IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.39 ⁜ Vol. 11, Issue 5, May 2022

DOI: 10.17148/IJARCCE.2022.115157

Human - Drivers Drowsiness Detection System

Vishnu Dinesh[1], Arun Prakash[2], Amal Dasan[3], Poojitha Reddy[4], Mr. Mohammed Zabeeulla[5]

Department of Computer Science and Engineering, Jain University, Bangalore, India[1-5]

Abstract: Nowadays, more and more professions require long-term concentration. Drivers must keep a close eye on the road, so they can react to sudden events immediately. Driver fatigue often becomes a direct cause of many traffic accidents. Therefore, there is a need to develop the systems that will detect and notify a driver of her/him bad psychophysical condition, which could significantly reduce the number of fatigue-related car accidents. However, the development of such systems encounters many difficulties related to fast and proper recognition of a driver's fatigue symptoms. One of the technical possibilities to implement driver drowsiness detection systems is to use the vision-based approach. Here we are detecting the driver's drowsiness by estimating the vision system.

Keywords: car accident, drowsiness, recognition.

I. INTRODUCTION

Whenever we are driving vehicles, we may feel drowsy at some point It's not something we like to admit but it's an important problem with serious consequences that needs to be addressed.1 in 4 vehicle accidents are caused by drowsy driving and 1 in 25 adult drivers report that they have fallen asleep at the wheel in the past 30 days. The scariest part is that drowsy driving isn't just falling asleep while driving. Drowsy driving can be as small as a brief state of unconsciousness when the driver is not paying full attention to the road. Drowsy driving results in over 71,000 injuries, 1,500 deaths, and $12.5 billion in monetary losses per year. Due to the relevance of this problem, we believe it is important to develop a solution for drowsiness detection, especially in the early stages to prevent accidents.

Additionally, we believe that drowsiness can negatively impact people in working and classroom environments as well. Although sleep deprivation and college go hand in hand, drowsiness in the workplace especially while working with heavy machinery may result in serious injuries similar to those that occur while driving drowsily.

Our solution to this problem is to build a detection system that identifies key attributes of drowsiness and triggers an alert when someone is drowsy before it is too late.

II. PROBLEM DEFINITION

According to current technology, monitoring drivers while driving is quite complex computation and expensive equipment and it is also not comfortable to wear while driving. For example, EEG, ECG to check the frequency and rhythm of the heart.

As a new solution, a drowsiness detection system which uses a camera placed in front of the driver is more suitable to be used but the physical signs that will indicate drowsiness need to be located first in order to come up with a drowsiness detection algorithm that is reliable and accurate.

Lighting intensity and while the driver tilts their face left or right are the problems that occur during detection of eyes and mouth regions.

Detecting faces at night is also a problem faced by the model because of less light.

Wearing sunglasses while driving is another problem for detecting the drowsiness of a person while driving.

As our solution, we are planning to add a night vision camera for better detection at night, and we plan on adding a dataset which has classes for sunglasses and tilt face.

III. RELATED WORK

[1] In this paper, Varun Chaudhary, Ziyad Dalwai and Vikram Kulkarni worked on a visual approach where it will detect the drowsiness of the driver and alert the driver using an alert system. They used open source libraries like OpenCV to detect the faces and DLib to map out the coordinates of the face, in this work they used mouth and eyes coordinates. This method was tested on two genders and the accuracy was 97.5

- [2] In this paper, Feng You, Xiaolong Li, Yunbo Gong, Haiwei Wang and Hongli Li proposed a real-time driving drowsiness detection algorithm that considers the individual differences between drivers. The authors propose a new parameter, EAR, based on the Dib toolkit, to assess the state of the driver's eyes. They designed a deep cascaded convolutional neural network model named DCCNN, which avoids the process of artificial feature extraction in

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.39 ⁜ Vol. 11, Issue 5, May 2022

DOI: 10.17148/IJARCCE.2022.115157

traditional face detection algorithms, to obtain the face of a driver in live video. A unique classifier based on SVM is trained for a specific driver and the state of eyes is judged with the application of the pre-trained classifier during driving. Experimental results show that the accuracy of face detection can reach 98.8%.

