E-ISSN 2581 – 7957
P-ISSN 2277 – 3916
CVR Journal of Science and Technology, Volume 23, June 2022
DOI: 10.32377/cvrjst2318

Limitations of CNN-Model and Enhanced AI-Model for Driver Drowsiness Detection

S. Nikhila[1] and V. Sidda Reddy[2]
[1]PG Scholar, CVR College of Engineering/ IT Department, Hyderabad, India
Email: sistla.nikhila@gmail.com
[2]Assoc. professor, CVR College of Engineering/ IT Department, Hyderabad, India
Email: siddareddy.v@gmail.com

Abstract: **Contemporarily driver drowsiness is one of the major universal facts of road accidents across the globe. Integrating enhanced Information Technology (IT) for instance Neural Network (NN), Computer Vision (CV), and Image Processing (IP), enormously reduce road accidents from drivers' drowsiness while driving. Advanced computer vision technology, artificial neural network, and intelligent cameras dynamically predicate as well as alert the drivers when they are in drowsiness. Drowsiness Detection (DD) has been an important research domain in real-time biomedical and traffic signal applications. Recently various Deep Learning (DL) algorithms implemented to study and diagnose fatigue situations in Electroencephalograms Signals (EEGs). The primary objective proposed in the present article is to study a survey of literature on drivers' drowsiness detection based on driver behavioral measures by using computer vision and artificial neural network algorithms. Furthermore, in this article, traditional drivers' drowsiness architecture models, limitations, challenges, and conventional neural network algorithms have been addressed. As regards to study also addressed the role of enhanced convolution neural networks and limitations in the present and future scenario to detect and alert drivers' drowsiness during driving which leads to reduced road accidents in the future globally.**

Keywords: **Computer vision, convolutional neural network, drowsiness detection, deep learning, and information technology**

I. Introduction

Due to the extensive efforts of research studies and government organizations over the last three decades, changes in driving conditions and driver safety have been observed. The present study: addressed various major driver drowsiness behavioral measures that cause around 1.3 million population die every year in traffic accidents, as well as another 20 to 50 million, who are being suffered from non-fatal injuries [1]. Fig 1 illustrated various levels of driver drowsiness based on behavioral measures during driving.

Face Symptoms	Output
Eyes Open and no Yawning	Not Drowsy
Eyes Blinking and Yawning	Less Drowsy
Eyes closed over 1.5 seconds	Drowsiness

Figure 1. Levels of Drowsiness

Drowsiness and exhaustion, which occur shortly after high speed and alcoholism, are the contributing factors of traffic injuries in a number of industries, including aviation, the army, and transportation [2]. However, in recent years, Drowsiness Detection (DD) research has piqued people's curiosity [6, 7]. In the present Covid-19 pandemic, when medical equipment is ubiquitous, this is a major problem [8].

Drunken driving is a major cause of road accidents all over the globe, but especially in the USA. According to statistical analysis from the National Highway Traffic Safety Administration (NHTSA), about 90,000 crashes were caused by drowsy driving between 2015 and 2017, with around 4000 persons killed [1]. There are various factors that cause people to fall asleep while driving; one study found that driving for a lengthy period of time produces a loss of self-control as well as a focus [1]. Drowsiness will impair a driver's ability to notice the environment and drive safely. Drivers, on the other hand, will not stop driving even if they fall asleep, according to the majority of them "I will be fine, I can continue driving". According to the National Sleep Foundation (NSF), several earlier indicators of drowsiness, such as frequent blinking, yawning repeatedly, eye closing continually, mouth opening, and/or keeping his/her head up, can alert a motorist to stop and rest [2].

II. Related Works

Two pruning techniques for artificial neural network structure optimization are discussed in this study [3]. Nonlinear dynamic systems will be modeled using these networks. The accuracy of these algorithms will be tested and compared using real steam turbine data from a thermal power plant. To capture the behavior of a selected real system, four distinct models will be created: a neural network model, an adaptive neuro-fuzzy model, and two optimized neural network models utilizing the Optimal Brain Damage and Optimal Brain Surgeon algorithms.

