GeolocationAlert System For Drowsiness Detection

[1]Megha N, [2]Ranjitham M, [3]Rufina Mariam, [4]Jahnavi K, [5]Srinidhi N.N

Department Of Computer Science Engineering, REVA University, Bangalore.

Abstract.

India alone has had over three hundred thousand road accidents in 2020. Almost 30% of the road accidents have been proven to be fatal because of lack of aid. In order to fix this issue a location-based drowsiness detection system has been proposed. By leveraging points of interest on the face recognised by the camera and communicating to Convolutional Neural Network, the proposed system identifies drowsiness and sends in multiple alerts to not only alert the driver but also his emergency contacts. The given CNN-based model can be used to develop a highly accurate and user-friendly real-time driver sleepiness detection system for various hardware devices and smartphones.

Keywords. face detection, CNN algorithm, driver drowsiness, deep learning, artificial intelligence

1. INTRODUCTION

There has been a significant increase in the demand for modern forms of transportation over the years, which prompted a quicker consolidation of a report that states the vehicles that are popular in the area, called the car-parc data. A total of 63.8 million automobiles will be sold globally by 2020. This was a decrease from the previous year, as the industry saw a downward trend as the global economy slowed owing to the coronavirus outbreak that had swept across all countries. In the following year, global car sales increased to roughly 66.7 million vehicles in 2021, resulting in a rise in the graph.While the automotive industry has revolutionized the lives of people by making things notably easier, it also brought in adverse repercussions like traffic accidents.

India had 3,54,796 road accidents in 2020, with 1,33,201 people killed and 3,35,201 wounded. These numbers have been given out by the National Crime Records Bureau’s research. In this study, fatigued driving was found to be responsible for around 20% to 30% of traffic accidents. Due to which, drowsy driving is a consequential and often overlooked danger in traffic incidents. Owing to the alarming rise in accidents by the year, a fatigue or drowsiness detection system made for drivers has been a topic of research and implementation but even though it’s been a hot topic to research about, not many proposals cover a very important aspect of the drowsiness detection system, location based alerts.

While it is beneficial to alert the driver that he is not in his best frame of mind to drive, it is also incredibly useful and precautionary to send a location alert to the driver's emergency contact or employer so that if something untoward does happen, help can arrive immediately.

When it comes to detection methods, it can be categorized into two different types of methods such as - subjective methods and objective methods. A driver must take part in the subjective detection method's evaluation, which involves self-questioning, appraisal, and questionnaire completion and is connected to the driver's subjective opinions.The information is then used to foresee the quantity of vehicles which are being driven by drivers who are sleepy, permitting them to productively detail their courses more. The objective location strategy, then again, doesn't need drivers' input since it persistently analyzes the driver's physiological status and driving way of behaving. The data accumulated is used to decide if the driver is tired. Furthermore, objective detection may be classified into two categories: contact and non-contact. Because it does not require a sophisticated camera, non-contact is less expensive and more suitable than contact, allowing it to be utilised in a wider range of cars.

To address the inconsistencies and issues associated with road accidents caused by a person's weariness, we have developed a detecting system based on the CNN algorithm. This system has a 95% accuracy rate, making it a very good system for alerting the location when the driver is drowsy on the market. Furthermore, the system is low-cost to deploy, allowing it to be extensively adopted and reach a large number of drivers.

2. LITERATURE SURVEY

Sleep-related crashes account for a considerable fraction of all car collisions worldwide and due to this academics and automakers have devised a variety of remedies, ranging from identifying trends in driving behaviours to measuring the driver's brain waves and vitals while driving. The majority of these solutions are underpinned by statistical and machine learning-based prediction algorithms.The study takes into account current technologies as well as research on the project's topic. It's an attempt to have a better grasp of the research efforts that have gone into this subject, as well as to establish where we should focus our efforts when working on this project.

