

Distracted Driver Detection with Deep Convolution Neural Network

Shivang Singh

B.tech (CSE)

Galgotias Univesity

Greater Noida, India

Jatin Singh

B.tech (CSE)

Galgotias University

Greater Noida, India

Divyansh

B.tech (CSE)

Galgotias University

Greater Noida, India

Department of Computer Science and Engineering Galgotias University

Plot No 2, Sector – 17 A, Yamuna Expressway, Greater Noida , Gautum Buddh Nagar , UP, India

Submitted TO Mr. K. Suresh

Assistant Professor School of Computing Science and Engineering, Galgotias University Greater Noida, India

Abstract

Driver basic cognitive process and distraction are the most causes of road accidents, several of that end in fatalities. to cut back road accidents, the event of data systems to find driver basic cognitive process and distraction is crucial. Currently, distraction detection systems for road vehicles aren't nevertheless wide obtainable or ar restricted to specific causes of driver basic cognitive process like driver fatigue. Despite the increasing automation of driving because of the supply of more and more refined help systems, the human driver can still play an extended role as supervisor of car automation. With this inmind, we tend to review the printed scientific literature on driver distraction detection strategies and integrate the known approaches into a holistic framework that's the most contribution of the paper. supported printed scientific work, our driver distraction detection framework contains a structured outline of reviewed approaches for police investigation the 3 main distraction detection approaches: manual distraction, visual distraction, and psychological feature distraction. Our framework visualizes the total detection info chain from used sensors, measured knowledge, computed knowledge, computed events, inferred behaviour, and inferred distraction sort. Besides providing a sound outline for researchers fascinated by distracted driving, we tend to discuss many sensible implications for the event of driver distraction detection systems that may additionally mix totally different approaches for higher detection quality. we expect our analysis is helpful despite - or perhaps attributable to - the good developments in machine-controlled driving.

Date of Submission: 02-05-2022

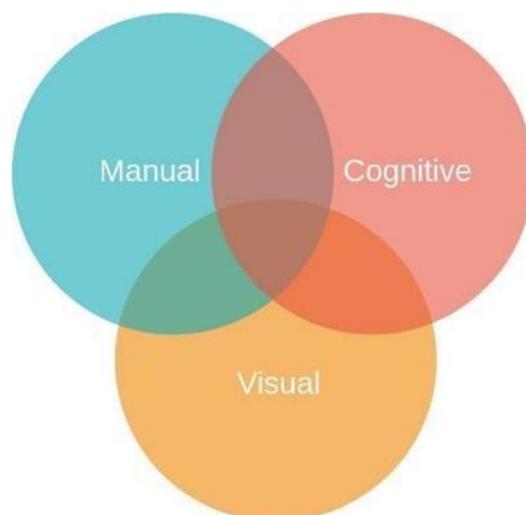
Date of acceptance: 15-05-2022

I. INTRODUCTION

Distracted driving refers to any activity that turns the driving force aloof from the safe driving task. samples of activities which will cause distraction throughout driving ar texting or talking on portable, modify the radio, drinking, reprimand a traveller and hair or makeup [1]. within the past few decades, the quality of living has improved that caused the families to possess their own cars. Thus, individual typically drive their cars for business or trip. However, long distance driving typically makes drivers bored. However, some drivers do another activities instead of specializing in driving that causes loads of accidents. Since the previous couple of years the quantity of automobile accidents has enhanced attributable to distracted driving. National main road Traffic Safety Administration of United State (NHTSA) has been proclaimed three,477 deaths and 391,000 injuries cases in 2015 because of distracted driving [2]. In US, distracted driving thought-about as a significant killed reason by a daily rate of nine cases and one thousand hors de combat cases [3]. At present, major corporations are commencing to work with new systems referred to as Advanced Driver help Systems (ADAS) by making techniques to alert the driving force once detection activities are occur. From this deep learning takes a part of ADAS techniques.

Deep learning (DL) may be a study of artificial neural networks and alternative Machine learning algorithms that hold quite one hidden layer. Deep learning has varied structures like Convolutional Neural Networks (CNNs), Deep Neural Networks (DNNs) and perennial Neural Networks (RNNs). Deep learning has been applied to several fields like pc vision, speech recognition and linguistic communication process [4]. per the Centre for malady management and hindrance (CDC) [5] among styles of driver distractions ar visual, manual and psychological feature. Visual distraction refers to any activity that takes the driver eyes on the road. Example of visual distraction includes checking portable for notifications, searching for things

within the automobile, and observant billboards. psychological feature distraction refers to any activity that causes the mind to lose specialize in the driving. samples of psychological feature distractions ar being attentive to the radio, reprimand passengers within the automobile, or being lost in thought. Manual distractionrefers to activities that cause the drivingforce to require their hands off the handwheel. samples of manual distractions embody adjusting the stereo player, ingestion whereas driving, or smoking. As shown in Figure a pair of, multiple styles of distractions will be alongin a very task. as an example, talking on a portable whereas drive entails manual and psychological feature distractions. To date, there are 3 primary styles of modality wont to acknowledge distracted drivers [5]: a) Physiological knowledge like cardiogram (ECG) and encephalogram (EEG). b) Vehicle management knowledge like pedal positions and handwheel movements. c) Visual knowledge like eye movements, body movements and pictures or videos of the driver's facial expressions.