The prevalence of diabetes in adults aged 20–79 years was estimated at 8.8% in 2015 and is expected to increase to 10.4% by 2040 [4]. Diabetes' high incidence in adults has significant social, economic, and developmental ramifications. Governments are under increasing pressure to establish policies that reduce the risk factors for type II diabetes and gestational diabetes, as well as gestational diabetes, and also to ensure that certain diabetics have adequate access to care. Taking on the global impact of diabetes is a huge undertaking, and the IDF continues to be

Received on 08-09-2021, Revised on 21-02-2022, Accepted on 14-06-2022.

a voice for diabetics, by informing individuals and governments about the activities that can be taken to prevent and manage the disease.

Figure 2. Example of Training Data

The topic of the comparison of optimization strategies employing gradient descent, gradient descent with momentum, Adam, and learning rate investigate decay in conjunction with previous optimization techniques for feature extraction utilizing deep neural networks in this article [5]. To correctly recognize the sex of a person caught, this study analyses many settings of these strategies as well as techniques over themselves using a mixture of regularization approaches that vary in their use from no regularization, L2 regularization, and dropouts.

Convolutional networks can be significantly deeper, more accurate, and efficient to train if they feature fewer links between layers near the input as well as those near the output [6]. In this research, the authors embrace this discovery and introduce the Dense Convolutional Network (Dense Net), feed-forward networking in which every layer is connected to any other layer. Our network contains L (L+1)/2 direct connections, whereas standard convolutional networks with L layers have L connections—one between each layer and its succeeding layer. All previous layers' feature maps are utilized as inputs into each layer, and their own feature maps are used as inputs into all subsequent layers.

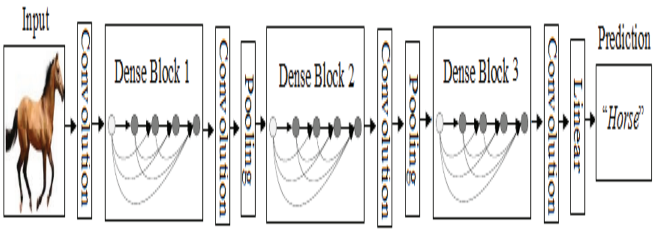

Figure 3. Deep dense-net three Dense Blocks

It's more difficult to train deeper neural networks for image identification using deep residual learning. They propose a residual learning strategy for training significantly deeper networks than previously used networks [7]. They explicitly reformulate the layers as learning residual functions with reference to the layer inputs, rather than learning unreferenced functions. They show extensive empirical evidence that residual networks are easier to tune and can achieve accuracy from greater depth. They look at residual nets on the Image-Net database with a depth of up to 152 layers, which is 8 layers deeper than VGG nets but still has less complexity. An ensemble of these residual nets scores 3.57 percent error on the Image-Net test set. The ILSVRC 2015 classification task awarded this effort first place. There's also a CIFAR-10 analysis with 100 and 1000 layers. Many visual recognition tasks require a high level of representation depth. They can only achieve a 28 percent relative improvement on the COCO object identification dataset because of their extraordinarily deep representations. Our contributions to the ILSVRC and COCO 2015 contests used deep residual nets as the foundation, and they won first place in the Image-Net detection, Image-Net localization, COCO detection, and COCO segmentation tasks. Rethinking the inception architecture for computer vision, most state-of-the-art computer vision solutions for a wide range of tasks used convolutional networks as their foundation [8]. Since 2014, very deep convolutional networks have become popular, resulting in significant improvements in a variety of benchmarks. Although the increased model size and computational cost typically result in immediate quality gains for most tasks (as long as enough labeled data is provided for the training), computational efficiency and low parameter count are still enabling factors for various use of the cases such as mobile vision and big-data scenarios. They are looking into ways to scale up networks in a way that makes the most of the extra computing by using appropriately factorized convolutions and aggressive regularization.

