INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact factor: 4.295
(Volume 4, Issue 3)
Available online at: www.ijariit.com

Machine learning based classifier model for autonomous distracted driver detection and prevention

Ashu Yadav
ashuyad1234@gmail.com
Army Institute of Technology, Pune, Maharashtra

Naveen Sharma
itsnaveen01@hotmail.com
Army Institute of Technology, Pune, Maharashtra

Yadvendra Yadav
yadvendra.yadav1995@gmail.com
Army Institute of Technology, Pune, Maharashtra

Jasaram Choudhary
jrchoudhary001@gmail.com
Army Institute of Technology, Pune, Maharashtra

Praveen Hore
praveen.hore@gmail.com
Army Institute of Technology, Pune, Maharashtra

ABSTRACT

Recent researches and surveys have provided us with the evidence that distracted driver is a major cause of vehicle crashes all around the world. In-vehicle information systems (IVIS) have raised driver safety concern and thus, detecting distracted driver is of paramount importance. The project (or paper) shows a method of real-time distraction detection and initiates safety measures. In the realization of this project we have used Web-Cam, Raspberry Pi (a low cost, small size computing device), along with concepts of deep learning and convolutional neural networks. We classify drivers into multiple categories of distraction, some of them are texting, drinking, operating IVIS etc. Web-Cam feeds the classifier with real-time images of the driver of a particular vehicle. The system also constitutes a buzzer alarm which rings once the distraction is detected.

Keywords: *Machine Learning, Convolutional Neural Network, Classification, Hyper parameters.*

1. INTRODUCTION

Distracted driving is characterized as a movement which redirects a man's concentration or consideration from his fundamental errand of driving. These sorts of exercises incorporate utilizing a cell phone, eating and drinking, discussion with co-travelers, self-preparing, perusing or watching recordings, modifying the radio or music player and notwithstanding utilizing a GPS framework for exploring areas. Among the greater part of the above, cell phone utilization is said to be the most diverting component. Diverted driving has been distinguished as an essential hazard factor in street activity wounds. Cell phone use has formed into an essential wellspring of driver diversion as it can prompt drivers to take their consideration off the street, consequently making vehicle tenants more helpless against street crashes. The utilization of cell phones while driving causes four kinds of commonly non-selective diversions – visual, sound-related, subjective and manual/physical. While visual diversions make drivers turn away from the roadway, manual diversions require the driver to grasp their hands off the guiding wheel; sound-related diversions cover those and Sounds that are critical for the driver to hear while driving and intellectual ones incite the driver to consider an option that is other than driving.

A system with web-cam integrated to raspberry pi running python classifier can be used to capture the image and classifying it into either distracted state or safe driving state. If the driver is in a distracted state, a buzzer alarm is generated. For classifier, we have obtained the dataset of drivers driving in different states. This is fed in as our training data set and with an open source machine learning python library Scikit-Learn a classifier is generated to predict the distracted state of the driver. Following states of the driver is to be predicted: texting, talking to co-passengers, phone call, looking left or right, reaching back seat, self-grooming, operating IVIS and eating or drinking. We aim at building an integrated system of webcam and classifier model based on Convolutional Neural Network which would classify images based on different states of the driver. The training set used for building the model has been taken using a static driving

simulator with real human subjects performing a specific secondary task.

2. MOTIVATION AND INITIAL RESEARCH

As it is referred to, road crashes have raised as one of the best reasons for death among the most profitable age gatherings and formed into a noteworthy general wellbeing emergency over the world. As indicated by the World Health Organization (WHO), road crashes execute 1.2 million individuals and for all time debilitate another 50 million consistently. Over the previous decade, road crash has turned into the tenth driving reason for death on the planet and is anticipated to ascend to the fifth position by 2030. India is the main supporter of worldwide road crash mortality and horribleness figures. Consistently, 16 lives are lost to road crashes in India. In the most recent decade alone, India lost 1.3 million individuals to road crashes and another 5.3 million were debilitated forever. While it has only 1% of the world's vehicles, India represents more than 10% of worldwide road crash fatalities. As per the 'Street Accidents in India', 2015 report of Ministry of Road Transport and Highways (MoRTH), 146,133 individuals were slaughtered in road crashes in 2015 alone including 12,589 kids. This number isn't just the most noteworthy that India has ever recorded ever; however, it speaks to a 53.9% expansion in the course of the most recent decade, and almost a 10-overlay increment since 1970. The WHO sorts driver diversion as a critical hazard factor for road crash wounds. The United States Department of Transportation terms occupied driving as a standout amongst the most risky driver practices and a pandemic which has expanded with the multiplication of cell phones.

