Vidhyayana - ISSN 2454-8596
An International Multidisciplinary Peer-Reviewed E-Journal
www.vidhyayanaejournal.org
Indexed in: Crossref, ROAD & Google Scholar

14

Driver Drowsiness Detection Using Inceptionv3 with Automatic Whatsapp Message Sender

Manas Ohara

Student, School of Computer Science, MIT WPU.

manasoharak@gmail.com

9511646396

Chaitali Gadekar

Student, School of Computer Science, MIT WPU.

chaitaligadekar50@gmail.com

9765293589

Abstract

Human Driver drowsiness is one of the main reasons for road accidents in the world. To prevent such accidents, a driver drowsiness detection system is proposed in this research paper. InceptionV3, a deep learning architecture, is used to classify the driver's facial expressions and detect drowsiness. The system is integrated with a real time frame capturing camera, which captures the driver's face, and the model processes the images in real-time to identify drowsiness of the human driver. Once the system detects that the driver is drowsy, an automatic WhatsApp message is sent to a predefined contact to alert them of the situation. This proposed system yields higher accuracy in drowsiness detection, and the automatic WhatsApp message sending feature can provide timely assistance to prevent potential accidents.

Index Terms- Driver drowsiness, InceptionV3, Deep learning, Facial expression, Real-time processing, Automatic message sending on WhatsApp.

I. INTRODUCTION

Driver drowsiness is a major contributor to road accidents in India, which lead to numerous fatalities and injuries each year. Common reasons for driver fatigue include long hours of driving and inadequate rest. To prevent such accidents, it is crucial to develop effective measures to detect and alert drivers in case of drowsiness. Recent advancements in deep learning techniques and computer vision has made major development of driver drowsiness detection systems that can monitor a driver's behaviour and mitigate drowsiness-related accidents.

This study proposes a driver drowsiness detection system that employs the InceptionV3 deep learning architecture to detect drowsiness based on facial expressions. In real-time, the system captures the driver's face using a camera, which is processed using the InceptionV3 model. The system can accurately detect drowsiness and is equipped with an automatic message sender that utilizes WhatsApp to notify a designated contact about the driver's drowsy state.

This proposed system aims to enhance road safety in India by reducing accidents caused by driver drowsiness. The integration of automatic message sending using WhatsApp can provide timely assistance to the driver, thus preventing potential accidents.

II. IDENTIFY, RESEARCH AND COLLECT IDEA

The literature survey/review conducted here covers multiple research papers and existing state-of-the-art systems proposed or implemented for Driver distraction recognition systems. This survey aims to study the existing methods and evaluate and propose further investigation and in potential areas.

Human Driver Distraction/Drowsiness Detection with a Camera Vision System [1,5]: this system implements System Interface to monitor driving focus and recognize the driver's distraction from driving based on the visual inputs provided. The system also monitors the lane and provides a lane tracking module which helps to track the lane of the vehicle, primarily cars and trucks. Together, the system tracks both lanes and the driver's attention using the visual inputs from the sensory data. The system operates the rule and support vector machine (SVM) distribution method, combining vision and lane data. This detects the visual and cognitive functions of the driver's eyes. The outcomes showed a positive rate of more than 80% in visual noise detection and 68-86% in emotional noise detection. But this approach of coupling lane tracking and driver distraction is limited to the driver's attention and lane-based roads. This system can have shortcomings for many out of the scope and catastrophic real-world situations. Therefore, a more reliable yet simpler axioms or parameters are required to improve the accuracy.

Lane Deviation based approach: This approach does not involve any facial tracking but a simple pattern finding algorithm which monitors lane deviation and tracks the pattern to recognize if the driver is going off the lane, assuming the driver is drowsing.

Another method is brain activity tracking and inferring the drowsiness and sleep through brain wave signals acquired through electronic sensors, such as Electrocardiogram, and Electroencephalogram data. Also, Electrooculography plays a key role. The received signals are divided into three primary states that are alpha, delta and theta. The accuracy for this method is around 90% but the issue with this approach is feasibility. To accomplish this, the peripheral sensors for collecting signals should be stuck to the human driver all the time. But this can be practically uncomfortable and not feasible for the long term.