Multi-access edge computing is a low-latency approach (MEC). Instead of using a central cloud server, MEC technology provides processing and storage capabilities to the edge of the mobile network. MEC has been tested in a variety of mobile applications and has shown to be faster than cloud-based solutions. Over 5G networks, the implementation of MEC-based DDD systems will result in real-time judgements, ensuring the driver's safety[1].

Dua et al [2] identifies the driver's fatigue. The four models which are based on deep learning techniques make up this architecture are ResNet, AlexNet,VGG-FaceNetand FlowImageNet. Face expressions, head motions, hand gestures and behavioural traits such as head, eye, and lip movements are all extracted using these models.The VGG-FaceNet model identifies and extracts face features, whereas AlexNet model adjusts for various ambient and background variables. Head gestures and behavioural characteristics are extracted using FlowImageNet, while hand gestures are extracted using ResNet. The models chosen by Dua create four outputs from the RGB film of the drivers, which are sent to an ensemble technique termed simple averaging, which is then followed by a SoftMax classifier. 's approach has an overall accuracy of 85, according to Dua et al.

Yassine uses the same "awake" and "drowsy" categories as our method, but excludes the "low-vigilant" videos. The authors employ roughly 100–120 photos per recording in this example. The fatigue detection model is a basic CNN that was built from the ground up and consists of five convolutional layers, a flatten layer, and a dense layer. Although the overall accuracy on the test photos is 69 percent, due to the high number of false positives (56 percent of the images examined), this model cannot be considered for a genuine implementation on an ADAS. It should be mentioned that, according to the author's code, the random train/test split of data (80% train/20% test) was conducted after the photos were extracted. As a result, accuracy might have been harmed if frames from the same movie were utilised in both the train and test sets. Because each video is shot under different settings, the network may be learning to distinguish the scenario depicted at a given frame (person, perspective, lighting) rather than the subject's exhaustion [3].

The feasibility of establishing a hybrid measurement-based detection system, such as vehicle-based, physiological, and behavioural signals was investigated by Gwak et al[4]. This inquiry involved a total of sixteen persons. From the calculated data and videos, a total of 80 features were retrieved. There were three components to the study.

At first, the drivers' physiological signs, driving execution, and social tiredness markers were recorded utilizing a driving test system and checking framework.After that, two distinct classification methods were used in order to classify: RF classifier and majority voting, employing logistic regression, SVM, and KNN. Backward feature selection was carried out consecutively in the case of a majority vote, followed by classification.Four approaches are utilised to categorise films into one of the three available categories in [5], the study where the UTA-RLDD was described. The HM-LTSM network, with 65.20 percent accuracy, achieves the highest global accuracy among these approaches, with global accuracies ranging from 57 to 65 percent. The accuracy on both awake and sleepy footage is impressive, hitting an impressive 80 percent in both. Despite the encouraging results, false alarms would be raised in 19% of the cases, therefore the system might be modified to lower this percentage. It's also fascinating to compare them to a human judgement baseline, in which four volunteers evaluated each video's sleepiness degree. The accuracy of human judgement was 57.8%. The rate of false positives is greater in 37 percent of situations, however, when employing human judgement, the system notifies the driver needlessly.

Sun [6] is the first to introduce DCNN based on CNN to recognise important human face points, due to the advancement of deep learning. Themodel detects five facial essential points, but it does it quickly. Zhou [8] used FACE++, which optimises DCNN and can detect 68 facial key points, to gain a greater accuracy for facial key point identification, however this approach contains too much of a model and the operation of this algorithm is quite hard.To enhance different layers of CNN, Wu [7] presented Tweaked Convolutional Neural Networks (TCNN), which is based on Gaussian Mixture Model (GMM). However, TCNN's robustness is too reliant on data. Deep Alignment Network (DAN) was put in place by Kowalski [9] to distinguish face important points, and it outperforms most algorithms.Large models and calculations based on complicated functions are required by

DAN. DriCare uses Dlib [9] to differentiate critical spots on the face in order to meet the real-time performance requirement.