3. GOALS AND OBJECTIVE

The primary goal of this project is to find a webcam which is compatible with a raspberry pie to capture an image of driver and secondary is to select a well-suited machine learning classifier

A model which can classify this image into various predefined

Classes to predict the distracted state of the driver. The accuracy of this system must be high, and the response time must be slow. Also, the entire system can be integrated with autonomous prevention system such as generating alarm or auto-breaking system to prevent an accident. The system can be improved by extended to real-time analysis of driver by reducing the response time and increasing processing power.

4. HARDWARE PROPOSED

In our project, we have proposed to use

1) Logitech web-cam: 3MP, HD recording (1280 x 720), embedded microphone with noise reduction.

2) Raspberry Pi 3 with 32 GB SD Card: OS used is Rasbian, Single board computer; 1.2 GHz processor can drive an HD desktop.

3) 5V Active Alarm Buzzer Module for raspberry pi with transistor drive module 8550.

5. SOFTWARE PROPOSED

We are designing a classifier model based on machine learning which will take real-time image data to classify the image. We are also working to improve the reaction time and efficiency of the model as it has worked on real-time data.

Further specifications of software parts are-

1) Language Used: Python 2.7 or above.

2) Open source tools and libraries: Anaconda, Scikit learn, Numpy, H5py, Tkinter, etc.

6. EXISTING SYSTEM

Previously in this project there was no hardware was used and the software was also very limited to classify the images already present in the database.

7. SYSTEM'S KEY FEATURES

Our system is a real-time, cost-efficient system with an integrated buzzer provides an alarm to alert the distracted driver, and has a minimalistic design. We provide the user with free software and a reliable device.

8. MATH

The classification of driver's images into various classes will be facilitated using a multi-layer perceptron also known as fully connected convolutional neural network which will use SoftMax activation function in the output layer. The term Fully Connected implies that every neuron in the previous layer is connected to every neuron on the next layer.

Convolution is a numerical idea utilized for the most part in Digital Signal Processing when managing signals that take the form of a time series. It can be mathematically represented as:

$$\begin{aligned} y[n] &= x[n] * h[n] \\ &= \sum_{k}^{n} x[k] * h[n+k] \;\; where\ k \\ &\rightarrow [-\infty, \infty] \end{aligned}$$

Though conventionally called as such, the operation performed on an image with CNN's is not strictly convolution, but rather a slightly modified variant called cross-correlation, in which one of the inputs is time reversed.

$$\begin{aligned} y[n] = x[n] * h[n] \\ = \sum_{k}^{n} x[k] \\ * h[n-k] \;\; where\ k€[-\infty, \infty] \end{aligned}$$

9. ECOSYSTEM SETUP

Creating an Ecosystem means installing and configuring following elements.

1) Take one raspberry pi kit and install one general purpose operating system like Raspbian on external SD Card connected to it.

2) Configure other hardware devices to raspberry pi like buzzer and WebCam for generating alarm and for capturing driver images respectively.

3) Installing Anacondas python on the Raspberry Pi and setting up the configuration.

4) Installing the dependencies such as openCV, numpy, scikit learn and keras on Raspbian OS.

5) Keras by default comes with Tensorflow as backend but we have to shift the backend to theano.

6) Obtaining the dataset of driver images captured by the static simulator installed in vehicle and storing in the secondary storage of system.

7) Splitting the data into training and testing data and further splitting the training data into different classes that the model needs to be trained for. This can be done by porting the whole dataset on the raspberry pi.

8) After training the system is ready to install for capturing driver images and generate alarm if found distracted.

10. CONVOLUTIONAL NEURAL NETWORK

ConvoNets are made for processing data in the form of multi-dimensional arrays, for example a color image composed of three 2-Dim arrays which represents the 3-color spectrums. Following things needs to understand to completely understand Convolutional Neural Networks.

