Driver Distraction Detection and Early Prediction and avoidance of accidents using Convolutional Neural Networks

A.Christy1*, Prayla Shyry2, Meera Gandhi G.3, M.D. Anto Praveena4

1,2,3,4Sathyabama Institute of Science and Technology, India, Chennai-119
ac.christy@gmail.com

Abstract. In this fast-moving world, accidents in four wheeled vehicles occur due to the break failure or because of the carelessness or the fatigue of the driver. The driver pattern of the driver plays a major role in providing road safety as well as in fuel consumption. The distraction of drivers is found by installing various sensors which is used for gathering real time data. The behaviour of drivers under stress condition and their behavioural patterns for early detection and avoidance of accidents are found using convolutional neural networks. Convolutional Neural Networks are efficient classifiers in handling image processing and computer vision problem. The input dataset is a collection of driving behaviour of 10 different drivers collected from Kaggle. The behaviour of drivers under 7 distracted situations like texting, talking through phone, playing music, drinking, eating, doing make up and talking to passenger are considered. The batch normalization is used at the right of the input layer in order to avoid skewing of data at a direction. It is shown, the convolutional neural networks at 4 epochs have shown 99% accuracy.

1. Introduction

According to World Health Organization, 1.24 million people die due to road accidents every year. The heart of any vehicle relies on the amount of assistance provided to the driver by the vehicle’s support system. Autonomous cars provide technical solutions to tackle heavy traffic and congested locations in an efficient manner. Internet of Things (IoT) paves its way in this era in connecting various devices for intelligent data capture and processing. In this paper, we propose a novel method to gather data from four wheeled vehicles using sensors process them and alert the driver from the occurrence of accidents.

2. Literature Survey

Ahmed S. Shamsaldin et al (2019) has implemented a population-based algorithm named Donkey and Smuggler Optimization Algorithm (DSO). This optimization algorithm follows the concept of searching and selecting routes adopted by donkeys for transportation. Donkeys have the tendency to help each other. A donkey trying to cross a fence gets the support of another. This concept supports in
obtaining the optimal solution. If the best solution is over loaded, then instead of chopping of this solution, the second-best solution can be adopted until the first solution becomes feasible to the situation.

For hybridization and improved optimization Whale optimization Algorithm (WAL) can be adopted Farhad Soleimanian Gharehchopogh and HojjatGholizadeh (2019). Fengfei Wang et al (2019) has proposed an improved kNN algorithm which reduce the dimension of the vector space and improves the precision and speed of classification performed with KNN. Jingzhao Li and ZihuaChen (2019) has developed a security information system for intelligent analysis using perception module based on CNN algorithm for coal mine safety inspection. Wrapper and Embedded are the various Feature Selection methods, Girish Chandrashekar and Ferat Sahin (2014). Filter methods adopt ranking methods such as Pearson correlation coefficient which can detect linear dependencies among obtained and target defined as defined in eqn.(1).

\[
\mathbf{r} = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{\sqrt{[n\Sigma x^2 - (\Sigma x)^2][n\Sigma y^2 - (\Sigma y)^2]}} \quad \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots (1)
\]

Mutual information (MI) is another filter method, in which Information theoretic ranking criteria named shannon’s entropy is defined. It can measure the dependency between two variables and is defined in eqn. (2).

\[
\mathbf{H(X)} = - \sum_{i=1}^{n} p(x_i) \log_b p(x_i) \quad \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots (2)
\]

From eqn. (2) b is the base of the logarithm used. By observing the variable X, the uncertainty in the output is reduced. The time complexity for measuring uncertainty, distances, consistency or dependency is less expensive than the accuracy of a learning process. The wrapper method for feature selection can be classified as Sequential Selection Algorithms and Heuristic Search Algorithms.

Jiongming Jiang et al (2018) has proposed artificial bee colony optimization with Differential Evolution (DE), which can prevent the occurrence of local optimum. J.A. Healey and R.W. Picard (2005) has analyzed the data collected from ECG in order to find the driver stress level. Maziar Yazdani and FariborzJolai (2016) has proposed the population based Lion optimization algorithm (LOA), considering an optimization problem of N dimensions and the number of solutions in the search space as \( N_{\text{pop}} \). A Lion is represented as :

\[
\text{Lion} = [x_1, x_2, x_3, \ldots, x_N] \quad \text{and the Fitness value of each Lion is computed by evaluating the cost function as:} \quad \text{Fitness value of lion} = \mathbf{f} (\text{Lion}) = f(x_1, x_2, x_3, \ldots, x_N)
\]

Malinda Vania et al (2019) has adopted a segmentation method using Convolutional Neural Network (CNN) inorder to segment spine from CT images. CNN is efficient in applications related to image classification, object recognition and image segmentation. Seyed Mostafa Bozorgi and Samaneh Yazdani (2019) has proposed Improved Whale Optimization problem (IWOP), which overcomes the premature convergence of getting stuck with local optima performed by WOP. Sudipta Chowdhury et al(2019) has compared Adoptive Large Neighborhood Search (ALNS) with Ant Colony Optimization (ACO). It is proved that ALNS based algorithms are adoptive to the changing circumstances. Xingsi Xue and Junfeng Chen (2019) has adopted meta-heuristic Tabu search algorithm
for matching identical entities used across sensor ontologies. YanliLv et al(2018) has studied the attacks on networks using coordinated scan algorithm. The overall characteristic of the algorithm is studied by considering the scan sequence as a time series.