Another recent approach which has gained popularity is computer vision, thanks to the widely available datasets and Machine Learning Algorithms with good accuracy to find patterns. There are numerous ML algorithms for computer vision like SVM, CNN etc. The only catch for this approach where all the accuracy gets deviated is the axioms / parameters taken as inputs and the algorithm. The accuracy and feasibility to solve/improve for this particular problem needs a fine-tuned algorithm for this specific use case. This is what our proposed system is about. We are using InceptionV3 which is fine tuned for use cases like these and will yield best accuracy.

III. WRITE DOWN YOUR STUDIES AND FINDINGS

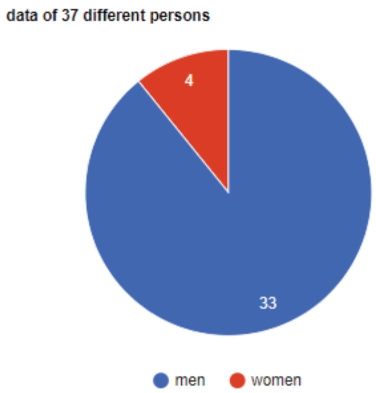
Data Acquisition: The dataset for this concerned project will be acquired from MRL Eye Dataset, a big dataset of human eye images from different angles, lighting conditions,eye types, infrared images in low and high resolution. For optimization purposes, the comparison of algorithms, the images are bifurcated into several categories, which is better for training

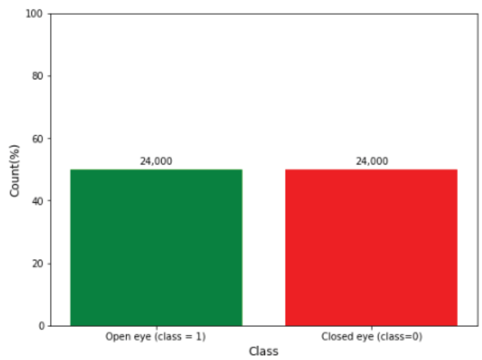
Vidhyayana - ISSN 2454-8596
An International Multidisciplinary Peer-Reviewed E-Journal
www.vidhyayanaejournal.org
Indexed in: Crossref, ROAD & Google Scholar

and testing classifiers [2].

The dataset is comprised of a plethora of annotations for various properties, including subject ID, image ID, gender (male or female), glasses (with or without), eye state (open or closed), reflections (none, small, or big), lighting conditions (good or bad), and sensor ID. This valuable data is related to 37 individuals, with 33 of them being male and 4 female. In addition, this remarkable dataset consists of a whopping 84,898 images, captured using cutting-edge sensors, such as Intel RealSense RS 300, IDS Imaging, Aptina.

Statistics of Dataset: In dataset, data of 37 different persons (33 men and 4 women) has been collected.