Deep Face Recognition, Face recognition is done from a single snapshot, or a group of faces tracked in a video is the purpose of this paper [9]. The use of a Convolutional Neural Network (CNN) for end-to-end task learning, as well as the availability of very large-scale training datasets, have both led to recent success in this area. They make two contributions: first, they show how a large dataset (2.6 million photos, over 2.6 thousand people) can be built using a combination of automation and human in the loop, and they analyze the trade-off between data purity and time; second, we walk you through the complexities of deep network training and face recognition to show you how to get equivalent state-of-the-art results on the common LFW and YTF face benchmarks. Deep Face Recognition with Keras, Face recognition has always been a difficult problem for science fiction and science [10]. To hide her identity, a woman colored her hair or wore a hat. To learn, deep learning tasks often expect to be fed several instances of a custom class (e.g. lots of pictures of someone). This is satisfactory for the face recognition problem because training should be done with a limited number of examples — usually, just one shot of a person exists. In addition, introducing new classes should not necessitate re-creating the model.

Driver fatigue detection survey from internal sources [11], the number of traffic accidents occurred every year owing to driver weariness based on Eye Tracking. As a result, a system to identify early driver drowsiness and issue a warning signal may be required to avert many road accidents, as well as to reduce personal suffering and save money. As a result, the authors created a method in which the camera may be mounted in front of the driver's seat for quick fatigue detection. As a result, if driver fatigue is detected while driving, this system will alert the driver immediately. They used video files collected from cameras in this system, then frames were taken from these video files and the eyes region was tracked to determine the distances between open and closed eyes. When the drivers' closed eyes are detected back-to-back for a few frames, the system determines that the driver is falling asleep and alerts the driver to save their life. According to the system design, the driver images or frames will be extracted from the video file and used to detect facial descriptors using the complexion-based technique on the frames.

Generally, the eyes are placed on the upper half of the face location, so the remaining lower half of the face portion will be removed for easy searching of eyes locations. Later from the top of the face, this system can calculate horizontal averages. The heavy changes in horizontal averages then can detect the eyes in closed or open states. If the eyes are closed for some consecutive frames, this system detects the driver's fatigue then it can generate the warning alarm to alert the driver.

Detecting drowsiness for drivers using a condition-adaptive representations learning approach, we develop a condition-adaptive representations teaching model based on a 3D-deep convolutional neural network for detecting driver sleepiness [12]. The suggested approach is built upon four different models: spatiotemporal representations training, scene state understanding, feature fusion, as well as drowsiness detection. Around the same time, the Spatio-temporal representation modeling collects characteristics that may define motions as well as appearances via video. The scene condition awareness identifies the scene conditions linked to different circumstances regarding the drivers and driving scenario, such as the status of wearing glasses, the lighting condition of driving, and the movements of facial features such as the head, eye, and nose as well as mouth.

Using two features taken from the previous models, feature fusion creates a condition-adaptive representation. Using the condition-adaptive representation, the detection model detects the driver's sleepiness [13]. The condition-adaptive representation educational approach will retrieve extra discriminative characteristics more unique to every scene situation than that of the generic representation, enabling the drowsiness detection approach to provide more accurate results in a variety of driving conditions. The suggested approach is tested using the video dataset NTHU Drowsy Driver Detection.

Various classification EEG signal models have been proposed [14] which are based on the Support Vector Machine (SVM) approach even performance challenges due to background noise. SVM benchmark SVM models predicated fatigues and drowsiness are the primary causes most of traffic accidents. Fatigue detection mainly focuses on the driver's facial behaviors for instance facial expression, eye blinking, and mouth yawning. Open CV and Dlib libraries were utilized to detect the expressions of drivers' faces. Benchmark SVM approaches predicated on five fatigue features such as count of a yawn, internal zone of the mouth opening, count of eye blinking, PERCLOS, and head are used to detect drowsiness from real-time video. At the same time, facial expressions were trained with SVM. In this study, SVM-based driver fatigue detection is recommended.

