International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 05 | May - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Driver Distraction Detection Using Machine Learning

Nikhitha A[1], Ramya K R[1], Swati R Hegde[1], Vaishnavi[1], Mrs. G V Sowmya[2]

[1]UG students, Dept. of IS&E, JNNCE, Shimoga, India.
[2]Assistant Prof, Dept. of IS&E, JNNCE, Shimoga, India.

---***---

Abstract - The problem of driver distraction has become serious, and this is due to physical distractions that can be dangerous on the roads. Convolutional Neural Networks (CNNs) are used in this research to introduce a new approach to identifying physical distractions. To train the designed CNN model, a large dataset of labeled images that depict different physical distractions such as eating, texting, and manipulating controls is utilized. We make use of the spatial hierarchies learned by CNNs that enable us to extract important information for distraction identification with high efficiency. The experimental results validate the proposed method's ability to recognize physical distractions in real driving scenarios effectively. Incorporating such models into in-car systems could therefore result in reduced distractions and consequently fewer accidents and additionally offer immediate warnings or interventions thus improving driver safety considerably.

Key Words: Distraction, Machine Learning, Convolution Neural Network, Road accidents.

1. INTRODUCTION

The main source of concern as regards road safety is driver distraction. This ranges from physical distractions such as talking on phone, drinking, doing grooming or adjusting the radio which can divert the driver's attention off the road leading to accidents. To put preventative measures in place and reduce related risks, real-time detection of these kind of distractions is important. One possible solution to this problem is Convolutional Neural Networks (CNNs) which are among the newest advances in machine learning that could be used to detect physical distractions from in-car videos. The main source of concern as regards road safety is driver distraction. This ranges from physical distractions such as talking on phone, drinking, doing grooming or adjusting the radio which can divert the driver's attention off the road leading to accidents. In order to put preventative measures in place and reduce related risks, real-time detection of these kind of distractions is important. One possible solution to this problem is Convolutional Neural Networks (CNNs) which are among the newest advances in machine learning that could be used to detect physical distractions from in-car videos. The study aims at finding various physical distractions commonly encountered by drivers, using CNN algorithms. Examples of these distractions include talking on a cell phone (with both hands), reaching behind, sipping water, watching TV, grooming, adjusting the radio, and interacting with other passengers. To create reliable systems that can accurately detect instances of distraction in real driving scenarios researchers train CNN models using labeled datasets that contain several Examples of these distractions. Traffic safety is significantly influenced when CNNs are utilized to recognize physical distractions. By automatically evaluating in-car camera data, CNN-based mechanisms can identify cases of distraction and immediately inform drivers or activate precautionary measures aimed at reducing risks. Such technologies could be incorporated in vehicles to reduce distracted driving-related accidents therefore averting injuries as well as fatalities.

2. Related Work

The article, "Driver Distraction Detection Methods: A Literature Review and Framework" [1] investigates the need to be aware of driver distraction and highlights the importance of focus in self-driving cars. It is concerned with different sensors employed to sense distractions and classifies them as intrusive or non-intrusive. The study also highlights that the recent developments in detection technologies are geared towards driver distraction aspects like eye trackers and exterior cameras. Furthermore, the issue of vehicle speed must be taken into account when detecting distractions.

Gradient Boosting Classifier is a competition-winning idea dealing with hard-to-predict observations through iterative boosting of weak learners and loss function optimization [2]. It does better than other classifiers which are such as ANN, RF, and LR by having uniformity across all the classes and using micro scores for imbalanced datasets. The results from the simulation tell us that the algorithm can be used to derive multi-class models with good performance in different classes as shown by precision-recall and ROC curves.

The technology used in the study included a 2017 Land Rover Discovery test vehicle, fitted with GPS, steering angle sensors, speed sensors, a CAN bus system, and a central recording computer [3]. The ds PACE system integrated the speed, steering angle, and GPS channels while a Logitech C520 web camera on the windscreen observed the driver's behavior. In addition to this, there was another computer sitting in the passenger seat that recorded videos from the camera as well as one LED light synchronized time series data between the camera and the vehicle's CAN-Bus system. The experiment was carried out at the MUEAVI testing ground in a closed environment for controlled flexible driving conditions during research.