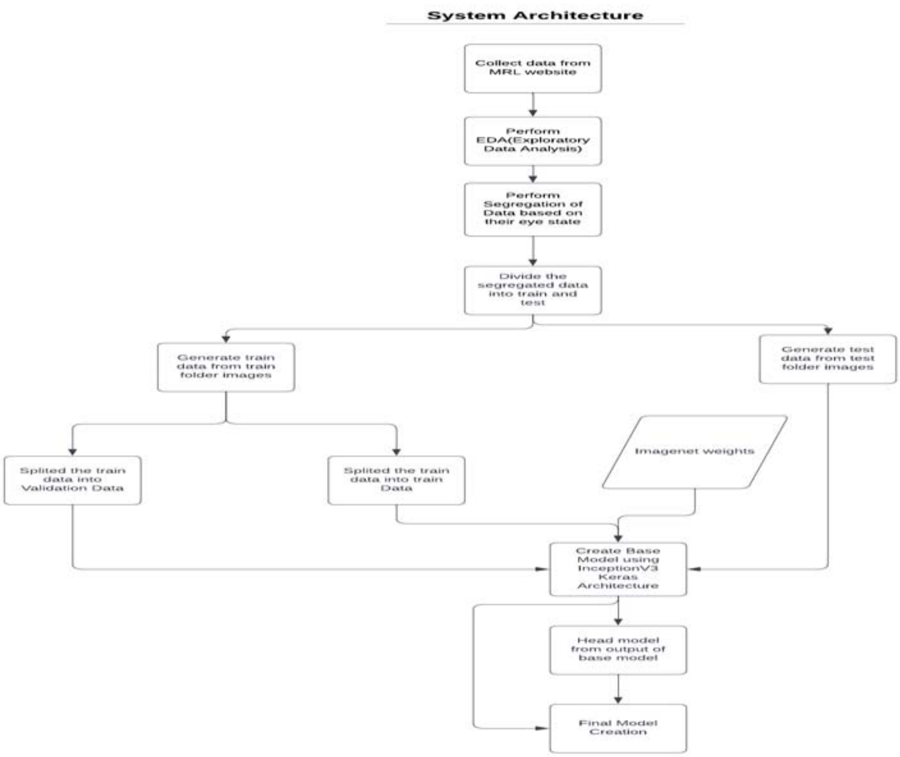


Below graph shows the number of closed and open eyes in the dataset. The classes in the dataset are balanced. Each class is represented by 24,000 images.

The dataset contains:

1. Train:
   a. close eyes (40.4k images)
   b. open eyes (41.3k images)
2. Test:
   a. close eyes (1566 images)
   b. open eyes (1657 images)

Vidhyayana - ISSN 2454-8596
An International Multidisciplinary Peer-Reviewed E-Journal
www.vidhyayanaejournal.org
Indexed in: Crossref, ROAD & Google Scholar



Proposed Methodology: The MRL Eye Dataset features images captured using different devices and under varying lighting conditions, making it an ideal tool for testing multiple features and trainable classifiers. The images have been sorted into separate categories to simplify algorithm comparisons, and the dataset comprises eye images from 37 individuals (33 men and 4 women). Each person's eye images are conveniently stored in individual folders, with the folder names reflecting the subject names [3]. The images are labeled using a standardized format below mentioned:

<subject_id><image_id><gender><glass_state><eye_state><reflection_status><lightning condition>_<sensor_type>.

Important attributes which have been considered during data preparation are eye_state (i.e whether eyes are closed or open) and lightning condition which will make the model robust as the model will work on any lightning conditions. The dataset will be segregated into open eyes and closed eyes based upon the file name with the help of *shutil* library. Later this data is manually bifurcated into training data and testing data dataset with 90% of the data considered as training data and the rest of 10% is considered as testing data. In the process

Model Building, every image from train data is transformed into 6 images which are rescaled 1./255 ratio of pixel as in dataset due to variation sensor types the image resolution varies. And that images are stored as trained data generators in the model building process. The data which has been stored is stored on the basis of categorical class mode, as we are looking for eye open or close state only. Further that generated data is splitted into train data and validation data. Validation data is used to validate whether trained data is correct or not. With the help of the same process, the test data is generated for test folder images. Using 'InceptionV3 - a keras application 'base model is created. While building the base model, weights for the neural network are taken from 'ImageNet is an image database that is structured based on the WordNet hierarchy, where each node in the hierarchy is represented by a large collection of images. Currently, only the noun categories are included in the hierarchy. This database contains hundreds and thousands of images for each node, making it a valuable resource for training and evaluating computer vision models. From the output of the base model, the head model is created and the neural network is densed, flattened, dropped out accordingly with the help of softmax and *ReLU* activation function. Once the model is created, callbacks are generated through keras library and the model is compiled with back propagation optimizer 'Adam' [5]. Now the loss values are computed and the accuracy is evaluated to identify any issues with the training data set. This evaluation can be used to also assess the model's performance and address the issues in the training process itself. Then we evaluate the trained model's performance on both validation and test datasets using the Keras API. Next, we evaluate the percentage it has learned to predict both validation and test datasets. This helps us to get the accuracy of the trained model in percentage and we can go for further fine tuning of the model. Once the model training is done, now we test the model with real time data. For this first need to define face and eyes detection objects in the images or videos and classify them as two cascade classifiers. One will be 'face_cascade' and another will be 'eye_cascade'. By using OpenCV library, we capture the driver's video and convert it into grayscale frames. This makes it easier to recognise the object in the frame regardless of the lighting conditions. Now, the cascade classifiers will help to recognise the face and then based on that recognise eyes as well and overlay the frames with rectangles. Then the eye's object data is extracted and sent to the

Deep Learning Model as an input. The Model then finds patterns and predicts if the human driver's eyes are open or closed. If the eyes are closed for more than 5 frames, then the code will play an alarm and send notification to WhatsApp. This will help to alert the driver immediately. A detailed elaboration of the same can be seen in the flowchart diagram below.