A large number of academics are working to find a remedy to road accidents caused by drowsy drivers. The numerous study findings have been divided into several categories.

Zhang et al. (2020) developed a sleepiness detection system for drivers on the basis of Karolinska Sleepiness Scale (KSS) [10]. A Mixed Effect Ordered Logit (MOL) model is combined with a Time Cumulative Effect (TCE) model in the proposed approach (TCE). The MOL-TCE model was compared to non-MOL-TCE models in an experimental study, and the findings demonstrate that the recommended model is 62.84 percent more accurate compared to the present models. Anevaluative method to detect driver fatigue was created by McDonald et al. [11]. The approach, which was combined with the Dynamic Bayesian Network algorithm (DBN), gave way to a false-positive rate which is much lower than the previous PERCLOS method for identifying driver drowsiness.Wang et al. [12] compared the performance of classical machine learning and deep learning algorithms to picture categorization. The research was carried out on both large and small datasets, including the MNIST and COREL1000 datasets. The conclusions derived showcased that the traditional form of machine learning functions tremendously better on small datasets, whereas deep learning performs better on large datasets.Kumar et al. concentrated on the utilisation of embedded systems and signal processing technologies to create surveillance systems [13]. For improved vehicle control, the system focuses on three deciding factors: identifying driver tiredness, alcohol intake, and crash detection. The experimental findings suggest that this technology is more efficient and accurate than the present analogue system.

Moujahid et al. compared an efficientface descriptor for identifying driver sleepiness to the NTH Drowsy Driver Detection (NTHDDD) dataset. In terms of performance, the proposed framework has been proved to be as effective as a convolutional neural network.

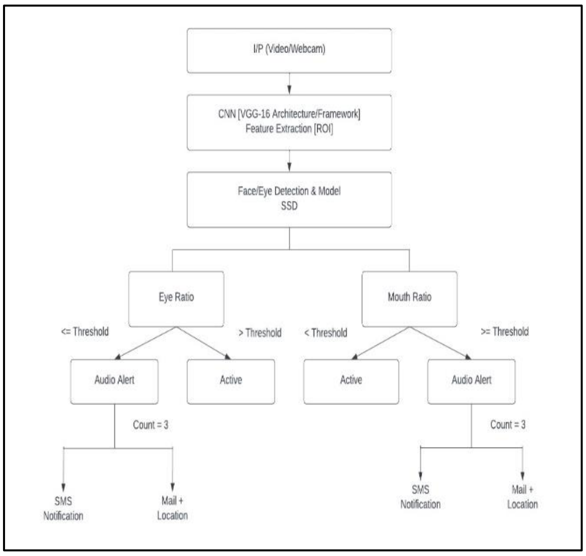
Using the wavelet packet transform, Phanikrishna et al. [15] developed an automated classification system for identifying driver drowsiness.From the driver's single-channel Electro-Encephalogram (EEG) data, the wavelet packet transform was created. The suggested model performs the real-time sleep analysis with a precision of 94.45%. Li et al. [16] suggested a technique for detecting urban crimes by assessing facial expression and emotion correlations. The Facial Expression Recognition (FER) strategy was created and used to decide a user’s emotion in view of their look, and the outcomes were contrasted with the Kernel Density Estimation (KDE) way to deal with show a connection among mind-set and driving example.

The project implemented is an effective way of extracting eye features and determining the drowsiness state of the driver.

3. METHODOLOGY

The system suggested is designed for automating the surveillance over any person who is on the road, driving. The entire project is done real time so that any sort of mishaps on the road can be controlled efficiently.