Article [15] addressed the hypo-vigilance prediction approach for Unnamed Combat Aerial Vehicle (UCAV) based on EEG signals to reduce the occurrence of hypo-vigilance. Drowsiness Detection and Warning System for Automobile Drivers Using Smartphones. This study describes a smartphone-based method for detecting sleepiness in car drivers. The suggested system detects sleepiness in three stages. The first step employs a modified eye state classification algorithm to calculate the percentage of eyelid closure (PERCLOS) from pictures acquired by the front camera. During nighttime driving, the technology illuminates the driver's face with near-infrared lights. In the case that PERCLOS exceeds the threshold, the second step uses the voiced to unvoiced ratio derived from the microphone's speech data.

A final verification stage uses a touch response within a predetermined time frame to declare the driver sleepy and raise an alert. The gadget keeps a log file of the metric's periodic occurrences, together with the accompanying GPS locations. The technology outperforms previous sleepiness detection methods in three ways. For starters, the three-stage verification procedure increases the system's reliability. The second benefit is that it is implemented on an Android smartphone, which is more easily available to most drivers or taxi owners than other general-purpose embedded platforms. The third benefit is the usage of SMS service to notify both the control room and the passengers.

III. Drowsiness Detection System

The general architecture of driver drowsiness detection is shown in Figure 4. Various EEG and ECG face detection techniques were employed in the Facial Detection phase to identify the face regions from the input photos. The research study [16] proposed an EEG-based approach that explored Hypo-vigilance real-time approach to predicate drowsiness, fatigued state, and alertness in biomedical applications. The ECG-based model [17] addressed non-invasive signals that analyze Heart Rate Variability (HRV) and applied filters to the ECG data to extract 13 statistical features. Extracted features trained by classification models SVM, KNN, and Ensemble to predicate driver fatigues state.

Face detection algorithms are divided into two categories: feature-based and image-based. Image-based algorithms for face detection have used statistical, neural network, and linear subspace methods. Different eye area detection techniques were employed in the second stage to detect and extract the eye region from facial photographs. After finding face regions, normalization is performed in the preprocessing stage to reduce the impacts of illumination.

Histogram equalization can be used to adjust the contrast discrepancies between face photographs. Feature extraction was applied to the input ocular region images in the third stage. The fourth phase in detecting driver's drowsiness is a classification, which employs a classifier to divide, using the characteristics obtained in the preceding two steps, divide pictures into asleep and non-sleeping groups.

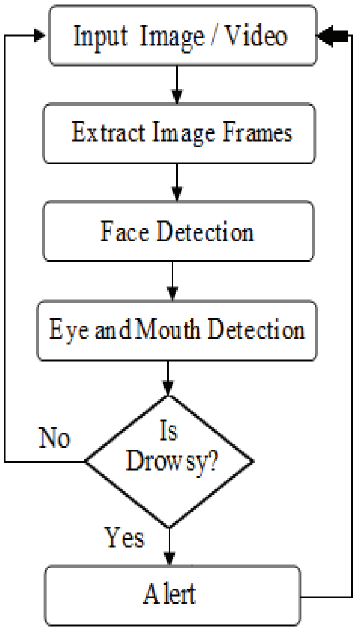

Figure 4. General Architecture of Drowsiness Detection.

IV. Dataset Selection

Many datasets have been used in various articles, although the majority of them are not public or realistic. As a result, each study used a distinct dataset to tackle the drowsiness detection problem. Unfortunately, the lack of a uniform dataset makes comparing different algorithms impossible. Several techniques, such as the ULG Multimodality Drowsiness (DROZY) database, achieve high validation accuracy of 91.6 percent and 95.8% by utilizing private datasets; others use public datasets, although they are unrealistic and lack sufficient training samples addressed in [18],[19].

V. CNN Model

Driver's drowsiness also cited sleepiness and fatigue which cause physical and mental tiredness [20]. The CNN model explores eye state categorization and drowsiness level detection in real-time video based on eye and mouth symptoms. Figure 5 depicts the system workflow and the drowsiness detection mechanism. The overview of the offline learning process is the initial stage in the system workflow. This stage summarizes the processes involved in the CNN training process for classifying open and closed eyes. Following the offline training, the online operation process can recognize the eye and mouth states in real-time video. As a result, symptoms from these states are used to predict the level of drowsiness.