Machine Learning Models and Techniques -

1) Keras: It is a user-friendly deep learning library that simplifies the development and deployment of neural networks. With pre-built customizable layers and models, it supports various optimization algorithms, loss functions, and metrics. Keras integrates with TensorFlow, CNTK, and Theano, making it a preferred choice for deep learning practitioners.
2) Activation functions:
    a) The ReLU activation function is a simple yet effective function used in deep learning neural networks. It returns the input value for positive inputs and 0 for negative inputs, making it a computationally efficient and easy to optimize choice.
    b) Softmax is a popular activation function used in deep learning for multi-class classification tasks. By transforming a vector of real numbers into a probability distribution, Softmax produces outputs that can be interpreted as the likelihood of each class. Due to its versatility and ability to handle complex decision boundaries, Softmax is widely used in neural network architectures.
3) Transfer Learning: It is a technique in machine learning that utilizes a pre-trained model's learned features to solve a new task. By using the pre-trained model's weights and biases, it reduces the amount of training data required for good performance on the new task [4].
4) Back-propagation: It is a type of supervised learning algorithm that optimizes the weights of the neural network using gradient descent. Backpropagation works by computing the gradient of the loss function with respect to each weight in the network, then propagating the error backwards from the output layer to the input layer.
5) InceptionV3: Inception v3 is a deep CNN that uses convolutional and pooling layers to extract complex features from images. It is pre-trained on the large-scale ImageNet

dataset, making it ideal for fine-tuning for various image classification tasks. Inception v3 is widely used and supported by deep learning frameworks like TensorFlow and Keras in both academia and industry.

IV. APPLICATIONS AND FUTURE SCOPE

The drowsiness detection system is not limited to cars and can be installed in various transportation vehicles including motorbikes, trucks, and more. Currently, the model is trained on an existing dataset of eyes from MRL, but the system can be enhanced by incorporating various sensors to collect driver data in real-time. Moreover, the model can be improved by including yawn detection as a feature. When the system detects drowsiness, it can be programmed to automatically send a WhatsApp message with the driver's current location. Additionally, the system can be further enhanced by sending automatic location coordinates to the nearest Traffic Control Centre or Toll Plaza on highways. This will alert the authorities in real-time and allow them to take immediate action to prevent accidents.

V. CONCLUSION

Driver drowsiness is an important factor in traffic crashes, and deep learning algorithms show promise for developing effective driver drowsiness detection systems. InceptionV3 architecture and algorithms based on CNN and RNN have been proposed for driver drowsiness detection to achieve high accuracy and low false alarm rate. However, further study and research is beneficial to assess actual effectiveness of these systems for preventing accidents.

VI. REFERENCE

[1] https://www.academia.edu/38928274/REAL_TIME_SLEEP_DROWSINESS_DETECTION_Project_Report

[2] https://www.researchgate.net/publication/336878674_DRIVER_DROWSINESS_DETECTION_SYSTEM

[3] A. M. Malla, P. R. Davidson, P. J. Bones, R. Green and R. D. Jones," Automated video-based measurement of eye closure for detecting behavioral microsleep," 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology, Buenos

Aires, 2010, pp.6741-6744. doi: 10.1109/IEMBS.2010.5626013

[4] M. I., B. Sharada and P. Nagabhushan, "Graph based features for recognition of handwritten Devanagiri numerals," *2016 International Conference on Communication and Signal Processing (ICCSP)*, Melmaruvathur, India, 2016, pp. 1710-1715, doi: 10.1109/ICCSP.2016.7754458.

[5] M. I. Bhat, B. Sharada, S. M. Obaidullah and M. Imran, "Towards Accurate Identification and Removal of Shirorekha from Off-line Handwritten Devanagari word Documents," *2020 17th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, Dortmund, Germany, 2020, pp. 234-239, doi: 10.1109/ICFHR2020.2020.00051.