This study income to specialize in the detection of distracted drivers via visual knowledge. There are varied strategies wont to find distracted drivers with pc vision. Among the foremost notable are [5]: a) Threshold: the best means, that theworth of a feature is compared to a predetermination threshold. b) ancient machine learning: like Support VectorMachine (SVM). c) Deep learning: deep learning has gained a lot of attention recently in distraction detection. There are several recent studies that addresses the matter of police investigation distracted drivers. Jin et al. designed a system to find psychological feature distraction victimisation solely vehicle management knowledge [6].The authors claimed that distracted drivers typically move the handwheel and apply the pedals otherwise than traditional drivers. city et al. used a proportion of eye closure (PERCLOS) to find driver temporary state [7].The frequency and period of a driver'seye glimpse for a minor task ar wont to turn out a complete activity of eyes off the road [8].Pohl et al. planned a distracted driving detection system supported gaze direction and head position [9].The instant distraction decided and also the distraction level was classified. Cray et al. introduced a hidden Andrei Markov model (HMM) based mostly methodology to find distracted driving activities. Their methodology needs the detection of the driver's face and right arm [10]. Abouelnag et al. provided a dataset and introduced a time period distraction detection approach employing a combination 5of 5} Alex internet and five Google internet models with hand, face, and skin options [11]. during this study, main focus is placed on police investigation driver's behaviour via a digital camera put in within the automobile cabin. The digital camera field-of-view covers the higher body ofthe driving force sitting on the driver'sseat. Deep learning is used to classify the photographs taken by the digital camera and confirm whether or not or not the driving force is distracted.

II. DATASET DESCRIPTION

The dataset used in the study is taken from a public Kaggle challenge by State Farm [12].The dataset consists of 22,400 training and 79,727 validation labelled images (640x480 full colour) of driver behaviours. There are a total of ten classesof behaviours provided in the dataset.

Table 1 tabulates the ten distinct behaviours. Figure 2 illustrate a sampleimage for each of the ten classes.

Table. 1 List of distracted driving activities Class

Class	Behaviour
C0	Safe driving
C1	Texting using right hand
C2	Talking on the phone using right hand
C3	Texting using left hand
C4	Talking on the phone using left hand
C5	Operating (Adjusting) the radio
C6	Drinking
C7	Reaching behind
C8	Hair and makeup
C9	Talking to passenger



III. PROPOSED DEEP LEARNING MODEL

Convolutional Neural Network

Convolutional Neural Network (CNN) could be a class of deep learning model that has proved to be terribly effective in areas like image recognition and classification. CNN are booming in distinguishing faces [13], objects and traffic signs except for powering vision in robots and self-driving cars [14]. Figure three illustrates the fundamental design of the CNN. The CNN design involves multiple layers of operations performed on the input image. A typical CNN model includes multiple convolutions followed by a pooling operation one once another. These layers of convolution and pooling in succession if conjointly called the feature extraction layers. The ensuing vectors from the feature extraction layers are unit then two-dimensional and stacked to the classifier layers.

Transfer Learning

Transfer learning could be a technique in deep learning wherever one model that's trained on a task is re-purposed to suit another task. for instance a model that's trained on pictures of house objects are often re-purposed to classify the kind of piece of furniture. All the models employed in this study is originally trained on the ImageNet object classification dataset that Comprised of one,000 categories of objects. so as to use transfer learning, the weights on the highest layers of the pretrained models are unit retrained with the State Farm dataset. These are unit 3 general ways to retrain the pretrained models exploitation the State Farm dataset of distracted drivers. the primary strategy is to retrain solely the last classifier layer of the pretrained model. this can be the foremost uncomplicated strategy exploitation quantity} amount of your time and process power in exchange for model accuracy. The second strategy is to retrain the previous few layers of the model as well as the classifier layer. This strategy might end in improved accuracy likewise as a lot of intensive process power and takes longer time to run. The last strategy is to retrain the complete model from scratch. this can be the

smallest amount fascinating methodology as a result of it takes the foremost time and process price. Figure four illustrates the variations among the 3 careful ways. during this study Strategy one (S1) and Strategy a pair of (S2) is totally evaluated on every of the pretrained model.

