ISSN 2277-2685
IJESR/Sep. 2023/ Vol-13/Issue-3/1-10
Mr.MD.Gulzar *et. al.*, / International Journal of Engineering & Science Research

Driver Drowsiness Monitoring using Convolutional Neural Networks

Mr.MD.Gulzar[1], Pateru Srinivas[2], Madhire Venkata Saritha[2], Gagguri Pavan Kumari[2] Modala Nagaraju[2] [1]*Asst. Professor,Computer Science and Engineering, CMR Engineering College, medchal, T.S, India* [2]*B.Tech, Computer Science and Engineering, CMR Engineering College, medchal, T.S, India*

Abstract: The development of computer vision has helped drivers by enabling autonomous self-driving vehicles and other technologies. Driver weariness and sleepiness account for 20% of accidents. It presents a significant issue for which numerous solutions were been out. They can't be used for real-time processing, however. These systems' ability to manage variations in human faces and lighting situations presents their biggest difficulties. Our goal is to put in place a system of intelligent processing that will significantly minimise traffic accidents. We can recognise the driver's facial features using this method, such as the proportion of closed eyes, eye-to-mouth ratios, blink rate, yawning, head movement, etc. In this setup, a camera is used to continually observe the driver. Haar cascade classifiers are used to find the driver's face and eye. For the purpose of identifying whether the left and right eyes are both closed, eye pictures are taken and supplied to a specially constructed KNN. The eye closure score is determined in accordance with the categorization. An alert will sound if the driver is discovered to be sleepy.

1. INTRODUCTION

Numerous safety-related driving assistance programmes reduced the risk of four-wheeler accidents, and studies showed that fatigue was a key contributing factor. An automobile organisation said that the majority of fatal incidents (17%) will be ascribed to fatigued drivers. Numerous updates published by Volkswagen AG state that driver inattention causes 5-25% of all incidents. A reliable, intelligent driver drowsiness monitoring system is required because lack of attention impairs steering movements and slows reaction times. drowsiness also increases the risk of collisions. To prevent traffic accidents, an intelligent processing system is being developed. This may be achieved by periodically checking for sleepiness and alerting drivers to distractions in order to avoid accidents. According to a research review, there are three ways to identify tiredness in a driver: physiological, behavioural, and vehicle-based data. But in certain real-world situations,[1] [2] these techniques have drawbacks.

A Description of The Project: Calculate the length of the intensity change in the eye area as a basic method [3] [4] of eye detection. Other physiological aspects like blink rate, yawning, head movement, etc. are not taken into consideration. Real-time processing is not a good fit for it. The choice was chosen to frame the use of two-dimensional CNN [5] for distinctive activity. It disregards the link between space and time. It makes use of simulated datasets and ignores aspects like head movement. A useful set of learned features to classify driver drowsiness was supplied by the characteristic state learning conceptualizations [6]. In order to extract features, it employs a complicated framework. The hand-operated tag was applied to the chosen photo series and video data collection.

2. LITERATURE SURVEY

2.1 Existing System

1) Deep Neural Network for Recognising Human Faces

AUTHORS: Nidhi Saxena and Priya Gupta

In the fields of biometrics, information security, access control, law enforcement, smart cards, and surveillance systems, face recognition (FR), the technique of identifying persons using facial photographs, has a wide range of useful applications. It has been shown that Convolutional Neural Networks (CovNets), a subset of deep networks, are effective for FR. Before employing CovNets for real-time systems, various preprocessing processes like sampling must be completed. The network then completes all of the processes (feature selection, feature extraction, and training) after receiving complete pictures (all of the pixel values) as input. This is the cause of the sometimes difficult and drawn-out implementation of CovNets. CovNets are still in their infancy, and despite the excellent accuracy levels, they still have a long way to go. The research suggests a novel method for face recognition using a deep neural network (another kind of deep network). In this method, just the extracted face characteristics are supplied as input, rather than the raw pixel data.[7] [8] With accuracy of 97.05% on the Yale faces dataset, this reduces complexity.

Behaviour-Based Data Dispatcher

Mohan Sai Singamsetti and Mona Teja Kurakula are the authors.

Human existence is a sophisticated social system. Humans are unable to travel without being able to read other people. They recognise the faces to do this. Based on how the other person is feeling, one may select how to respond. While a person's mood may be determined simply watching his facial gestures (emotion). The project's goal is to develop a "Facial emotion Recognition" model in real time using a DCNN (Deep Convolutional Neural Network). Since it has been shown that DCNN perform more accurately than CNN (convolutional neural network), DCNN was used to build the model. Humans' facial expressions are incredibly dynamic and may vary in a matter of seconds, expressing everything from happiness to sadness to anger to fear to surprise to disgust and neutrality. Real-time emotion prediction is the goal of this study. Neural networks in our brains are in charge of all types of thinking (decision-making,comprehension).[9]This methodology makes an effort to educate the computer to acquire these categorization and decision-making abilities. At the same time, it can simultaneously categorise and forecast the many faces and many emotions. We use models that have been refined across tens of thousands of datasets in order to achieve improved accuracy.