- [3] In this paper, Satori Hachisuka first evaluated the effectiveness of the proposed DCCNN in Face Detection Data Set and Benchmark (FDDB). The authors then discuss the correlation of EAR and the size of eyes to describe the individual differences of drivers. The authors introduce two types of data sets to conduct experiments. The first one is FDDB, which contains the annotations for 5171 faces in a set of 2845 images. The proposed algorithm detects the drowsy state of driver quickly from 640∗480 resolution images at over 20fps and 94.80% accuracy. That is, when the proportion of eyelids covering the pupil is over 80 percent, the eyes are identified as closed in the current frame. It can be seen from the table that the algorithm the authors proposed improves the accuracy effectively. Experimental results show that the accuracy of face detection can reach 98.8%.

- [4] In this paper, Purvika Bajaj, Renesa Ray, Shivani Shedge, Sagar Jaikar and Pranav More , have trained their system using 3 different kinds of deep neural nets in order to know which of these would be a suitable model to be implemented in our system.They have used the pre-trained models namely VGG16, Inception and ResNet50 and had compared their performances.VGG16, Inception and ResNet50 is having accuracy about 92%, 83% and 50% respectively.

- [5] In this paper, Avigyan Sinha, Aneesh R P and Sarada K Gopal proposed a new methodology for approximation of the locations of facial key points with convolutional networks having three carefully designed levels . There are two benefits: first, the texture context data over the whole face is applied to find every key point. Second, for the reason that networks

- are skilled to predict all of the key points simultaneously, the geometric constraints amongst key points are implicitly encoded.Yolo is a deep learning algorithm for pattern matching. As it represents “You Only Look Once” (YOLO), this algorithm efficiently works in object detection as a one stage detector . In our work, Yolov3 has been also used for face detection as its speed is more than 1000 times that of R-CNN and 100 times that of Fast R-CNN. Accuracy of the system is 97% for 20 frames per seconds
- [6] In this paper, Charlotte Jacobe de Naurois, Christophe Bourdin, Anca Stratulat and Emmanuelle Diaz used different ANNs to detect a drowsiness level or to predict when a driver’s state will become impaired. The best models (those whose rates of successful detection or prediction are the highest) used information about eyelid closure, gaze and head movements and driving time. Performance on prediction is very promising, since the model can predict within 5 min when the driver’s state will become impaired. In this paper they achieved an accuracy of 95%.
- [7] In this paper, Rateb Jabbara, Khalifa Al-Khalifaa, Mohamed Kharbechea, Wael Alhajyaseena, Mohsen Jafaric and Shan Jiang used Deep Neural Networks. Using the Min-Max Scaler algorithm to change the range between 0 and 1, Defining the neural network model ,Adding layers and dropout to the model:,The first fully connected layer with rectifier20 function. The input layer has 67*2 ,136 nodes and the number of neurons used is 100.,A Dropout to prevent over-fitting21, dropout rate is set to 20%, The first hidden layer with rectifier function. The number of neurons used is 10. A Dropout to prevent over-fitting. The dropout rate is set to 20%.The second hidden layer
- with rectifier function. The number of neurons used is 10.A Dropout to prevent over-fitting. The dropout rate is set to 20%. The third hidden layer with rectifier function. The number of neurons used is 10.A Dropout to prevent over-fitting and the dropout rate is set to 20%. A softmax21 function to get output class label problems. In this paper they achieved accuracy of 81%.
- [8] In this paper first, G. Roshini, Y. Kavya, M. Suna and N. Sunny captured the video of the driver and extracted images from the video and extracted the eye portion separately using facial processing. This system will take the recorded video as input. Convolution neural network(CNN) is used to train this model. In CNN, each neuron in one layer is connected to all neurons in the next layer. Digital image processing is used to process digital images through various algorithms. To access the webcam, they made an endless cycle that captures every image. By using method CV2.Videocapture(0) from the OpenCV library to use the camera and get the user images