From the input image, the convolutional layer will extract valuable characteristics. To increase the non-linear characteristics in each image, each convolutional layer is coupled to RELU activation. Then, it's not only to maintain the primary features but also to reduce the size of the photos, a max-pooling layer is applied. This aids the CNN model in reducing the quantity of irrelevant data while identifying the eye states.

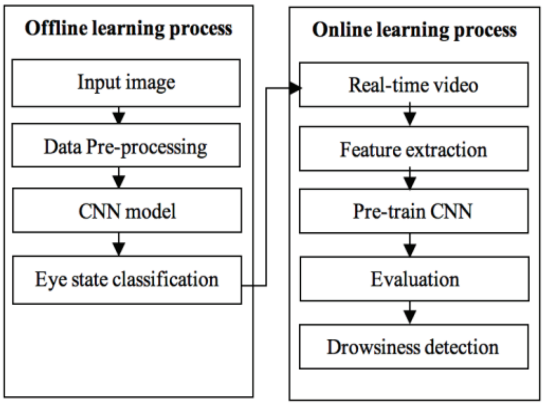

Figure 5. CNN Workflow.

VI. The CNN Based Drowsiness Detection System

Algorithm Steps:

(1) The Viola-jones face detection algorithm is utilized to recognize the faces in the photos, which is then passed on to the Viola-jones eye detection algorithm as input

(2) Once the face has been recognized, the Viola-jones iris recognition approach is utilized to extract the eye area from the face pictures, which is subsequently supplied as input to CNN.

(3) Deep features are extracted using a CNN with four convolutional layers, which are then sent to a fully connected layer.

(4) In CNN, the SoftMax layer classifies the pictures as sleepy or non-sleepy.

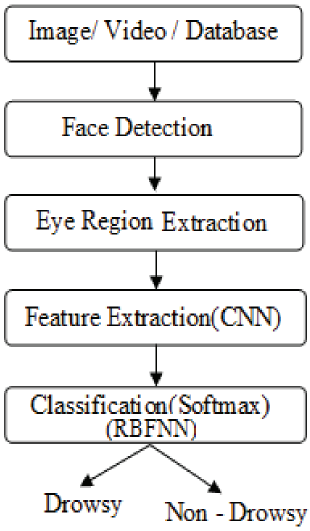

Figure 6. Drowsiness Architecture

VII. Discussion

Despite strong results, when a neural network is tested on a completely new person who was not included in the dataset, it struggles to detect drowsiness in a person's facial expression, implying that the person is awake the majority of the time. A useful gesture that is recognized by others is a small head droop. On the other hand, all films of captured subjects, including those produced afterward were effectively recognized.

VIII. Conclusions

This study indicates that in order to develop an application that could assist drivers in identifying, they were able to create a design that fits virtually flawlessly for varied frames despite their mental state of sleepiness. A network must be retrained for each new person in order to achieve acceptable outcomes for the subject. This article identified the limitations of CNN models and suggested an enhanced AI model for driver drowsiness detection. In the future, it must be necessary to enhance robust AI models to predicate and alert driver fatigues to reduce road accidents.

References

[1] Sahayadhas, A.; Sundaraj, K.; Murugappan, M.; Palaniappan, R. Physiological Signal based Detection of Driver Hypo vigilance using Higher-Order Spectra. Expert Syst. Appl. 2015, 42, 8669–8677. [CrossRef]

[2] Ghandour, A.; Hammoud, H.; Al-Hajj, S. Analyzing Factors Associated with Fatal Road Crashes: A Machine Learning Approach. Int. J. Environ. Res. Public Health 2020, 17, 4111. [CrossRef] [PubMed]

[3] J. Cigánek and J. Osuský, "Structure optimization of artificial neural networks using pruning methods," 2018 Cybernetics & Informatics (K&I), 2018, pp. 1-6, doi: 10.1109/CYBERI.2018.8337554.