3) Eye tracking-based driver fatigue detection

Mandalapu Sarada Devi, author

According to international data, driver weariness contributes to a significant proportion of traffic accidents. Thus, a system that can identify approaching driver weariness and send out early warnings might aid in averting numerous accidents, which would assist to save money and lessen suffering for individual victims. In an effort to create a system that can detect driver drowsiness, the authors have placed a camera right in front of the driver's face. If driver weariness is found, a warning signal is sent to let them know. The video files that the camera captured have been edited by the writers. Frames are extracted from a video file. Once the eyes have been identified from each frame, the open or closed status of the eyes may be determined by calculating the distances between intensity changes in the eye region. The technology determines that the driver is dozing off if the eyes are closed for five consecutive frames and sends out a warning signal. The algorithm has been suggested, put into practise, tested, and confirmed to function adequately.

4) A Deep Network-Based System for Feature Representation Learning-Based Driver Drowsiness Detection

Sanghyuk Park, Fei Pan, Sunghun Kang, and Chang D. Yoo are the authors.

According to statistics, weariness is a factor in 20% of all traffic accidents. Drowsy detection is a vehicle safety algorithm that may wake up a motorist who is nodding off in the hopes of averting a collision. In order to acquire useful characteristics and identify tiredness given an RGB input video of a driver, the deep drowsiness detection (DDD) network is proposed in this study. The DDD network is made up of three deep networks that work together to provide global resilience against background and environmental fluctuations and to learn regional head and facial motions that are crucial for accurate recognition. To identify sleepiness, the three networks' outputs are combined and sent to a softmax classifier. On the benchmark dataset for NTHU-drowsy driver detection, experimental findings demonstrate that DDD achieves a detection accuracy of 73.06%.

2.2Proposed System

Our suggested approach will provide a means of keeping track of driver sleepiness. By providing an input driver picture to a specially tailored CNN, the prior system's drawback of extracting just a few hand-crafted features is solved. A camera will now be used to continually observe the driver. A series of frames are created out of the video that was recorded. Using prepackaged classifiers termed haar cascade classifiers, which are accessible in opencv, the face and eye are identified for each frame. Eye pictures are retrieved, transmitted via a sequence of 2D CNN layers (5x5, 3x3 kernel valid padding), max-pooling layers (2x2), and lastly, the fully connected dense layer, which determines whether or not the eyes are closed. Eye closure is used to determine a score. The technology detects drowsiness if both eyes are closed continuously for 15 frames, at which point an alarm sound is played to warn the driver. Using specially created CNN, driver drowsiness is properly classified, and the normalisation problems in the old model are removed[10].

2.2 Description of Implementation

The modules include Preprocessing and labelling, Data Augmentation, Enhanced CNN, Face and eye recognition, and Triggering Alarm. The Face and eye identification module of OpenCV uses a load technique to supply pre-trained models.

The OpenCV information folder contains the pre-trained forms. The system will make use of previously practised Haar models to examine the ocular image. The crucial XML file is loaded using the Cascade Classifier's technique. The multiscale approach then generates a bound rectangle for the observed eyes. We preprocessed and labelled webcam photos of eyes as having open or closed eyes to create our dataset. The retrieved ocular pictures were normalised, reduced in size to 24x24 pixels, and made into grayscale images. For the Data Augmentation module, we transformed the existing data rather than gathering new data. Using Keras, which takes the ranges for rotation, brightness, shear, zoom, etc. as parameters, we supplement the data. As demonstrated in Figure 2, each CNN layer contains several arguments that may be adjusted and carry out different operations on the input data. With a 3x3 kernel size, we applied the relu activation function in convolutional layers. Fully linked layers use the Softmax activation algorithm to provide the results of either open or closed eyes. Adam optimizer is used to train CNN.

