Distracted Driver Detection

Dr Jyoti Kaushik, Ankit Mittal, Mohit Soni, Aditya Singh

Department of Computer Science Engineering, Maharaja Agrasen Institute of Technology, India

---

ABSTRACT

Inattentive driving is a leading cause of road accidents, and as a result, there has been an increase in the development of intelligent vehicle systems that can assist with safe driving. These systems, known as driver-support systems, use various types of data to track a driver's movements and provide assistance when necessary. One major source of data for these systems is photographs of the driver that are taken with a camera inside the car, which may include images of the driver's face, arms, and hands. Other types of data that may be used include the driver's physical state, auditory and visual aspects, and vehicle information. In this research, we propose the use of a convolutional neural network (CNN) to classify and identify driver distractions. To create an efficient CNN with high accuracy, we used the Visual Geometry Group (VGG-16) architecture as a starting point and modified it to suit our needs. To evaluate the performance of our proposed system, we used the StateFarm dataset for driver-distraction detection.

Keywords: Convolutional Neural Network (CNN), VGG-16, Confusion matrix.

1. INTRODUCTION

The number of cars sold yearly exceeds 70 million, and there are more than 1.3 million fatal vehicle accidents yearly. India is responsible for 11% of all traffic-related fatalities worldwide. 78% of accidents are attributed to drivers. It is heartbreaking to see that road accidents are a leading cause of death for young people between the ages of 5 and 29 according to the WHO report. Tragically, the number of fatalities keeps increasing each year due to driver distraction. In India, 17 people lose their lives every hour due to vehicle accidents. Driver distraction is the most common cause of motor vehicle accidents and is defined as any activity that diverts the driver's attention away from their primary task. The driver is the most crucial component of the vehicle's control system, which includes steering, braking, acceleration, and additional processes. All traffic participants must be able to complete these essential activities safely. This research suggests a method for detecting driver distractions that recognizes various forms of distractions by watching the driver through a camera. Our objective is to create a high-accuracy system that can track the driver's movement in real-time and determine if the driver is operating the vehicle safely or engaging in a certain type of distraction. The system will classify them appropriately using adequate Machine Learning based on their activities.

2. LITERATURE SURVEY

The reviews of some of the pertinent and important works from the literature for detecting distracted driving are summarised in this section. The major cause of manual distractions is the usage of cell phones.

Researchers employed a support vector machine (SVM)-based model to identify cell phone use while driving [1]. The dataset used for the model focused on two driving activities: a motorist with a phone and a driver without a phone, and depicted both hands and faces with a predefined assumption. Frontal photographs of the drivers were utilised with an SVM classifier to detect the driver's behaviour. Subsequently, other researchers studied the detection of cell phone use while driving. They employed a camera mounted above the dashboard to create a database and a Hidden Conditional Random Fields model to identify cell phone usage. Zhang et al. [2] mainly utilised hand, mouth, and face features for this purpose. In 2015, Nikhil et al. [3] developed a dataset for hand detection in the automotive environment and utilised an Aggregate Channel Features (ACF) object detector to achieve an average precision of 70.09%.

Driver distractions can be detected using a variety of visual indicators and mathematical models. This study examines the use of machine learning (ML) algorithms to identify driver distractions. ML models are trained to recognize certain patterns associated with distracted driving, such as pupil diameter, eye gaze, head posture, facial expressions, and driving posture. These models are then used to identify driver distractions in real time.[4]

A support vector machine-based (SVM-based) model was created to identify drivers using cell phones while operating a vehicle by extracting features from an image. The driver's face was depicted in frontal view images for the dataset, with and without a phone.

Researchers have been developing datasets and object detectors to detect hands in an automotive environment. In one study, an object detector with aggregate channel features was used to achieve an average precision of 70.09% [1]. Another study focused on detecting a driver using a cell phone. The authors used AdaBoost classifiers and a histogram of gradients (HOGs) to locate the landmarks on the face, and then extracted bounding boxes from the left to the right side of the face [2]. Segmentation and training were used to achieve 93.9% accuracy at 7.5 frames per second. To further explore distracted driving, a more comprehensive dataset was created that took into account four different activities: safe driving, using the shift lever, eating, and talking on a cell phone [3]. The contourlet transform and random forest were used by the authors to achieve an accuracy of 90.5%. Additionally, Faster R-CNN [4] was suggested and a system with a pyramid of the histogram of gradients (PHOG) and multilayer perceptron was used to produce an accuracy of 94.75%.