Table. 2 List of pre-trained modelsevaluated [16]

Model	Input size(W x H)	Depth	License
Xception [17]	299 x 299	126	Neural network could
VGG19 [18]	224 x 224	26	design containing neuron
VGG16 [18]	224 x 224	23	input bound information Oxford the layers area unit inte
ResNet50 [19]	224 x 224	168	every different and hav
InceptionV3 [20]	299 x 299	159	weights. As we tend to tra
InceptionResNetV2[21]	299 x 299	572	the weights get updat
MobileNet [22]	224 x 224	88	suggests that the model options of our dataset.
DenseNet121 [23]	224 x 224	121	neural network is sam
DenseNet169 [23]	224 x 224	169	networks except for picture
DenseNet201 [23]	224 x 224	201	pictures as input to the
NASNetLarge [24]	331 x 331	-	consists of input layer, out
NASNetMobile [24]	224 x 224	-	variety of hidden layers.

IV. Literature Study

Transfer Learning

Pre-trained models were used as a start line rather than ranging from scratch [6]. It had many edges. The pretrained models we tend to used, are trained on a awfully massive dataset (ImageNet), that contains one.2 million pictures with one thousand classes. To adapt the model we tend toights to our dataset we roughly mounted initial seventy % of the layers (made them untrainable) and trained last thirty %. the rationale being that the initial layers of the model embrace edge detection and form detection modules, that area unit generalized for any image recognition application and these become{increasingly| progressively more and alot of} more abstract within the final layers, creating it a lot of specific to the appliance. within the distracted driver situation, the last layer provides associate output one amongst the ten categories for a given image.

Convolutional Neural Network

be a stratified we tend to to the network, rconnected to e some initial in the network ed and this has learnt the Convolutional e as neural. So, we offer CNN model. It put layer and Hidden layers embrace the Convolution layer, Pooling layer, corrected Linear Units layer, Dropout layer and totally Connected layer. [7]

Input Layer The input layer holds raw constituent values of the photographs. In our case, pictures area unit coloured with resolution of 640*480 pixels that area unit scaled right down to 224*224 to scale back coaching time.\.

Conv Layer

The Conv layer contains a group of learnable filters of little dimensions. These filters area unit touched throughout entire region of input image and at every location a real number is dotted the weights of filter and little region at a lower place the filter. For our project if twelve filters area unit used then output dimension would be 224*224*12.

Pooling Layer

The Pooling layers reduced the second dimensions of input volume to stop from overfitting or to avoid computation inefficiencies. this can be done by applying atiny low filter to computer file on every depth slice. There numerous sorts of pooling filters like gamma hydroxybutyrate pooling that choose gamma

hydroxybutyrate worth below filter, average pooling, etc.

ReLU Layer

It applied activation perform to every component to extend non-linearity of the model. $\max(0,x)$ is associate example of activation perform.

Dropout Layer

The Dropout layer is additional to stop the model from overfitting. it's a regularization methodology that at random sets some activation values to zero to get rid of some feature detectors. In our models we've got additional a dropout layer with worth of zero.5.

4.2.6 FC Layer

In this layer every vegetative cell is connected to outputs from previous layer. This layer provides the ultimate prediction for every category. In our project there are unit ten categories, thus FC layer contains ten neurons.

V. CONCLUSION

Thus, once attempting out many CNN models, our greatest ensemble was created once averaging the possibilities generated by VGG-16, VGG-19 and InceptionV3. the ultimate log loss that we tend to got was zero.795. we tend to solely used processor provided by Google Cloud Platform for this project. If we might have had access to a lot of computing resources, we tend to may have tried to boost our results with the assistance of the following: 1. exploitation KNN to search out out the K nearest neighbors of a picture so generating the ultimate likelihood by considering the typical of the possibilities of those pictures. This approach works well owing to high correlation between the photographs. 2. attempting ResNet-50 and ResNet-152. These 2 CNN models are unit wide used for image classification issues and will offer smart ends up in this situation. 3. Cropping out components from the photographs which offer a lot of data like hands, eyes, etc. to boost accuracy.

REFERENCES

- [1]. Report on Road Accidents in India 2016- Ministry of Road Transport & Highways (MoRTH), Government of India pp. 1-2 <http://morth.nic.in/showfile.asp?lid=2904>
- [2]. Kaggle. A brief summary <https://www.kaggle.com/c/state-farmdistracted-driver-detection>
- [3]. Hssayeni, Murtadha D; Saxena, Sagar; Ptucha, Raymond; Savakis, Andreas, "Distracted Driver Detection: Deep Learning vs Handcrafted Features", Society for Imaging Science and Technology, Imaging and Multimedia Analytics in a Web and Mobile World 2017, pp. 20-26(7)
- [4]. Karen Simonyan, Andrew Zisserman, "Very Deep Convolutional Networks for Large Scale Image Recognition", arXiv preprint arXiv:1409.1556
- [5]. CS231n Convolutional Neural Networks for Visual Recognition <http://cs231n.github.io/convolutional-networks/>

REFERENCES

- [6]. Report on Road Accidents in India 2016- Ministry of Road Transport & Highways (MoRTH), Government of India pp. 1-2 <http://morth.nic.in/showfile.asp?lid=2904>
- [7]. Kaggle. A brief summary <https://www.kaggle.com/c/state-farmdistracted-driver-detection>