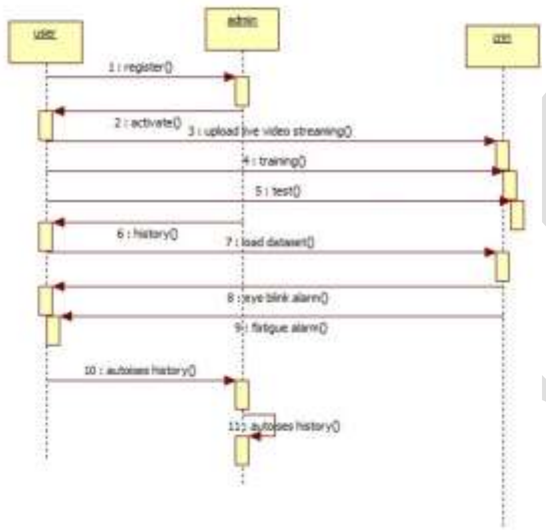

Fig 1: System Flow

3. SYSTEM ANALYSIS&DESIGN

3.1 Face detection and feature point location make up section.

Face identification is difficult in real-world situations because of shifting driving postures and uncontrolled environmental elements like lighting and occlusion. The global face features, including the positions of the left and right eyes, nose, and corners of the mouth, can be extracted by using the depthcascading multitasking MTCNN framework. Face detection and alignment can be done simultaneously, and the internal relationship between the two is exploited to improve performance. Figure 2 depicts the MTCNN's structural layout. The three cascaded subnetworks that make up the MTCNN—P-Net (proposal network), R-Net (refined network), and O-Net (output network)—are used to recognise faces and feature points at various stages of refinement.

P-Net: To acquire photos of various sizes, an image pyramid is first built. These photos are then sequentially entered into the P-Net. To identify whether a face is a part of a 12 12 area at each place, a fully convolutional network is used. This yields a bounding box for the candidate face area and its regression vector.

The candidate face window is calibrated using the frame regression vector in 'en, and nonpolar large value suppression is used to get rid of the candidate face areas that are significantly overlapping [28, 29].

R-Net: the picture size is modified to 24 x 24 and the candidate face region from the P-Net input is used. By using bounding box regression and nonmaximum value suppression, the candidate face window is filtered. The network structure adds a connection layer to the P-Net in order to achieve a more precise face location.

O-Net: Like the R-Net, the O-Net scales the picture to 48 by 48 and screens the candidate face window to determine the final face position and five feature points.

The three-layer cascade network used by the MTCNN to recognise faces concurrently conducts face categorization, bounding box regression, and feature point localization. It has excellent toughness and is appropriate for actual driving situations. Figure 3 displays the outcome.

3.2. State of the Eye and Mouth Recognition

ROI Extraction: Typically, to detect a condition of weariness, most eye detection techniques simply extract one eye.However, utilising information from only one eye may easily result in error when the driver's head moves.Therefore, the suggested technique extracts a two-eye picture to ascertain whether the eyes are open or closed in order to gather additional eye information and precisely recognise the eye state.

Using the MTCNN network, the location of the driver's left and right eyes is determined. Here, the left eye is in position a1 (x1, y1), the right eye is in position a2 (x2, y2), the space between the left and right eyes is measured by the letter d1, and the width of the eye image is measured by the letter w1. 'e height is h1, according to the proportion of the face "three courts and five eyes," the binocular pictures are intercepted, and the correlation between width and height is stated as follows:

When speaking and yawning, a driver's mouth area changes substantially. The MTCNN network is used in this instance to determine the locations of the left and right corners of the mouth. The left corner of the mouth is located at location b1 (x3, y3), while the right corner is located at position b2 (x4, y4). Similar to the eye region extraction, the mouth image's width is w2, its height is h2, and the distance between its left and right corners is d2.

3.3 System Architecture: The system's intended flow is shown in Figure 2. A webcam is used to watch the driver in the first stage. A series of frames is created from the video input. Haar cascade classifiers are used in each frame to identify the driver's face and eyes. Images of the identified eyeballs are saved to create a data collection for CNN. Additionally, the system has provisions for the creation of the ocular data set for CNN model training. to improve the image's training and the quantity of data sets It is done to supplement the data. Then, a variety of image preparation operations, including grayscale conversion, resizing and normalising, etc., are applied to the pictures of both eyes. In order to forecast eye closure, it is then input into a pre-trained CNN model composed of convolution layers, max-pooling layers, and dense layers. The prediction is used to determine a score. An alarm will sound to warn the motorist if the system detects that they are asleep.