3. System Architecture

Studies on driver behavior depend upon simulators or based on data collected from multiple sensors. The car is equipped with GPS, accelerometer, gyroscope and magnetometer. The purpose of these sensors is to monitor real time data. The system for modeling the system is depicted in Fig.1.

![System Architecture](image)

In this implementation, four sensors, GPS, accelerometer, Gyroscope and Magnetometer in order to monitor and control the vehicle from varying speed, applying brake and variation of lane changes. GPS is a satellite based global positioning system, which is used to find the position of location on earth using latitude and longitude values. Smart Phone Hardware sensor converts physical quantity which could be interpreted using devices. They are measured along x, y and z direction as m/s² values. Gyroscope can either detect or measures the orientation of the device from its angular rate. It functions adopting the principles of angular momentum and is represented in 3-axis by pitch, roll and yaw. Magnetometer measures the magnetization or magnetic strength of materials like Ferro magnet.

The pitch angles are received using sensor fusion and X-axis of the Accelerometer. Lane turn or change is recorded using yaw angles and Y-angles of the accelerometer. The location and speed limit is obtained through GPS. Harsh brakes are observed if brake is applied to more than 20 mph within 2 sec.

The GPS provide information to the sensor regarding the latitude, longitude value, max speed allowed, traffic details, etc. The data collected by the sensors are converted to normalized world coordinates. This conversion is mandatory in order to obtain device coordinates inside the vehicle. Converted sensor data is then pre-processed. The attributes best suited for prediction are obtained using Convolutional neural networks.

4. Convolutional Neural Networks

Raw sensor data records the values of 3 axes and time stamp, which indicates the time of data collection. The sensor data cannot be passed to the classifier, as it may contain missing data. A sliding window with the sensor data is hence sliced to a time frame of one-second in order to create the attributes required for classification. The process continues as the first frame is send across sliding window. A feature vector represents the details of sliding window collected from sensor data as depicted in Fig. 2. The feature vector is updated whenever a driving event happens.
The proposed model receives the sensor data as the input layer, middle layer and a classifier layer. Let the number of dimensions of the input data sample be $D_x$ and the number of classes on the output layer be $N_y$. The middle layer of the Convolutional neural network with Convolutional, Pooling and recurrent layers along with a fully connected network is depicted in Fig. 3.

Data extracted from the sliding window is passed into the Convolutional layers, which are then passed through the pooling layer. Multiple levels of Convolutional layers are used for feature extraction, which ignores the noise present in the data collected through the sliding window. The output layer is used to obtain class probability distribution for identification of driver behaviour. The Convolutional layer and the pooling layer initiate Convolutional operations on sliding window sensor data. Each class of outputs produced by the Convolutional layer termed as feature map records the extracted features. The number of feature map obtained from the (l-1) of the Convolutional layer is $n_{l-1}$ having the feature map size as defined in equation 3.

$$m_{l-1} = w_{l-1} * h_{l-1}$$

$$n_{l-1} = n_{l-1} * m_{l-1}$$

The total number of neurons in the layer (l-1) is defined in eqn (4) and the feature map k from the output layer l of the Convolutional layer is defined in eqn. (5).

$$n_{l-1} = n_{l-1} * m_{l-1}$$
\[ X^{(l,k)} = \sigma \left( \sum_{p=1}^{P_{l-1}} W^{(l,k,p)} \otimes X^{(l-1,p)} + b^{(l,k)} \right) \tag{5} \]

where \(\sigma\) indicates ReLU activation function \(W^{(l,k,p)}\) and indicates 2D filter mapping from \(p^{th}\) feature map on the \(l-1^{th}\) layer to the \(k^{th}\) feature map of the \(l^{th}\) layer. The convolutional neural network with ReLU activation function could extract the essential features through the feature map whereby avoiding over-fitting. The output of pooling layer is obtained from eqn(6).

\[ X^{(l+1)} = \text{down} \left( X^{(l)} \right) \tag{6} \]

5. Results and Discussion

Installation of cameras and sensors inside vehicles has helped us in observing driver behaviour in an accurate manner. Data is a collection of 10 different states of drivers with one safe driving mode and 9 distracted driving modes. The dataset is taken from the state farm obtained through Kaggle. Data extraction and pre-processing is obtained by 9 categories. C0 indicates safe driving, c1: texting – right, c2: talking on the phone – right, c3: texting – left, c4: talking on the phone – left, c5: operating the radio, c6: drinking, c7: reaching behind, c8: hair and makeup and c9: talking to passenger. A sample of 4 images taken from different distractions is depicted in Fig.4.