With the help of a webcam, images are captured and processed as the input. Our selected regions of interest are the eyes and the mouth. A CNN classifier processes those images and a eye-aspect ratio and the mouth aspect ratio is formulated. If the ratio is lesser than a specific threshold value for eyes then an audio alert is played. If the mouth aspect ratio is greater than the threshold then the yawn counter increases. When the yawn count reaches a specific value, an audio alert is sent. If three audio alerts are sent to the driver then the driver's emergency contacts are notified via SMS and email. In conjunction to that the real time location of the driver is also sent. Fig 1 seamlessly depicts the workflow of the project.



An endless loop has been developed to access the camera, which catches every frame and stores the picture in a frame variable. This is done by the means of OpenCV.

SSD is intended for object recognition proof continuously, and finding regions of interest in the proposed project is used. R-CNN is quicker since it uses a locale proposition organization to develop limit boxes and afterward utilizes those containers to sort things. The whole interaction runs at a pace of 7 frames each second, which is viewed as forefront concerning accuracy. Undeniably not exactly is expected for constant handling. SSD speeds up the cycle by eliminating the requirement for the locale proposition organization. To make up for the deficiency of accuracy, SSD offers an assortment of developments, for example, multi-scale elements and default boxes. These upgrades permit SSD to match the exactness of the Faster R-CNN while utilizing lower-quality pictures, coming about in considerably quicker execution. It accomplishes ongoing handling speed and even surpasses the Faster R-CNN regarding exactness.

To empower discoveries of different scales, the SSD network utilizes highlight maps from many layers of a changed VGG16 system.

A CNN Classifier is used to assess if the driver is sleepy because the attention is mostly on the eye and the mouth. Convolutional Neural Networks are a type of deep neural network used inimage classification and computer vision. A primitive neural network is made up of three layers: a layer which is used for taking in input, a hidden layer, and a layer which gives the output. Two or more hidden layers can be found in a deep neural network. Convolution layers are followed by a fully connected neural network in a convolutional neural network.

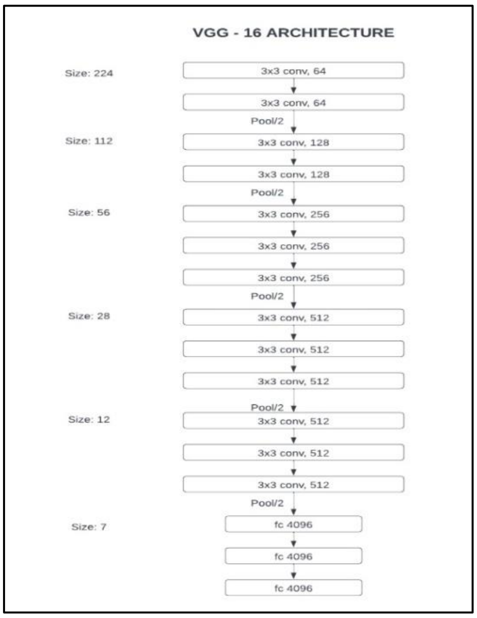
The image's characteristics are extracted via the convolution layers (input). A tiny filter or kernel analyses the picture, extracting features such as vertices and horizontal lines, and creating a feature map. Following the convolution layer, the pooling layer is the next layer. The feature map obtained by the convolution layer is effectively downsampled by the pooling layer.The precise site of the feature is included in a feature map produced by the convolution layer. This might lead to overfitting. A filter is applied to the features map by the pooling layer. It only uses the information relevant to that filter that is decided on the basis of the pooling technique selected, either the average of the values coming under the filter or the maximum.

This reduces the feature map's spatial dimension and turns the feature's precise spatial information to rough data.This reduces the risk of overfitting. Any convolution and pooling layer can be bracketed together on the grounds of the complexity of the input data set. After that, the last pooling layer is flattened and converted into a one-dimensional array, which is then passed on to the fully linked layers anticipating the output.

The convolution neural net (CNN) architecture applied in the proposed system is VGG16. As one amongst the most prominent vision models to exist, it is highly competent in completing the job. Rather of having a large number of hyper-parameters, VGG16 focused on 3x3 filter convolution layers with a stride 1 and always used the same padding and maxpool layer of a 2x2 filter stride 2. The convolution and max pool layers are placed in the same location throughout the design. It features two FC (completely connected layers) and a softmax for output in the end. The architectural flow is depicted in Figure 2.