1) Feature (Convolution Kernels): Similarly, as its exacting importance infers, an element is a particular and helpful perception or example acquired from the info information that guides in playing out the coveted picture examination. The CNN takes in the highlights from the info pictures. Regularly, they rise over and again from the information to pick up noticeable quality. For instance, when performing Face discovery, handling two eyes on the face as features.

2) Receptive Field: It is illogical to associate all neurons with every single conceivable district of the information volume. It would prompt an excessive number of weights to prepare and create too high a computational intricacy. Subsequently, rather than associating every neuron to every conceivable pixel, we determine a 2-dimensional locale called the open field stretching out to the whole profundity of the contribution, inside which the included pixels are completely associated with the neural system input layer. It's over these little areas that the system layer cross-segment (each comprise of a few neurons (called profundity segment)) work and deliver the actuation delineate.

3) Zero-Padding: Zero-Padding alludes to the procedure of symmetrically adding zeroes to the info lattice. It's a normally utilized change that enables the measurement of contribution to be acclimated to our prerequisite. It is for the most part utilized as a part of planning the CNN layers when the measurements of the information volume should be protected in yield volume.

4) Hyper parameters: In CNN's, the properties relating to the structure of layers and neurons, such as spatial game plan and responsive field esteem, are called hyperparameters. hyperparameters interestingly determine layers. The principle CNN hyperparameters are the open field, zero-Padding (P), the information volume measurements and walk length.

5) Convolutional Layer: Every neuron relates to a district of the information volume called the responsive field. For instance, for an information picture of estimations 28x28x3, if the open field is 5x5, by then every neuron in the Conv. Layer relates to a region of 5x5x3 in the data volume. In this way every neuron will have 75 weighted wellsprings of information. For an estimation of R, we have a cross-zone of neurons completely committed to taking responsibility from this area. Such a cross-section is known as an importance parcel. It stretches out to the whole criticalness of the Conv. Layer.

6) The Relu (Rectified Linear Unit) Layer: ReLu alludes to the Rectifier unit, the most normally conveyed enactment work for the yields of the CNN neurons. Tragically, the ReLu work isn't differentiable at the inception, which makes it difficult to utilize it with back engendering preparing.

7) Pooling Layer: The Pooling Layer is generally set after the convolutional layer. Its essential utility lies in decreasing the spatial measurements of the Input Volume for the following Convolutional Layer. It doesn't influence the profundity measurement of the Volume.

8) The Fully Connected Layer: The Fully Connected Layer is arranged precisely the way its name suggests: it is completely associated with the yield of the past layer. Completely associated layers are commonly utilized as a part of the last phases of CNN to interface with the yield layer and build the coveted number of yields.

11. CONCLUSION

This project develops a model based on Convolutional Neural Network to predict distracted states of the driver and thus this application could be used in order to acquaint driver for wrong driving deeds which could finally help in accident prevention. This project can be implemented in developing systems that would be able to alert the driver if found distracted and help in preventing the accident. This project works on classifying images, but it could be extended for real-time monitoring application that could detect a distracted state in a video feed and generate an alarm. As we know that most of the accidents that take place in today's world take place due to distracted states of the driver then why not to take a step in preventing it with the use of technology. Our project aimed at achieving that by building this model

12. ACKNOWLEDGMENT

We might want to accept this open door to thank our inner guide Prof. Praveen Hore for giving all of us the assistance and direction we required. We are exceptionally thankful to him for his kind help. His important recommendations were extremely useful. We are likewise thankful to Prof. Sunil R. Dhore, Head of Computer Engineering Department, Army Institute of Technology, Pune for his basic help, proposals.

At last our extraordinary on account of Staff Members for furnishing different assets, for example, a research facility with all required programming stages, nonstop Internet association, for Our Projects.

13. REFERENCES

[1] S. Chen, B. Mulgrew, and P. M. Grant, "A clustering technique for digital communications channel equalization using radial basis function networks," IEEE Trans. on Neural Networks, vol. 4, pp. 570-578, July 1993.
[2] J. U. Duncombe, "Infrared navigation—Part I: An assessment of feasibility," IEEE Trans. Electron Devices, vol. ED-11, pp. 34-39, Jan. 1959.