IV. OBJECTIVE

Suggests a way to detect fatigue and drowsiness while driving.
Investigates the physical changes of fatigue and drowsiness.
Develop a system that uses eye closure and yawning as a way to detect fatigue and drowsiness.
Provide alertness before any problem occurs.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.39 ⁜ Vol. 11, Issue 5, May 2022

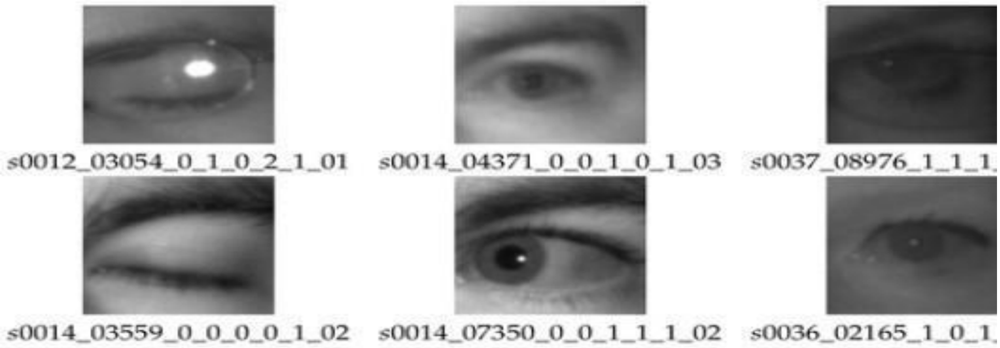
DOI: 10.17148/IJARCCE.2022.115157

Dataset

Mrl Eye - 2018 : The detection of eyes and their parts, gaze estimation, and eye-blinking frequency are important tasks in computer vision. In last years, we have been solving these tasks in the area of driver's behaviour, which causes the acquiring of a lot of testing data that was acquired in real conditions. Therefore, we introduce the MRL Eye Dataset, the large-scale dataset of human eye images. This dataset contains infrared images in low and high resolution, all captured in various lightning conditions and by different devices. The dataset is suitable for testing several features or trainable classifiers. In order to simplify the comparison of algorithms, the images are divided into several categories, which also makes them suitable for training and testing classifiers.

In the dataset, we annotated the following properties (the properties are indicated in the following order):subject ID; in the dataset, we collected the data of 37 different persons (33 men and 4 women)

- ❖ image ID; the dataset consists of 84,898 images
- ❖ gender [0 - man, 1 - woman]; the dataset contains the information about gender for each image (man, woman)
- ❖ glasses [0 - no, 1 - yes]; the information if the eye image contains glasses is also provided for each image (with and without the glasses)
- ❖ eye state [0 - closed, 1 - open]; this property contains the information about two eye states (open, close)
- ❖ reflections [0 - none, 1 - small, 2 - big]; we annotated three reflection states based on the size of reflections (none, small, and big reflections)
- ❖ lighting conditions [0 - bad, 1 - good]; each image has two states (bad, good) based on the amount of light during capturing the videos
- ❖ sensor ID [01 - RealSense, 02 - IDS, 03 - Aptina]; at this moment, the dataset contains the



CNN (Convolutional neural network)

A convolutional neural network consists of an input layer, hidden layers and an output layer. In any feed-forward neural network, any middle layers are called hidden because their inputs and outputs are masked by the activation function and final convolution. In a convolutional neural network, the hidden layers include layers that perform convolutions. Typically this includes a layer that performs a dot product of the convolution kernel with the layer's input matrix. This product is usually the Frobenius inner product, and its activation function is commonly ReLU. As the convolution kernel

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁘ Impact Factor 7.39 ⁘ Vol. 11, Issue 5, May 2022

DOI: 10.17148/IJARCCE.2022.115157

slides along the input matrix for the layer, the convolution operation generates a feature map, which in turn contributes to the input of the next layer. This is followed by other layers such as pooling layers, fully connected layers, and normalization layers.