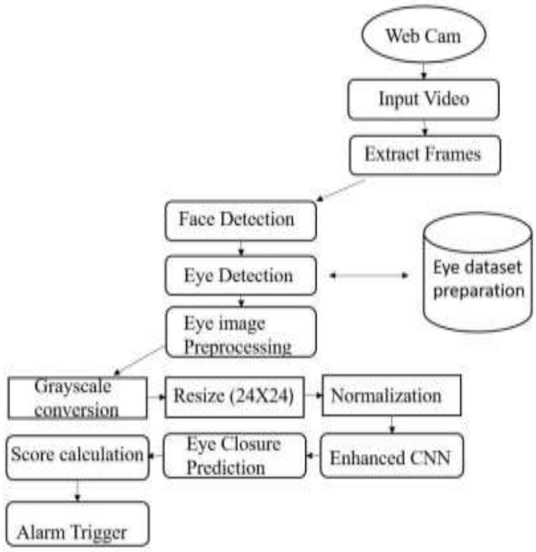

Fig 2 : System Architecture

3.4 Data Flow Diagram : Whenever a new system is developed, user training is required to educate them about the working of the system so that it can be put to efficient use by those for whom the system has been primarily designed. For this purpose the normal working of the project was demonstrated to the prospective users. Its working is easily understandable and since the expected users are people who have good knowledge of computers, the use of this system is very easy.

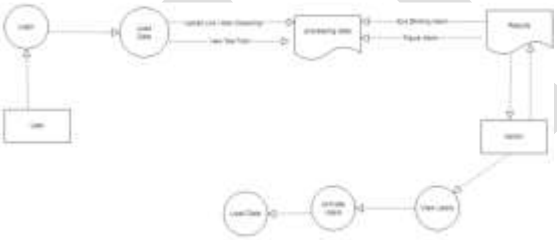

Fig 3: Data Flow Diagram

Fig 4 Use Case UML Diagrams

4. IMPLIMENTATION AND RESUTS

5. CONCLUSION

Effective KNN architecture is required for a model for drowsiness sensing, which is intended to monitor tiredness based on eye closure. For both open and closed eyes, the implementation first began preparing picture collections. The custom-designed KNN training uses 75% of the data set, and the remaining A quarter of the dataset is used for testing. The face and eyes are first recognised in each of the frames created from the information video. The improved KNN provided an automatic and efficient learning feature that helps us classify when eyes open or close. An alarm is set off to notify the driver if the eyes close 15 times in a row. A training accuracy of 97% and a testing accuracy of 67% are provided by the suggested KNN. Additional facial attributes may be introduced in further works to improve detection accuracy. Additionally, we can integrate face feature extraction with data on driving patterns from on-board diagnostic sensors.

6. FUTURE SCOPE

A training accuracy of 97% and a testing accuracy of 67% are provided by the suggested KNN. Additional facial attributes may be introduced in further works to improve detection accuracy. Additionally, we can integrate face feature extraction with data on driving patterns from on-board diagnostic sensors.

7. REFERENCES

[1] Dr. Priya Gupta, Nidhi Saxena, Meetika Sharma, Jagriti Tripathi,Deep Neural Network for Human Face Recognition InternationalJournal of Engineering and Manufacturing, vol.8, no.1, pp. 63-71.
January 2018.
[2] Jeyasekar A, Vivek Ravi Iyengar,Based on Behavioural Changes using ResNet , International Journal ofRecent Technology and Engineering (IJRTE), vol. 8, no. 3, pp. 2 5-30,2019.

[3]Conference (IACC), Gurgaon, pp. 995-999, 2014.

[4] Ki Wan Kim, Hyung Gil Hong, Gi Pyo Nam and Kang Ryoung Park,

.[5] Luigi Celona, Lorenzo Mammana, Simone Bianco, RaimondoSchettini,-Task CNN Framework for Driver Face Berlin, 2018.

[6] Mandalapu Sarada Devi and Dr. Preeti R Bajaj,Detection Based on Eye Tracking , First International Conference onEmerging Trends in Engineering and Technology, vol.1, pp. 649-652,2008.

[7] Sanghyuk Park, Fei Pan, Sunghun Kang and Chang D. Yoo, Driverdrowsiness detection system based on feature representation learningSpringer International Publishing, Computer Vision ACCV 2016 Workshops.

[8] Tawsin Uddin Ahmed, Sazzad Hossain, Mohammed ShahadatHossain,RaihanUl Islam ,Karl Andersson, Facial Expression

Recognition using Convolutional Neural Network with Data Th International Conference on Informatics, Electronics & Vision (ICIEV), Washington, USA,2019,

[9] Upasana Sinha, Kamal K. Mehta, AK Shrivastava,Real Time Implementation for Monitoring Drowsiness Condition of a Train Driverusing Brain Wave Sensor , International Journal of ComputerApplications, vol. 139, no.9, pp. 25-30, 2016.

[10]. Ch Prathyusha, Mrutyunjaya S Yalawar, An approach for detecting drowsy drivers in vehicle using CNN techniques, Journal International journal of recent technology and engineering,Volume 11,Issue 3, Pages-41-44, 2022, 10.35940/ijrte.C7246.0911322