Fig. 4 Driver distractions

Initially a simple model, named as linear model without any hidden layer is used. Total number of parameters used is 903,186 in which there are 903,180 trainable parameters and 6 non-trainable parameters and the output has produced 94% accuracy.

The batch normalization is provided at the input layer to avoid any input values. The output with softmax layer for 10 classes is studied with 4 epochs as shown in table 1.
Table 1 Accuracy from CNN

Epoch 2/3
7083/7083 [==============================] - 54s 8ms/step - loss: 0.0788 - acc: 0.9763 - val_loss: 0.0262 - val_acc: 0.9902
Epoch 3/3
7083/7083 [==============================] - 54s 8ms/step - loss: 0.0686 - acc: 0.9804 - val_loss: 0.0135 - val_acc: 0.9970

The validation from linear model has produced 94% accuracy whereas CNN with 4 epochs has produced 99% accuracy.

6. Conclusion
The performance comparison of our model with linear model and CNN is studied and by using the ReLU activation function in a CNN model has produced 99% accuracy. The proposed framework can learn features from the given dataset without much of preprocessing. Driver monitoring and analysis is the process of automatically extracting driver’s data (e.g., latitude, longitude, speed, acceleration, etc) and computes safety score. The live extraction of driver score can be obtained from sensor data instead of obtaining data from synthetic datasets.

References

[1] Colic, Aleksandar, Marques, Oge, Furht, Borko (2014),” Driver Drowsiness Detection”, Systemsand Solution
[2] Daminan Salapatek, Jacek Dybala, Pawelczapski and Pawel Skalski (2017), Driver Drowsiness Detection Systems, Proceedings of the Institute of Vehicles, 3(112) / 2017
[3] Ellen De Pauw, Stijn Daniels, Tom BrijsElke Hermans and Geert Wets (2014),” An evaluation of the traffic safety effect of fixed speed cameras”, Vol. 62, Pp.168-17
[4] Mercedes Ayuso, Montserrat Guillén Ana María, Pérez-Marin (2014), “Time and distance to first accident and driving patterns of young drivers with pay-as-you-drive insurance”, Accident Analysis and Prevention, Vol. 73, Pp. 125-131
[5] Esra Vural, Müjdat Çetin, Aytül Erçil, Gwen Littlewort, Marian Bartlett and Javier Movellan (2008), "Machine Learning Systems for Detecting Driver Drowsiness”, In-Vehicle Corpus and Signal Processing for Driver Behavior, Pp. 97-110
[6] Ian Y. Noy, David Shinar and William J. Horrey (2018),” Automated Driving: Safety Blind Spots”, Safety Science, Vol. 10, Pp. 68-78.
[7] S. Vaithyasubramanian, P. M. S. Chandu and D. Saravanan (2020). Intensify of Metrics with the Integration of Software Testing Compatibility. In Intelligent Computing in Engineering, Pp. 693-699. Springer, Singapore.
[8] Kristie L. Young and Paul M. Salmon (2012),”Examining the relationship between driver distraction and driving errors: A discussion of theory, studies and methods”, Safety Science, Vol. 50, Issue 2, Pp.165-174
[9] Markus Maurer, J. Christian Gerdes Barbara Lenz and Hermann Winner (20..),” Autonomous Driving-Technical, Legal and Social Aspects”, Springer Open.
[10] D. Saravanan, D. Joseph and S. Vaithyasubramanian (2020). Effective Utilization of Image Information Using Data Mining Technique. In Recent Trends and Advances in Artificial Intelligence and Internet of Things, Pp. 207-215. Springer, Cham.
[11]. Michael A Regan, Charlene Hallett and Craig P. Gordon (2011), “Driver distraction and driver inattention: Definition, relationship and taxonomy”, Accident Analysis and Prevention 43, Pp. 1771-1781
[12]. Muhammad Rizwan Ullah, Muhammad Aslam, Muhammad Imran Ullah and Martinez-Enriquez Ana Maria (2017),”Driver’s Drowsiness Detection Through Computer Vision: A Review”, 16th Mexican International Conference on Artificial Intelligence, pp 272-281
[13] Nanxiang Li and Carlos Busso (2013),” Modeling of Driver Behaviour in Real World Scenarios using Multiple Noninvasive Sensors”, IEEE Transactions on Multimedia, Pp. 1-25.
[14]. Oliver Carsten and Marieke H. Martens (2018),” How can humans understand their automated cars? HMI principles, problems and solutions”, Cognition, Technology &Work, Vol. 21, Issue1, Pp.3-20
[15]. Pushpa Choudhary and Nagendra R.Velaga (2017), “Modelling driver distraction effects due to mobile phone use on reaction time”, Transportation Research Part C: Emerging Technologies,Vol. 77, Pp.351-365
[16]. Ramon Zatarain-Cabada, Maria Lucia Barron-Estrada , Francisco González-Hernández And Hector Rodriguez-Rangel (2019),” Emotion Recognition Using A Convolutional Neural Network”, Advances In Computational Intelligence Pp 208-219