Convolutional layer

In a CNN, the input is a tensor with a shape: (number of inputs) x (input height) x (input width) x (input channels). After passing through a convolutional layer, the image becomes abstracted to a feature map, also called an activation map, with shape: (number of inputs) x (feature map height) x (feature map width) x (feature map channels).

Convolutional layers convolve the input and pass its result to the next layer. This is similar to the response of a neuron in the visual cortex to a specific stimulus. Each convolutional neuron processes data only for its receptive field. Although fully connected feedforward neural networks can be used to learn features and classify images captured by different sensors

RealSense RS 300 sensor with 640 x 480 resolution, IDS Imaging sensor with 1280 x 1024 resolution, and Aptina sensor with 752 x 480 resolution)three (Intel data, this architecture is

generally impractical for larger inputs such as high resolution images. It would require a very high number of neurons, even in a shallow architecture, due to the large input size of images, where each pixel is a relevant input feature. For instance, a fully connected layer for a (small) image of size 100 x 100 has 10,000 weights for each neuron in the second layer. Instead, convolution reduces the number of free parameters, allowing the network to be deeper. For example, regardless of image size, using a 5 x 5 tiling region, each with the same shared weights, requires only 25 learnable parameters. Using regularized weights over fewer parameters avoids the vanishing gradients and exploding gradients problems seen during backpropagation in traditional neural networks. Furthermore, convolutional neural networks are ideal for data with a grid-like topology (such as images) as spatial relations between separate features are taken into account during convolution and/or pooling.

Pooling layers

Convolutional networks may include local and/or global pooling layers along with traditional convolutional layers. Pooling layers reduce the dimensions of data by combining the outputs of neuron clusters at one layer into a single neuron in the next layer. Local pooling combines small clusters, tiling sizes such as 2 x 2 are commonly used. Global pooling acts on all the neurons of the feature map.[18][19] There are two common types of pooling in popular use: max and average. Max pooling uses the maximum value of each local cluster of neurons in the feature map,[20][21] while average pooling takes the average value.

Fully connected layer

Fully connected layers connect every neuron in one layer to every neuron in another layer. It is the same as a traditional multi-layer perceptron neural network (MLP). The flattened matrix goes through a fully connected layer to classify the images.

Receptive field

In neural networks, each neuron receives input from some number of locations in the previous layer. In a convolutional layer, each neuron receives input from only a restricted area of the previous layer called the neuron's receptive field. Typically the area is a square (e.g. 5 by 5 neurons). Whereas, in a fully connected layer, the receptive field is the entire previous layer. Thus, in each convolutional layer, each neuron takes input from a larger area in the input than previous layers. This is due to applying the convolution over and over, which takes into account the value of a pixel, as well as its surrounding pixels. When using dilated layers, the number of pixels in the receptive field remains constant, but the field is more sparsely populated as its dimensions grow when combining the effect of several layers.

Weights

Each neuron in a neural network computes an output value by applying a specific function to the input values received from the receptive field in the previous layer. The function that is applied to the input values is determined by a vector of weights and a bias (typically real numbers). Learning consists of iteratively adjusting these biases and weights.