The pre-calculated eye-aspect ratio should be less than threshold value and the mouth-aspect ratio must be more than threshold. If the former is violated then the driver is alerted by an audio file and if the latter is violated then the yawn counter steadily increases; when the yawn counter exceed a certain value, an alarm is played to the driver. However, if the score keeps increasing, the alarm is played thrice and an SMS is sent to the emergency contacts via an API.

Furthermore, the most important point is that this system sends the location of the driver as well using the geocoder module in python. The module allows to access the latitude and longitude of the driver, thereby giving the location via email.



4. EXPECTED RESULTS

The expected result accomplishes a number of things:
Accurate prediction

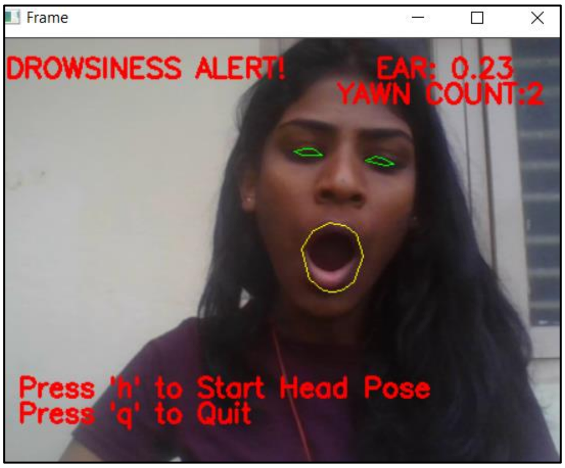


Fig 3

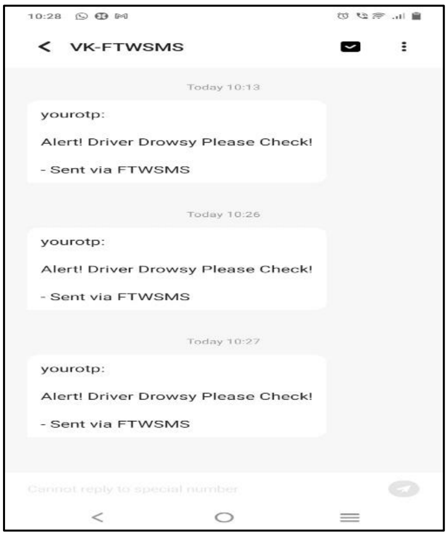


Fig 4

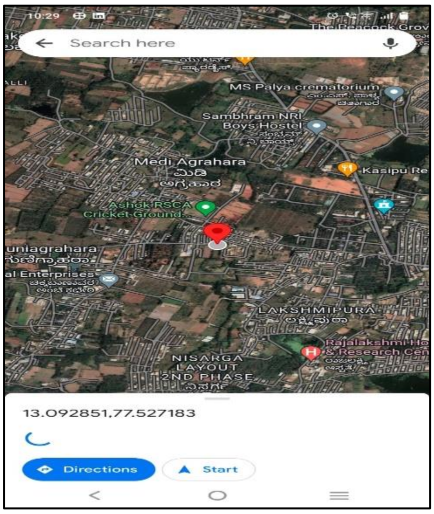


Fig 5

5. CONCLUSION

A CNN-based Deep Learning-based sleepiness detection system was described in the paper. The aim is to develop a lightweight system that could be employed in numerous devices and will most importantly alert the location of the driver as and when required. The proposed model was able to identify drowsy driving behaviour by recognising facial landmarks in collected photographs and feeding the information to a CNN-based trained Deep Learning model. The accomplishment was the development of a deep learning model that is relatively small in size yet extremely accurate. This technology might be readily integrated into next-generation car dashboards to provide increased driver-assistance systems, or even a hand-held device to intervene when drivers become drowsy.

