iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 Issue: VIII Month of publication: August 2021

DOI: https://doi.org/10.22214/ijraset.2021.37398

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

Detection of Driver's Distractedness Using Convolutional Neural Networks

Shailaja Pasupuleti[1], Shobana Padmanabhan[2]
[1, 2]Alliance University, Bangalore, India

Abstract: Deaths caused due to Road Accidents has always been an area of grave concern and happens to be one of the most critical problems in India as on date. The inattentiveness of the driver causes almost 80% of the casualties. Utilization of mobile phones, talking to passengers, reaching behind to grab something, and drinking while driving are some reasons drivers may lose attention. Distractions are considered of numerous types, out of which we focus on the manual distraction, which is based on the driver's posture. In particular, we consider nine distinct distracted states of the driver. Then, our goal is to detect whether a given image of the driver falls into one of these categories. Eventually, we will integrate an alarm that will alert the driver if she/he is detected to be in a distracted state. Accordingly, this paper presents a mechanism where we use convolutional neural networks; a deep learning technique, to classify driver images. The image dataset we use to train and test our neural network consists of the first dataset made available publicly at the Kaggle data source.
Keywords: Convolutional neural networks, Deep learning methods, Distracted driver, Drowsiness detection.

I. MOTIVATION

Driving is defined as a complex task requiring concurrent execution of various cognitive, physical, sensory, and psychomotor skills. Despite these complexities, it is usual for the drivers to engage in various non-driving-related activities while driving. With the increased usage of wireless communications like mobile phones, which are more sophisticated entertainment systems, and the introduction of technologies such as route navigation and the internet into vehicles, preoccupation with electronic devices while driving is becoming a common practice. Any activity that competes for the driver's attention while driving has the potential to degrade the driver's concentration and has severe consequences on road safety. Few examples of driver distraction activities are texting or talking on a mobile phone, adjustment of the FM radio, drinking, talking to a passenger, hairdo or makeup etc.[1] According to the latest census data, usage of mobile phones and cognitive distractions were reported to be as high as 79.3%. in India leading to road crashes and fatalities while mobile use during driving has increased by 33% compared to last year. This scenario necessities an automatic alarming system that can alert a person operating in this situation which is the focus of our research work.

II. INTRODUCTION

Distraction in common terminology is defined as "A diversion of attention away from activities critical for safe driving toward a competing activity."

According to the latest findings from CDC - Centre for Disease Control and Prevention , Distraction of drivers have been classified into three main types, namely: Visual distractions - activities of the driver which lead to driver-eyes off the road during driving like checking on mobile phone for messages or notifications, observing advertisements on bill boards, searching for items in car while driving etc. Second main area is of Manual distractions like doing another manual activity or taking the hand off the steering wheel. Few examples are - adjusting radio volume, eating and driving,smoking, playing with a kid etc and lastly the Cognitive distractions which refer to an activity that causes the mind to lose focus while driving. Few examples are listening to radio, talking to passengers, being lost in thought, day dreaming while driving etc..Figure below shows multiple types of distractions which can also coexist together in a task. For example, a person talking on a mobile phone while driving entails manual and cognitive distractions.

To date, there are three primary types of modalities used to recognize distracted drivers [5]:

1) Physiological data such as an ECG - Electrocardiogram & EEG - Electroencephalogram,
2) Vehicle control data such as position of pedals and steering wheel rotation and movements.
3) Visual data such as eye-ball movements, body posture changes and images or videos of the driver facial expressions.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VIII Aug 2021- Available at www.ijraset.com

Based on these, mobile phone usage and cognitive distractions were reported as the leading causes of road crashes and fatal injuries. Drivers perceived mobile phones and mental stress as top distractions while driving. Drivers who said cognitive distractions were more likely to have met with an accident than those who did not. The prevalence of distracted driving practices was as high as 75.6% for mobile phone use, 79.3% for cognitive distraction, and 62.2% for other distractions. The NHTSA - National Highway Traffic Safety Administration of the United States of America has declared that in the year 2015, a total of 3,477 deaths and 3,91,000 injuries have been reported due to distracted driving [2]. In the US, distracted driving has been classified as a single primary reason for death/kill with a daily rate of 9 points and 1000 injury cases [3].