The main focus of this project is to explore various types and techniques of image classification that can be used in detecting distracted drivers [4]. The process of training involves fitting a distribution of probability based on labeled training samples to estimate the class membership. Decision Tree uses simple decision rules to classify images. The findings reveal that SVM classifier can yield different accuracy when weight matrix initialization is changed. Various models such as Linear SVM, Decision Tree, and Two-layer Neural Network were employed for classification. Experimentation includes preprocessing, model implementation, and assessment that effectively detects distracted drivers using images.

For driver distraction detection, the proposed model incorporates a Channel Attention Mechanism to enhance feature selection [5]. The data is prepared for training and testing by using geometric augmentation, resizing, and normalization as preprocessing steps. On the AUCD2 dataset it achieved impressive accuracy rates of 98.97% while on the SFD3 datasets it got 99.58%, making state-of-the-art models like HRRN and ConvNet inferior to this option. Stacking Ensemble, VGG19, and Drive-Net were also compared to these models, proving that our proposal was far superior in this regard. Moreover, such an approach made efficient real-time driver drowsiness detection possible through the use of EfficientNetB0 architecture.

3. Methodology

The proposed System consists of different phases in the framework. It combines the data-driven approaches to build an effective driver distraction detection model.

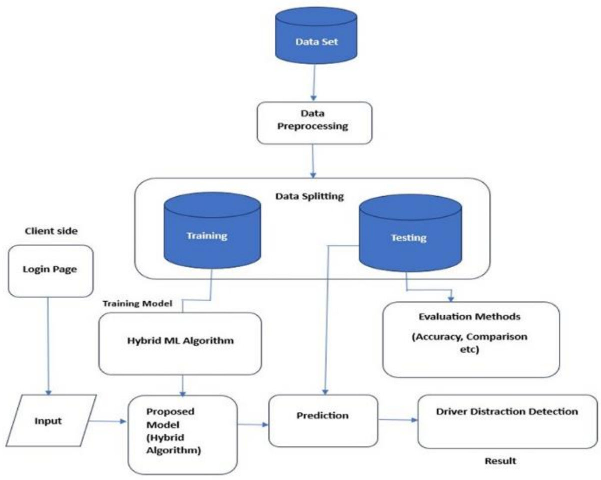

Fig-1: Block Diagram of Proposed Methodology

3.1 Dataset:

The data set is a compilation of pictures taken from different activities the driver may be engaged in such as safe driving, Texting-right, talking on the phone-right, Texting - left, talking on the phone - left, operating the radio, drinking, talking to the passenger, reaching behind and doing things like adjusting hair or makeup. A label is given for each similar activity type indicating what action the driver is performing. This dataset was created primarily for the development and evaluation of a model that will enable accurate detection and classification of driver behaviors. The dataset plays an important role in creating a driver assistance system. They are so kind that they can classify distracted drivers from non-distracted ones.

3.2 Data Preprocessing:

The data preprocessing is to ensure the quality and consistency of the input data. Preprocessing is carried out in two foremost steps image loading and resizing. In the image loading step, images are loaded using the OpenCV library. The get_cv2_image is responsible for carrying out this task, taking the way to the picture record as input. According to the declared color type, the function will process the image to either grayscale mode or to the color mode. This works efficiently in handling different color formats, which ensures that it is compatible with various types of inputs.

At first, the image is loaded and then resized depending on the estimated size according to OpenCV's resize function. This resizing step is very necessary to make all images have the same input dimensions as well that are required for successful training of machine learning models like VGG16 used in this context.

3.3 Data Splitting:

Splitting the data is a crucial step because without the proper data splitting, model training, validation, and evaluation can't be performed effectively. It is a process of splitting the data into the training data set, validation data set, and testing data set. Customarily data splitting certifies that the model on one set of data, validated on the other set of data, and finally tested on a completely independent set of data. The data splitting is done in such a way that 70% of the data is for training, 15% for validation, and 15% for testing.