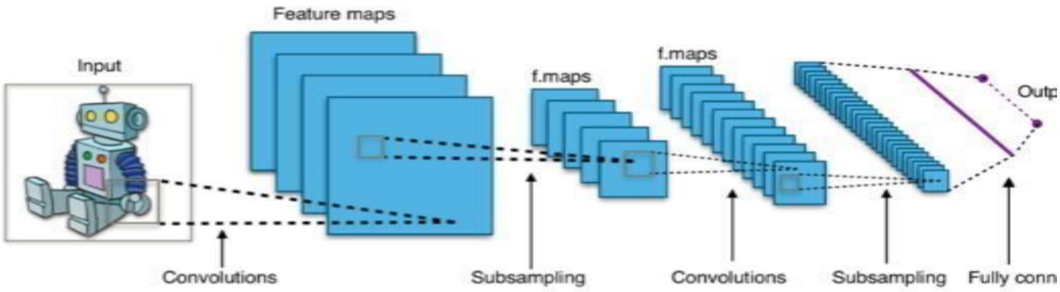
The vector of weights and the bias are called filters and represent particular features of the input (e.g., a particular shape). A distinguishing feature of CNNs is that many neurons can share the same filter. This reduces the memory footprint because a single bias and a single vector of weights are used across receptive fields that share that filter, as opposed to each receptive field having its own bias weighting.

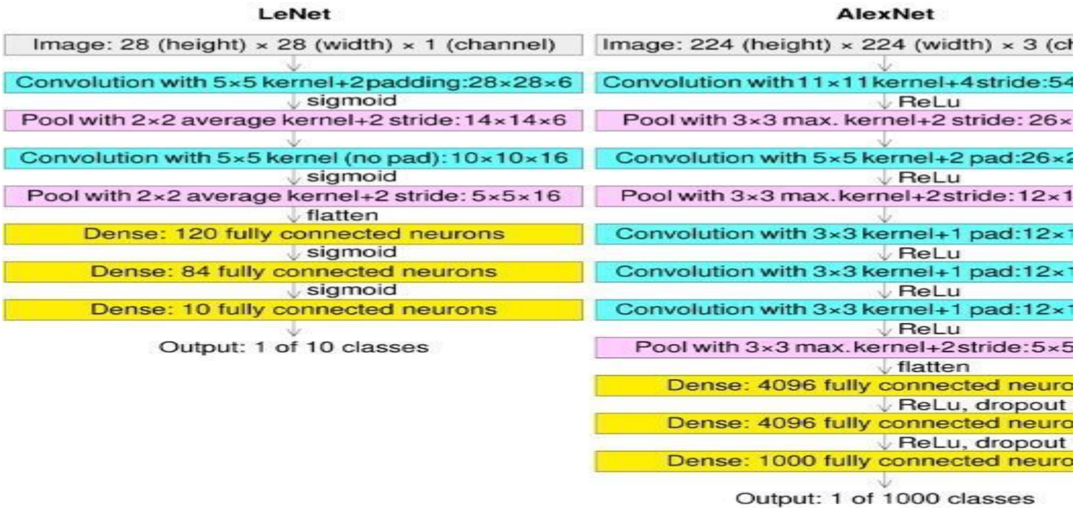
IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁘ Impact Factor 7.39 ⁘ Vol. 11, Issue 5, May 2022

DOI: 10.17148/IJARCCE.2022.115157





REFERENCES

[1]. [Feng You, Xiaolong Li, Yunbo Gong, Haiwei 2wang, Hongli Li, "A real - time drowsiness] detection algorithm with individual

[2]. consideration: a new algorithm to detect driving drowsiness which considers the individual difference of driving," IEEE, 2019. [Online].

[3]. [Satori Hachisuka, "Human and vehicle 3drowsiness detection by facial expression :] Using camera on dashboard for image

[4]. processing algorithm and a drowsiness detection algorithm," IEEE. ICBKE, 2018. [Online].

[5]. [Purvika Bajaj, Renesa Ray, Shivani Shedge, 4 Sagar Jaikar and Pranav More,] "Synchronous System for driver drowsiness

[6]. detection : drowsiness algorithm on android application also computer vision for cloud architecture in case of analysing real-time video streaming," IEEE. ICACCS, 2021. [Online].

[7]. [Avigyan Sinha, Aneesh R P and Sarada K 5Gopal, "Drowsiness detection system using] deep learning : give an alert alarm to the

[8]. driver to take necessary action and use deep learning techniques with CNN," IEEE. ICBSII, 2021. [Online].