The data processing that takes place in or by a vehicle may not be fully disclosed to the driver or passengers of the vehicle. The information might not be released promptly and might only be disclosed to the owner of the vehicle, who might not be the driver. As a result, there is a chance that the subject may not have enough functions or alternatives to control and to utilize their rights towards data protection and privacy. This is important to know since vehicles may have several owners throughout the course of their lives if they are sold or leased rather than owned.(European data protection, Any sort of data mining raises concerns about privacy, particularly online since the majority of data is obtained anonymously. When hosted in the cloud, facial recognition technology can offer optimum data security. However, good encryption is the only way to ensure data integrity. During the installation of drowsiness detection model both the owner, driver and the third-party people who controls the data should abide to rules and regulation and below paragraphs explains.

Data breaches can cause major privacy problems for the general public as well as the government. While data breaches are a big issue for individuals, technological advancements have resulted in gains in cybersecurity and growing usage of cloud-based data. Data saved in the cloud can be secured against malicious use with an additional layer of security such as encryption, Manufacturers of vehicles and equipment, service providers, and other data controllers should make it easier for drivers to have control over their rights, they should be allowed to restrict the data used for processing and also, they should get access to control the data during entire processing. It should be their choice on what information can be shared and what is to be limited, A profile managing system needs to be installed within the vehicle to save the choices of the drivers and make it easy for them to modify the settings any time. This will simplify the process for setting adjustments. The managing system should be able to allow the driver store and access the data for each data that is processed every time. The driver should have the access to stop data collection temporarily or permanently at any time, if the vehicle is put for sale, then the owner should have access to delete the entire data which is previously stored for processing process for setting adjustments. The managing system should be able to allow the driver to remove ,store and access the data for each data that is processed every time. The driver should have the access to stop data collection temporarily or permanently at any time. If the vehicle is put for sale, then the owner should have access to delete the entire data which is previously stored for processing.

When it came to treating victims, every kid at the ambulance had at least one call that remained indelible—maybe a multi-car crash on the highway, maybe a cardiac arrest or a house fire or a head injury—that introduced us to a world of grief we hadn’t known before, that took us behind the veil of our town. I recall responding to a daytime suicide, at a house not more than a mile from my own, and when we spilled out of the ambulance and hustled through the strobes in our bright uniforms, hoping to save the overcast day, fix the wrong, piece back the body—crazy-competent mini-adults that we were—one unimpressed police officer stopped us short on the doorstep.

"You’re not going in there," he said. When we insisted, he exhaled an exasperated sigh and added, "She slit her goddamn wrists in the tub, and you’re kids, and I’m not letting you in there." I remember we protested, outraged that he’d called us "kids," and we wouldn’t leave the scene, waging our own quiet sit-in, until we were finally called off by an adult adviser. But even as we worked ourselves into a bruit, I had this nightmare image of a submerged naked body, blood streaming from her wrists, face twisted in some ghoulish rictus.

Half an hour later, I was sitting back in calculus, trying to figure out a derivative.

Vehicle manufacturers, data controllers and service providers should take measures in order to protect and control the data to make sure it doesn't fall into unauthorised hands. The following are the data privacy measures should be taken for securing the data, By state-of-art algorithms, Communication channels should be encrypted, A unique encryption managing system should be installed in every vehicle, If the data is stored it should be encrypted every time it is stored. Encryption keys should be renewed regularly, Regularly verifying the devices which receives data. Protecting the integrity of data. The proposed non-intrusive model requires a smaller number of components for detection of drowsiness. The devices used are common and available easily in real world.

Smartphones are widely available these days, . It has become one of the most essential devices in our daily life. so, a driver or automobile owner does not need to acquire additional gear for this purpose. No extra add-ons are required for this system.

At present everyone is looking for healthy lifestyle and maintaining proper diet and exercise regularly. To keep updated with the health fitness smart watches are introduced in the market and are prevalent in the market at very low price. So, this won't be an extra burden for the owner or a driver to use a fitness watch. Because smartphones and watches have Bluetooth connectivity and Internet access, emergency situations may be handled quickly using SMS and other online services. With the use of handy devices, the drowsiness of the driver can be detected and can help in preventing the accidents is the special advantage of this model.

The initial cost and the maintenance cost for the used components is very low which can be an advantage for the owner/ driver to install the system . In the below TABLE the total cost involved..