3.4 Training:

This phase undergoes the loading of the pre-trained VGG16 model without the top layers. The model consists of a vgg_std16_model function which is alters the VGG16 model trained on ImageNet to the model for the specific task. The model is complied with an appropriate loss function and optimizer.

Before commencing the training, the pre-trained weights are transferred from the file. Model's parameter initialization is done by these weights, through which training will begin with a strong foundation based on previously learned representations. The set is used for training the model using the techniques like data augmentation to improve the generalization. In Training data has been iterated over batches, to update the model parameter each batch undergoes forward and backward propagation. The essential parameters, batch size and number of epochs will control the updates and the duration of training.

3.5 Evaluation Methods:

The trained model is evaluated based on the testing set to monitor its performance. Evaluation is usually held in two stages validation set evaluation and testing set evaluation, during the training phase some of data are held as a validation data set, through the model performance can be assessed periodically.

After the completion of the training process, the final performance of the model is evaluated by the testing data set, which still has been not used in the training and validation set. The weights of the trained model are loaded, and the predictions are compared with the testing data labels.

3.6 Hybrid Algorithm:

A hybrid algorithm combines the different machine learning techniques or a architecture to improve the performance of the model. The deep learning model is used for the feature extraction and machine learning model is used for the classification. The VGG16 model will extract a high-level feature from the input images, which will be later used for the classification. By using this Hybrid algorithm, the robustness and recognition task Can be achieved. Science the VGG16 model is used, which is a deep convolution neural network (CNN), which is a most powerful algorithm for the feature extraction. An image classification algorithm is already very powerful. But combining it with traditional machine learning methods will improve its performance even more. The reason why this VGG16 model is so popular (and it's highly effective in image categorization tasks) lies in the fact that it is a famous CNN architecture. This model has multiple convolutional neural layers, each followed by a max-pooling layer for hierarchical feature extraction. The final step of this process is the global average pooling and dense layers designed to enable classification using characteristics drawn from the above. As soon as it has been loaded, cross-entropy is utilized as a loss function, rmsprop as optimizer and accuracy as evaluation metric. This setup is crucial in training the model to identify images relevant to various activities a driver might engage in while driving.

3.7 Prediction:

The webcam takes the snapshot while carrying out the forecast and the image is processed for the forecast of driver behavior based on its content by using a pretrained VGG16 model. When an image is picked from the driver with use web cam, it will be used in real time recognition of the driver's action. This particular photo will then be used as data input when predicting what the outcome could be in terms of the driver's attitude.

The function gets_cv2_image function is used to capture some images; these images are preprocessed so that their sizes become standardized within a well-defined range. After preprocessing it uses the VGG16 model for prediction. Consideration has been given to the probability of matching instances across all classes and this will help you predict what is going to happen next among drivers. The driven class has been chosen based on the maximum probability out of several instances compared and only one class per frame can be predicted resulting in the final decision reached as mentioned above.

3.8 Result:

The model's performance is evaluated by using the validation dataset. The evaluate generator method is employed through which the test score and accuracy metrics are computed. How the model will generalize to unseen data and its overall performance is understood by this evaluation process. The test score represents the average loss, in a validation set how close the model predictions are close to the real answers.

4. Result and Future work

Conv2D and Max-Pooling layers in convolutional neural networks (CNNs) are of the utmost importance for machine learning-based driver distraction detection. Conv2D layers extract features from input images. A series of learnable filters is applied to certain areas of the input image to assist the network in identifying spatial patterns and traits—such as hand or facial movements—that are connected to distracted driving. These filters capture hierarchical representations of features by convolving across the entire image at several levels of abstraction. On the other hand, max-pooling layers preserve only the most crucial information by down sampling the feature maps created by conv2D layers, therefore reducing computational complexity. By selecting the highest value inside each pooling zone, max-pooling layers assist in retaining important information while rejecting less important elements. This aids in the network's generalization.