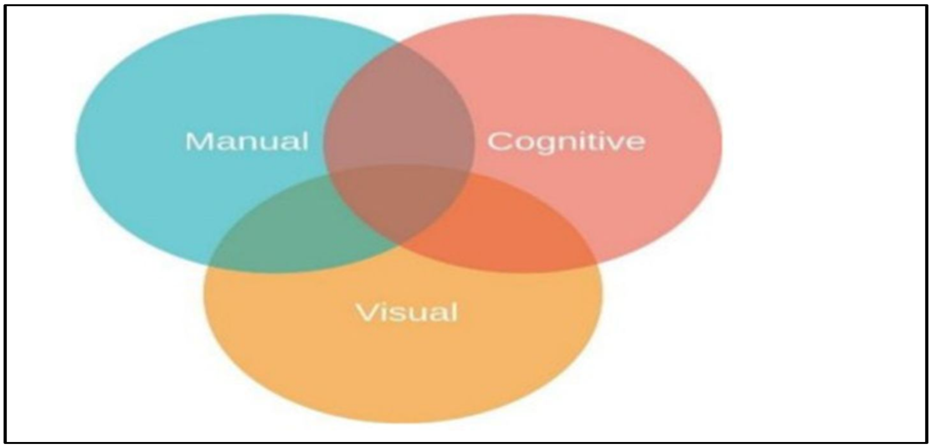


Fig:1 Main types of driving distraction

III. RELATED WORK

Earlier, a lot of research has been done in the area of cognitive and visual distractions and some work done in problem domain of distracted driver detection. Cognitive distraction focuses on the mental instability of the driver, while the visual distraction focuses on the "eyes off the road" behavior of the driver.

In [17], Yuang Liang et al. have conducted a study on detecting cognitive distractions in real time. They have utilized the concept of Support Vector Machines(SVM). SVM is a prevalent and efficient data mining method. The classification was based on specifications like eye-ball movement and driving behavior. The primary cause of visual distractions is the utilization of cell phones [2].

Motivated by the same, some researchers carried work on mobile phone utilization detection while driving vehicle. Zhang et *al*. [19] created a database using a dashboard mounted camera and used the model Hidden Conditional Random Fields to detect mobile phone usage, which basically operates on the face, mouth, and hand features of the person driving the vehicle.

In 2015, Nikhil *et al*. framed a dataset for hand detection in the automotive environment [5] and achieved a precision of 70.09% using(ACF) Aggregate Channel Features object detection. Authors used Method of Supervised Descent, Histogram of Gradients (HoG), and an AdaBoost classifier and achieved 93.9% classification accuracy.

Martin *et al*. [18] presented a vision-based analysis framework which recognizes in-vehicle activities. The hand-on-wheel information obtained using two Kinect cameras providing frontal and back views of the driver. Authors extended their research to include eye cues to the previously existing head and hands cues [13]—however, only three types of distractions considered here.

Zhao *et al*. [20] designed a more inclusive distracted driving dataset with the side view of the driver considering four activities, namely safe driving, operating shift lever, eating, and talking on a mobile phone. In 2016, Yan *et al*. [14] came up with a Neural Network CNN based solution that achieved a 99.78% classification accuracy.