6. REFERENCES

[1]. Abbas, Q.; Alsheddy, A. Driver fatigue detection systems using multi-sensors, smartphone, and cloud-based computing platforms: A comparative analysis. *Sensors 2021, 21, 56.*

[2] Dua, M.; Singla, R.; Raj, S.; Jangra, A. Deep CNN models-based ensemble approach to driver drowsiness detection. *Neural Comput. Appl. 2021, 33, 3155–3168.*

[3] Yassine, N. Artificial Intelligence Techniques for Driver Fatigue Detection. Ph.D. Thesis, *Oxford Brookes University, Oxford, UK, 2020.*

[4] Gwak, J.; Hirao, A.; Shino, M. An investigation of early detection of driver drowsiness using ensemble machine learning based on hybrid sensing. *Appl. Sci. 2020, 10, 2890*

[5] Ghoddoosian, R.; Galib, M.; Athitsos, V. A Realistic Dataset and Baseline Temporal Model for Early Drowsiness Detection. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Singapore, 8–10 December 2019.*

[6] J. Valmadre, L. Bertinetto, J. Henriques, A. Vedaldi, and P. H. S. Torr, "End-to-end representation learning for correlation filter based tracking," *in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), July 2017, pp. 5000–5008.*

[7] Y. Wu, T. Hassner, K. Kim, G. Medioni, and P. Natarajan, "Facial landmark detection with tweaked convolutional neural networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 12, pp. 3067–3074, Dec. 2018.*

[8] M. Kowalski, J. Naruniec, and T. Trzcinski, "Deep alignment network: A convolutional neural network for robust face alignment," *in Proc. IEEE Conf. Computer Vision and Pattern Recognition Workshops (CVPRW), July 2017, pp. 2034–2043*

[9] C. Ma, J. Huang, X. Yang, and M. Yang, "Robust visual tracking via hierarchical convolutional features," *IEEE Transactions on Pattern Analysis and Machine Intelligence, p. 1, 2018*

[10] X. Zhang, X. Wang, X. Yang, C. Xu, X. Zhu et al., "Driver drowsiness detection using mixed-effect ordered logit model considering time cumulative effect," *Analytic Methods in Accident Research, vol. 26, no. 9, pp. 100114, 2020.*

[11] Basha, S. M., Poluru, R. K., & Ahmed, S. T. (2022, April). A Comprehensive Study on Learning Strategies of Optimization Algorithms and its Applications. In *2022 8th International Conference on Smart Structures and Systems (ICSSS)* (pp. 1-4). IEEE.

[12] P. Wang, E. Fan and P. Wang, "Comparative analysis of image classification algorithms based on traditional machine learning and deep learning," *Pattern Recognition Letters, vol. 141, no. 11, pp. 61–67, 2021.*

[13] V. S. Kumar, S. N. Ashish, I. V. Gowtham, S. P. A. Balaji and E. Prabhu, "Smart driver assistance system using raspberry pi and sensor networks," *Microprocessors and Microsystem, vol 79,pp 1-11,2020.*

[14] A. Moujahid, F. Dornaika, I. Arganda-Carreras and J. Reta, “Efficient and compact face descriptor for driver drowsiness detection,” *Expert Systems with Applications, vol. 168, no. 12, pp. 114334, 2021.*

[15] V. Phanikrishna and S. Chinara, “Automatic classification methods for detecting drowsiness using wavelet packet transform extracted time-domain features from single-channel EEG signal,” *Journal of Neuroscience Methods, vol. 347, no. 3, pp. 108927, 2021.*

[16] Z. Li, T. Zhang, X. Jing and Y. Wang, “Facial expression-based analysis on emotion correlations, hotspots, and potential occurrence of urban crimes,” *Alexandria Engineering Journal, vol. 60, no. 1, pp. 1411–1420, 2021.*