Snapshot of loading the dataset into training is showed in the figure 4.1. The purpose of this preprocessing stage is to guarantee that the input data is uniform and prepared for training, which is essential for obtaining the best possible model performance and accuracy.

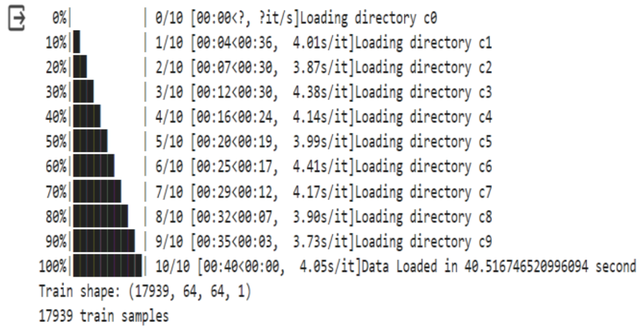

Fig-2: Loading the data for training

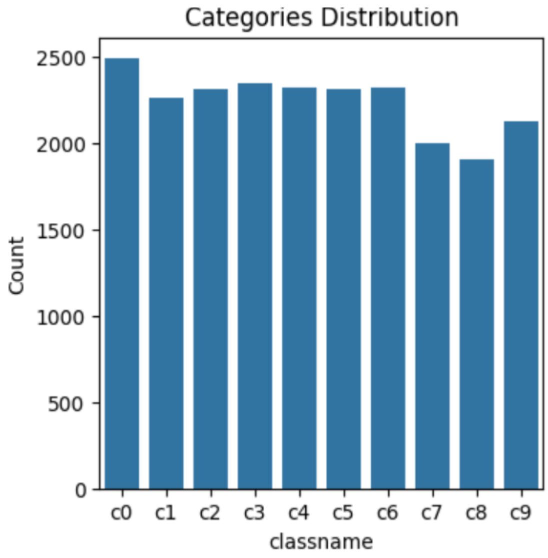

Fig-3: Snapshot of plot of data distribution in each category

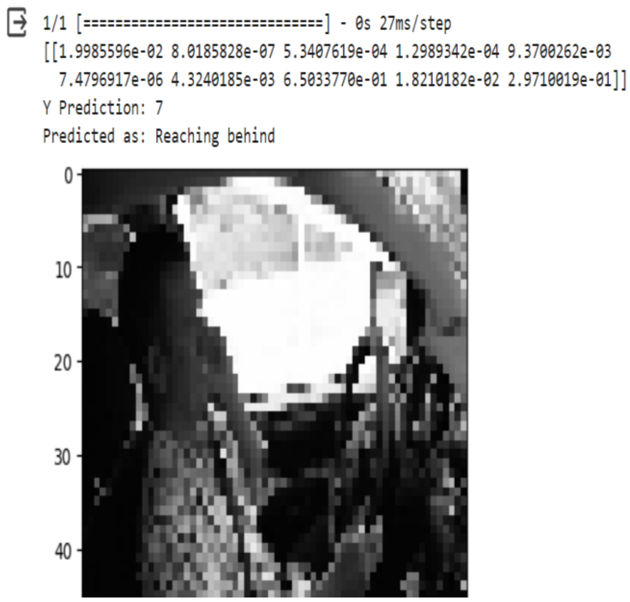

Fig-4: Predicted as reaching behind

Fig-5: Predicted as talking on the phone-left

Future developments in driver distraction involve the incorporation of sophisticated technologies that are capable of identifying cognitive and visual distractions. This means using state-of-the-art technologies like artificial intelligence and computer vision to track the driver's movements and surroundings in real time. Systems to recognize visual distractions can follow a driver's gaze and spot any distractions in the car or on the road. Cognitive distraction detection, on the other hand, looks at physiological signs, facial expressions, and eye movement patterns to determine how mentally taxed the driver is. Future car safety systems that combine these features will be able to reduce distractions and provide proactive alerts and interventions. This will improve overall road safety and lower the likelihood of accidents brought on by distracted driving.