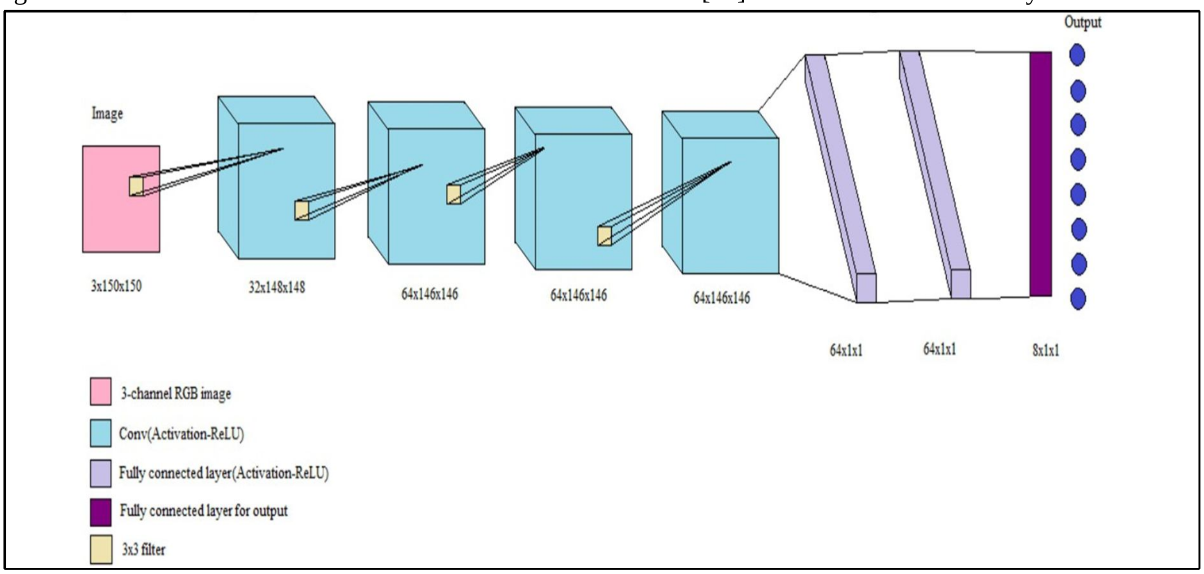
IV. DEEP LEARNING MODEL ARCHITECTURE

Convolutional Neural Networks (CNN) is a deep learning model which is very effective in image recognition and classification[13]. The CNN architecture involves performance of multiple layers of operations on the input image. A typical CNN model includes multiple convolutions followed by a pooling operation one after the other. The convolution layer, also known as the feature extraction layer, creates a feature map predicting the class probabilities for each input image feature. To perform the mapping, the layer applies a filter that scans the whole image, considering few pixels at a time.

An Activation function is used on the matrix formed upon viewing these filters. In our system, we use the ReLu Activation function that activates the node based on a specific predefined threshold value. Finally, a pooling operation is performed that selects the maximum element from the region of the feature map covered by the convolutional filter.CNN has successfully identified faces, objects, and traffic signs apart from powering vision in robots and self-driving cars [15]. The resulting vectors from the feature extraction layers are then flattened and stacked to the classifier layers.

1) *Convolution Layer:* The convolution layer extracts features from the image input and preserves the relationship between pixels by learning these image features using a filter or kernels. We used convolution of 3x3 kernel at various layers for training our model.
2) *Activation Layer:* We used ReLU as an activation function that activates a node if the input is above a particular threshold value.
3) *Max Pooling Layer:* Max pooling progressively reduces the representation's spatial size to reduce the number of parameters and computation in the network.
4) *Fully Connected Layer:* Input to fully connected layer is output from final pooling or convolutional layer which is flattened and then fed into a fully connected layer.

Figure-2 shows the architecture of convolutional neural networks [15]: CNN Architecture for key Point Detection



V. OUR APPROACH AND METHODOLOGY

This project aims to recognize unsafe behavior and send real-time feedback to the driver using short sound alerts, vibration in the steering wheel, etc. The presented system uses Convolutional Neural Network to extract crucial features in the warning system's training phase. The proposed dataset for training has been categorized into nine classes, labeled as Reaching behind, Drinking, Makeup, Talking phone left, Talking phone right, Operating Radio, Texting phone left, Texting phone right, and Yawning. We include images of the Yawning of the person while driving, which helps to predict the drowsiness and fatigue of the driver. So, the system is able to depict the distraction due to unsafe behavior and drowsy state of a driving person, which may at times lead to terrible road accidents.

In our system, initially, the images are captured using a high precision dashboard camera. These images are pre-processed and rescaled to maintain uniformity of the pictures using a computer placed inside the vehicle and are fed in a deep learning CNN model used to train these images as follows. The input is a three-channel 256 x 64 image, next would be two layers of convolutional filters of size 3 x 3 and 32 filters. A max-pooling layer follows with a stride of 2 x 2. Two convolutional layers follow of filter size 3 x 3 and 64 filters. Again, a max-pooling layer follows with a stride of 2 x 2. Finally, two fully connected layers are added; one with 64 units followed by a dropout of 50 percent and the final output layer of 9 units. Each entry in these nine units gives the probability of the image belonging to the respective class. The network is trained for 15 epochs with a fixed learning rate of 0.01 and batch size 32. Finally, the captured image of the driving person is classified into one of these classes categorized. Based on the results obtained, the driver is alarmed using a sound alert.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VIII Aug 2021- Available at www.ijraset.com

VI. DATASET

The dataset we used for the work is from the Kaggle data source. In April 2016, The State Farm's distracted driver detection competition on Kaggle was the first dataset to consider a wide variety of distractions and made available publicly [12]. They released their dataset of 2-Dimensional dashboard camera images for a Kaggle challenge conducted. The dataset had nearly 22400 training images and 79727 testing images. The resolution was 640 x 480 pixels. We framed a sample dataset of 1500 images by modifying a few parameters and adding images under the class of "yawning." Images are preprocessed by applying normalization and scaling techniques to make them appropriate for training the model. The training images had corresponding labels attached as follows:

A. Reaching Behind
B. Drinking
C. Makeup
D. Talking Phone Left
E. Operating Radio
F. Yawning
G. Talking Phone Right
H. Texting Left
I. Texting Right

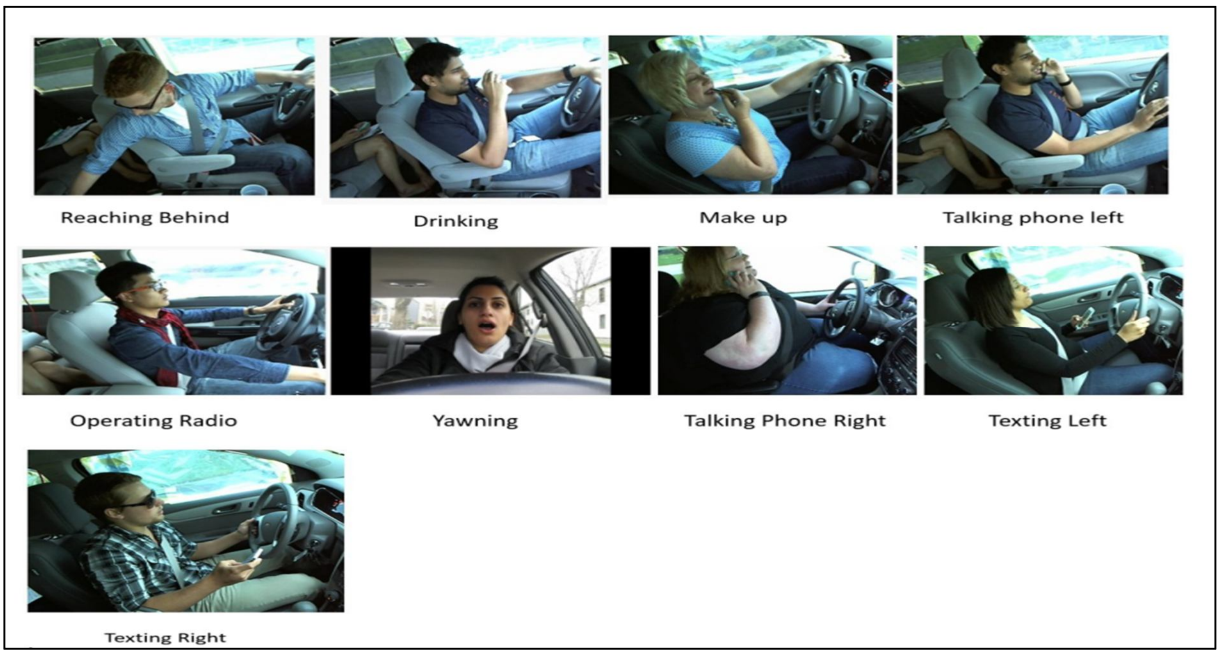


Fig:3 Labeled Training Images

VII. RESULTS AND ANALYSIS

We consider two metrics for evaluating our deep learning model—accuracy of the model's prediction and value of the loss function. Loss function is often used in training process to find the best (optimal) values for the model's parameters such as the "weights" assigned to the synapses in the neural network. Under these optimal parameter values, the model's predictions are compared against the known values. The percentage of correct predictions is defined as "accuracy". For our application, the accuracy is 99% with the training data and 97% with the validation data. The loss value, which shows the behavior of the model after every iteration of optimization, optimized using "Adam" optimizer and "categorical cross entropy". The loss value for our model is approximately 20% with respect to both the training and the validation data. The lower the loss function value, the better. The figure depicts the accuracy values as well as the loss function values for both the training and validation sets. Whether the loss function value can be reduced further by fine tuning the model is part of our ongoing work

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VIII Aug 2021- Available at www.ijraset.com

The accuracy of training and testing data are plotted as shown in the following figures:

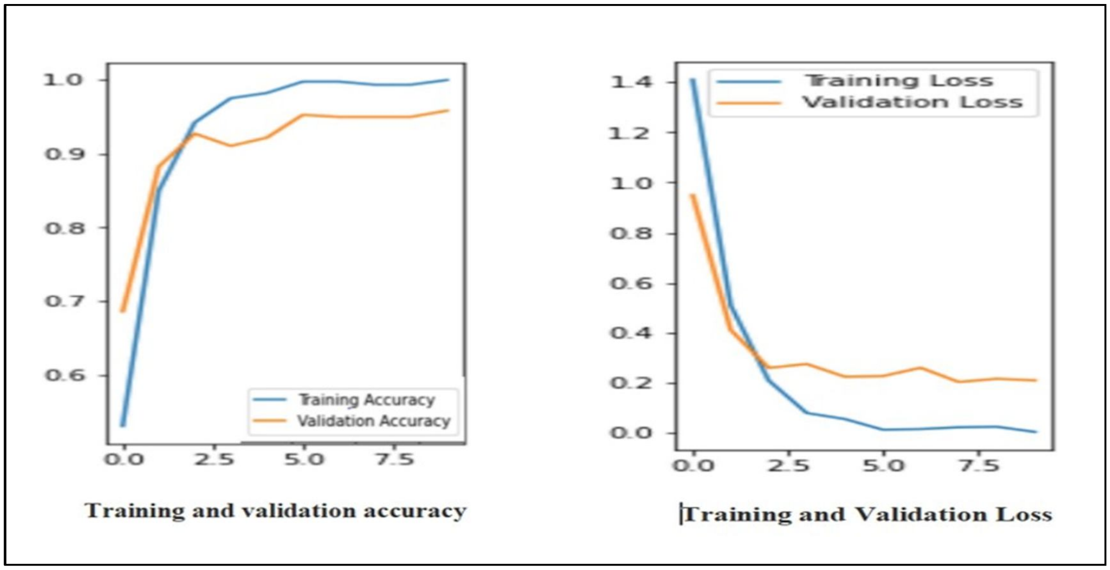


Fig:4 Training and Validation Accuracy

VIII. OUTPUT OF TRAINED MODEL -PREDICTING IMAGE:

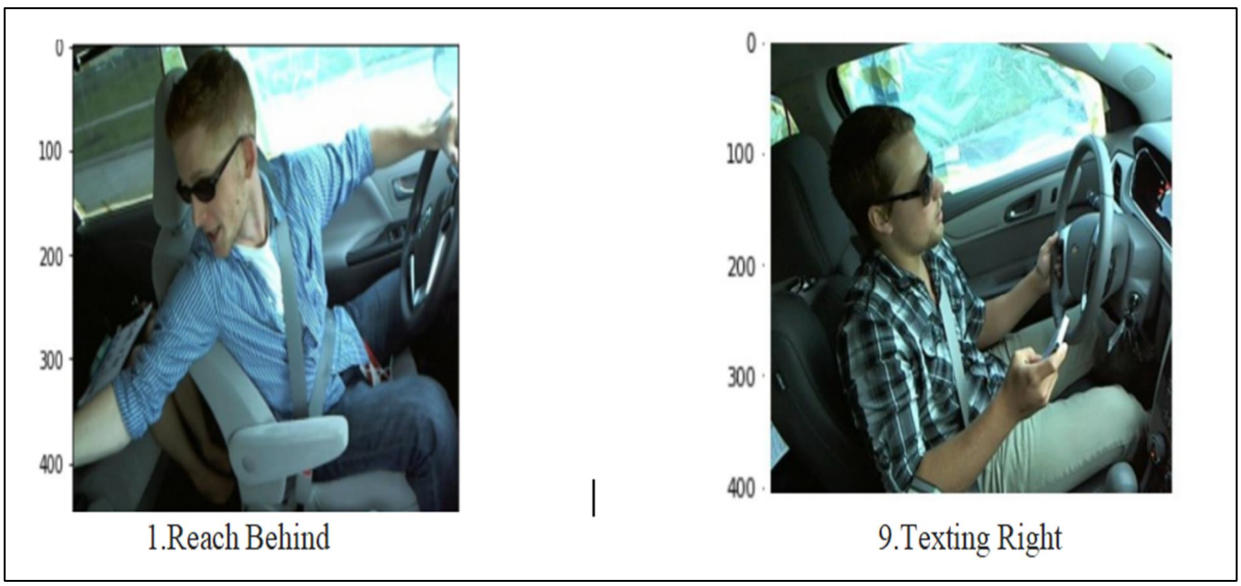


Fig: 5 Predicted image

IX. CONCLUSION AND FUTURE SCOPE

Our work uses the deep learning model of convolutional neural networks to predict driver distraction behavior. Experimental results depict a very high accuracy of the prediction. Apart from manual distractions, we also tried to portray fatigue and drowsiness in the person driving a vehicle, reducing the rate of accident-prone situations.