Initially, however, I remember a lot of time spent blowing air into those manikins, real lip to synthetic lip, thrusting palms down on fake chests loaded with thick springs, and, at the end, paper readouts issuing from a slot at the ribs, a ticker showing the peaks and valleys that gauged one’s efficacy at giving CPR. Repetition made for perfection on those fake bodies, though reality, I would soon find to my dismay, could be different. When the grandfather of the boy next door keeled over on the lawn, I lined my palm up on his sternum as I’d been instructed—and had succeeded at so many times before on the dummies—and with the first thrust felt three real ribs give way.

for setting up the system is calculated. Once the initial products used in this system have long life time has the recent smartphones are capable of updating themselves for the longer use. No additional maintenance is required for a smart phone and watches.

No special training is required for drivers in order to install and use the proposed detection system because almost every individual have good experience in handling mobile phones and watches these days.

The total cost for the production and the maintenance of the proposed model is mentioned in table2. From the above data it can be concluded that the proposed system is a low cost and one time investment model which doesn’t need extra capital for maintenance and known for its longevity, this paper we have discussed the use of various metrics to summarise the previous papers. To explain the evaluation criteria, we specify the following characteristics to determine a system's benefits and drawbacks, which were considered as areas of interest and concern in previous papers: Accuracy: the percentage of non- drowsy and drowsy classifications.

Eg: Privacy, Data size, Cost,Privacy, Vulnerability, Adaptability

**Growing up,** we had this sort of unusual thing in our town: an ambulance service operated by kids. It’s still there today, in fact—thriving. Then, it was housed in a defunct red train station that rattled every time a passing commuter train rushed by on its way to Manhattan. In winter, icy gusts came lunging through the walls. There was a garage with two ambulances, and off it, a cramped radio room. Inside the station was an open common area where presumably tickets had once been sold, but which now hosted our training sessions and organizational meetings. Upstairs there was a loft where the CPR manikins were stashed. Sometimes you’d forget and go up there at midnight to turn out a light and nearly have a heart attack at all those synthetic bodies laid out, staring dumbly at the ceiling.

loss:

Data loss might happen during recording because devices can fall or get detached, bad weather conditions can disrupt recordings or changes in lighting can make it difficult to record behaviour of driver. Loss of data does not mean that monitoring of drowsiness is stopped. As we employed an adaptive hybrid model with several drowsiness feature sources. The research done by (Samiee, Azadi et al. 2014) discussed about the continuation in monitoring drowsiness even though the data loss occurs sometimes. During driving the system should be able to work even loss of data occurs at any time. In this paper we have adopted the (Samiee, Azadi et al. 2014) method for continuing data monitoring even during the loss of data.

Time :

The time taken for extraction of features , pre-processing of images and the duration for the entire experiment is one of the important factors to be considered when monitoring driver's drowsiness. If the time length for extraction of features is more than detection of early symptoms of drowsiness is not possible. If the processing of system takes more minutes to detect by the time the alert is generated an adverse event can be happened by the. Delay in processing the images may even generate wrong results as driver may feel drowsy at any time.

For the proposed model with use of hybrid LSTM-ViT the model has achieved an accuracy of 94.57% and the drowsiness is detected at its early stages with the help of behavioural and physiological features.

The proposed hybrid model is reliable in real world conditions and the non-intrusive method will help in preventing the accidents by early detection of drowsiness of the driver. On

AI:

comparing with the previous models proposed by different researches our model has performed better and obtained satisfying results and can be implemented in real-time. The model has obtained an accuracy of 92 percent in detecting the drowsiness the driver. The model has been trained and two different datasets named UTA-RLDD for drowsy related videos and YAWNDD dataset which contains videos where yawning by people at different situations is recorded. The below table explains the advantages and disadvantages of the previous existing models and the proposed model.

**The accident—the first one**—occurred on the Wednesday night before Thanksgiving of my senior year in high school. It left one friend injured and one dead, and for a while afterward the whole thing seemed so surreal and impossible that all we could do—friends, family, anyone connected but not in the accident itself—was try to re-create the simultaneities of that evening, the first person at the scene, the shock of the couple at the nearby house from which the call was made for an ambulance, and then: who called whom, and who was where when they heard. Given our own shock, we couldn’t imagine the parents of the victims hearing those first words: There’s been an accident....

When the news reached my family that night, in that orbit of calls, my parents, perhaps like other parents among our friends, presumed their child might have been in the car, which wasn’t the case, though might have been, had I made a different decision earlier that evening. For us seniors, it was a free night with no school the next day, a holiday from everything, including our cursed college apps. Mine was spent with my girlfriend, so I missed the pre-party and then the ride to the real party. And so I missed the accident, too