5. CONCLUSIONS

By effectively utilizing CNN models, we have shown that it is feasible to identify and classify a range of physical distractions such as phone use, eating and control adjustments using visual cues obtained from in-car cameras. This development has far-reaching implications for road safety when designing proactive non-intrusive alert systems that can Step in whenever drivers get involved into any distracting acts. Moreover, incorporating CNN based detectors into motor vehicles would be a significant improvement in reducing distractions that might cause accidents. A future study could thus optimize CNN structures, enhance dataset variability and provide distraction detection by means of multimodal sensor fusion to improve precision and resilience. Essentially, the adoption of these distracting detection technologies powered by CNNs on a large scale can completely change road safety hence saving millions of lives and averting uncountable accidents. The achievement of all the rewards attendant upon this transformational breakthrough will require collaboration between researchers, automakers, public policy makers and other stakeholders in this endeavor.

REFERENCES

1. Alexey Kashevnik, Roman Shchedrin, Christian Kaiser, And Alexander Stocker, “Driver Distraction Detection Methods: A Literature Review and Framework”, IEEE Access, vol. 9, pp. 60063-60076,2021.

2. Ghandour R, Potams A.J, Boulkaibet I, Neji B, A1 Barakeh, “Driver Behavior Classification System Analysis Using Machine Learning Methods” Appl. Sci, vol. 11,2021. 10562.https://doi.org/10.3390/app112210562.

3. Cao, Dongpu & Fotouhi, Abbas & Auger, Daniel & Zhang, Zhaozhong & Velenis, Efstathios. (2020). “Driver distraction detection using machine learning algorithms” International Journal of Vehicle Design, Vol. 83, Issues 2-4, pp. 122-139, 2021.DOI:10.1504/IJVD.2020.115057

4. Demeng Feng Department of Materials Science and Engineering Stanford University, Stanford, “Machine Learning Techniques for Distracted Driver Detection”,Spring

5. Qingwan Xue, Xingyue Wang, Yinghong Li and Weiwei Guo, “Young Novice Drivers Cognitive Distraction: Comparing Support Vector Machines and Random Forest Model of Vehicle Control Behavior”. Sensors, vol.23, pp. 1-15,2023. https://doi.org/10.3390/s23031345.

6. Albadawi Y, AlRedhaei A, Takruri M, “Real-Time Machine Learning-Based Driver Drowsiness Detection Using Visual Featurs”. J.Imaging ,vol.9,no 91.2023http://doi.org/10.3390/jimaging9050091.

7. Taimoor Khan, Gyuho Choi and Sokjoon Lee, “EFFNet-CA: An Efficient Driver Distraction Detection Based on Multiscale Features Extractions and Channel Attention Mechanism”, Sensors,vol. 23,pp 1-18,2023 3835. https:// doi.org/10.3390/s23083835

8. Alberto Fernández 1, Rubén Usamentiaga 2, Juan Luis Carús 1 and Rubén Casado 2, “Driver Distraction Using Visual-Based Sensors and Algorithms”, Sensors 2016, vol. 16, pp. 2-44.1805; doi:10.3390/s16111805.

9. Hong Vin Koay, Joon Huang Chuah a, Chee-Onn Chow a, Yang-Lang Chang b “Detecting and recognizing driver distraction through various data modality using machine learning: A review, recent advances, simplified framework and open challenges” 0952-1976/© Elsevier Ltd.,2022.

10. Hesham M. Eraqi, Yehya Abouelnaga, Mohamed H. Saad and Mohamed N. Moustafa “Driver Distraction Identification with an Ensemble of Convolutional Neural Networks”, Hindawi Journal of Advanced Transportation Volume 2019, Article ID 4125865, pp. 12 https://doi.org/10.1155/2019/4125865.

11. Apurva Misra, (Member, IEEE), Siby Samuel, Shi Cao, (Member, IEEE), And Khatereh Shariatmadari, “Detection of Driver Cognitive Distraction Using Machine Learning Methods”,IEEE Access, vol.11,pp. 18001-18012. 10.1109/ACCESS.2023.3245122