In our ongoing work, we are working on installing the driver's distraction detection systems in cars. A dash board camera is installed in the vehicle to monitor the drivers driving and alert the driver in real-time by generating a beep sound whenever the driver is distracted. These kinds of systems would be especially beneficial for taxi/cab aggregators such as Ola and Uber. They can monitor their drivers in real-time and prevent potential accidents. Statistics about driver actions and behaviour can also be collected for future behavioral forecast and prediction if required. Car/Cab rental service providers like Avis, Hertz, Zoom cars etc can also use these systems to check on their customers.

REFERENCES

[1] Vasanti S, Neha P and Aniruddha H. Distracted driver detection using CNN and data augmentation Techniques. In: International Journal of Advanced Research in Computer and Communication Engineering(IJARCCE), 2018, 7(4):130-135.
[2] National Highway Traffic Safety Administration (NHTSA).Distracted driving 2015. Retrieved from https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812381; March 2017.
[3] Center for Disease Control and Prevention (CDC). Distracted driving. Retrieved from https://www.cdc.gov/motorvehiclesafety/distracted_driving/index.html; 30 March 2019.
[4] Parijat O. Artificial intelligence, machine learning, and deep learning. In: 15th International Conference onICT and Knowledge Engineering (ICT&KE), 2017.
[5] Duy T, Ha MD, Weihua S, He B and Nikhil et al. Real-time detection of distracted driving based on deep learning. IET Institution of Engineering and Technology, 2018, 12(10):1210-1219.
[6] Listing J, Qingming N, Haijing H, Huacai X, Yali W and Dongdong S. Driver cognitive distraction detection using driving performance measures. Discrete Dynamics in Nature and Society, 2012.
[7] Pooneh. and Reza. AZ. Drowsiness detection is based on the brightness and numeral features of the eye image. In: 2009 Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, IEEE,12-14 Sep 2009.
[8] Farber E, FoleyJ, Scotts. Visual attention design limits for ITS in-vehicle systems: the society of automotive engineers standard for limiting visual distraction while driving.Transportation Research Board.
[9] Pohl J, Birk W and Westervall L. A driver-distraction-based lane-keeping assistance system. In: Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering, 2007, 221(4):541-52.
[10] Craye C and Karray F. Driver distraction detection and recognition using RGB-D sensor. arXiv preprintarXiv:1502.00250. 2015.
[11] Abouelnaga Y, Iraqi HM, and Moustafa MN. Real-time distracted driver posture classification. arXivpreprint arXiv:1706.09498, 2017.
[12] Kaggle: State farm distracted driver detection. Retrieved fromhttps://www.kaggle.com/c/state-farm- distracted-driverdetection/data; 31 March 2019.
[13] ParkhiOM, Vedaldi A, and Zisserman A. Deep face recognition. In: MVC. 1(3):6, 2015. 14. Redmon J and Farhadi A. Yolov3: an incremental improvement.arXiv preprint arXiv:1804.02767, 2018.
[14] C. Yan, F. Coenen, and B. Zhang. Driving posture recognition by convolutional neural networks. IET Computer Vision, 10(2):103–114, 2016.
[15] Cornelisse D. An intuitive guide to convolutional neural networks. Retrieved from https://www.freecodecamp.org/news/an-intuitiveguide-to-convolutional-neural-networks-260c2de0a050; 24 April 2018.
[16] Keras applications. Retrieved from https://keras.io/applications/; 31 March 2019.
[17] Y. Liang, M. L. Reyes, J. D. Lee.: Real-time detection of driver cognitive distraction using support vector machines. IEEE transactions on intelligent transportation systems, vol. 8, no. 2, pp. 340-350 (2007)
[18] S. Martin, E. Ohn-Bar, A. Tawari, and M. M. Trivedi. Understanding head and hand activities and coordination in naturalistic driving videos. In 2014 IEEE Intelligent Vehicles Symposium Proceedings, pages 884–889, June 2014.
[19] X. Zhang, N. Zheng, F. Wang, and Y. He. Visual recognition of driver hand-held cell phone use based on hidden crf. In Proceedings of 2011 IEEE International Conference on Vehicular Electronics and Safety, pages 248–251, July 2011
[20] C. H. Zhao, B. L. Zhang, J. He, and J. Lian. Recognition of driving postures by contour let transform and random forests. IET Intelligent Transport Systems, 6(2):161–168, June 2012. 1151