[4] Krizhevsky Alex, Sutskever Ilya, and Geoffrey E. Hinton, "ImageNet Classification with Deep Convolutional Neural Network", Advances in Neural Information Processing System 25, 2012.

[5] Kepesiova, Z., Ciganek, J., Kozak, S.: Driver drowsiness detection using convolutional neural networks. In: 2020 Cybernetics & Informatics (K&I) (2020).

[6] *Gao Huang, Zhuang Liu, Laurens van der Maaten, Kilian Q. Weinberger*; Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 4700-4708

[7] Murugan, S.; Selvaraj, J.; Sahayadhas, A. Driver Hypovigilance Detection for Safe Driving using Infrared Camera. In Proceedings of the Fifth International Conference on Inventive Computation Technologies (ICICT), Tamilnadu, India, 26–28 February 2020; pp. 413–418.

[8] Chaari, L.; Golubnitschaja, O. Covid-19 pandemic by the "real-time" monitoring: The Tunisian case and lessons for global epidemics in the context of 3 PM strategies. EPMA J. 2020, 11, 133–138. [CrossRef]

[9] 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. [CrossRef]

[10] Houssaini, A.; Sabri, A.; Qjidaa, H.; Aarab, A. Real-Time Driver's Hypovigilance Detection using Facial Landmarks. In Proceedings of the International Conference on Wireless Technologies, Embedded and Intelligent Systems (WITS), Fez, Morocco, 30 May 2019; pp. 1–4.

[11] M. S. Devi and P. R. Bajaj, "Driver Fatigue Detection Based on Eye Tracking," 2008 First International Conference on Emerging Trends in Engineering and Technology, 2008, pp. 649-652, doi: 10.1109/ICETET.2008.17.

[12] Thomas, L.; Gast, C.; Grube, R.; Craig, K. Fatigue Detection in Commercial Flight Operations: Results Using Physiological Measures. Procedia Manuf. 2015, 3, 2357–2364. [CrossRef]

[13] Neri, D.; Shappell, S.; DeJohn, C. Simulated Sustained Flight Operations and Performance, Part 1: Effects of Fatigue. Mil. Psychol. 1992, 4, 137–155. [CrossRef]

[14] Hu, J.; Wang, P. Noise Robustness Analysis of Performance for EEG-Based Driver Fatigue Detection Using Different Entropy Feature Sets. Entropy 2017, 19, 385.

[15] Choi, Y.; Kwon, N.; Lee, S.; Shin, Y.; Ryo, C.; Park, J.; Shin, D. Hypovigilance Detection for UCAV Operators Based on a Hidden Markov Model. Comput. Math. Methods Med. 2014, 2014, 567–645. [CrossRef]

[16] Boudaya, A.; Bouaziz, B.; Chaabene, S.; Chaari, L.; Ammar, A.; Hökelmann, A. EEG-Based Hypo-vigilance Detection Using Convolutional Neural Network. In Proceedings of the International Conference on Smart Living and Public Health (ICOST), Hammamet, Tunisia, 24–26 June 2020; pp. 69–78.

[17] Stanley, P.; Prahash, T.; Lal, S.; Daniel, P. Embedded based drowsiness detection using EEG signals. In Proceedings of the IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI), Chennai, India, 21–22 September 2017; pp. 2596–2600.

[18] Murugan, S.; Selvaraj, J.; Sahayadhas, A. Detection and analysis: Driver state with electrocardiogram (ECG). Phys. Eng. Sci. Med. 2020, 43. [CrossRef]

[19] Zhang, L.; Liu, F.; Tang, J. Real-Time System for Driver Fatigue Detection by RGB-D Camera. Assoc. Comput. Mach. 2015, 6. [CrossRef]

[20] Dinges, D. An overview of drowsiness and accidents. J. Sleep Res. 1995, 4, 4–14. [CrossRef]