In recent years, the StateFarm distracted driver identification competition on Kaggle has become the first publicly available dataset to consider a wide range of distractions. This dataset outlines ten postures to be detected, including safe driving and nine distracting behaviours. To accurately identify these postures, many researchers have turned to a combination of handcrafted feature extractors, such as SIFT, SURF, and HoG, and classical classifiers like SVM, BoW, and NN. It is clear, however, that CNN's have proven to be the most successful technique for achieving high accuracy. Although the dataset is available for public use, the rules and regulations of the competition restrict its use to competition purposes only.

3. METHODOLOGY

CNNs have made significant strides in recent years in a number of applications, including image classification, object identification, action recognition, natural language processing, and many more. Convolutional filters/layers, Activation functions, Pooling layer, and Fully Connected (FC) layer are the fundamental components of a CNN-based system. These layers are effectively stacked on top of one another to create a CNN. Since 2012, CNNs have advanced quite quickly thanks to the availability of massive amounts of labelled data and computational capacity. In the field of computer vision, many architectures like as AlexNet, ZFNet, VGGNet, GoogLeNet, and ResNet have developed standards. In this study, we investigate and tweak the Simonyan and Zisserman [16] VGG-16 architecture for the job of distracted driving detection.

Convolutional Neural Network (CNN):

A Convolutional Neural Network (CNN) consists of an input layer, a convolution layer, a pooling layer, a fully connected layer, and an output layer (see Figure 1). The input layer takes in images of the driver's current state and the convolution layer uses these images to extract features. The pooling layer calculates the feature values from the extracted features. Depending on the complexity of the images, the convolution and pooling layer can be extended in order to extract more details. The fully connected layer combines the output from the earlier layers into a single vector, which can then be used as input for the next layer. Finally, the output layer categorises the plant disease based on the input.

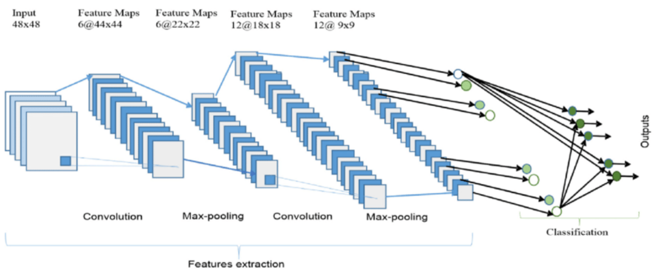

Figure: 1

VGG - 16 and VGG - 19:

VGG is a standard deep Convolutional Neural Network (CNN) architecture consisting of multiple layers. It was developed by the Visual Geometry Group and is the basis of a large number of object recognition models. VGG-16 and VGG-19 consist of 16 and 19 convolutional layers respectively and are capable of outperforming many other models on a wide range of tasks and datasets. It remains one of the most popular architectures for image recognition

the VGGNet-16 supports 16 layers and can classify images into 1000 object categories, including keyboard, animals, pencil, mouse, etc. Additionally, the model has an image input size of 224-by-224.

The concept of the VGG19 model (also VGGNet-19) is the same as the VGG16 except that it supports 19 layers. The “16” and “19” stand for the number of weight layers in the model (convolutional layers). This means that VGG19 has three more convolutional layers than VGG16.

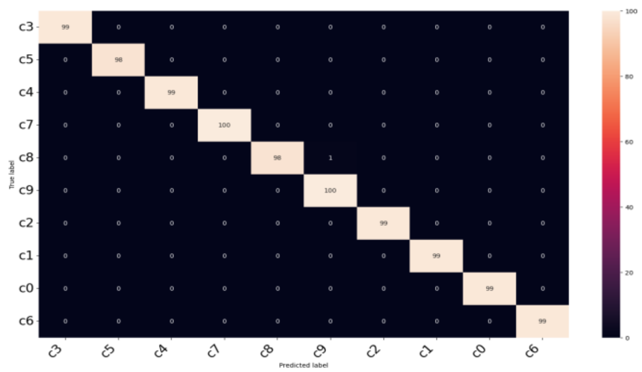

Figure - 2

4. DATA SET

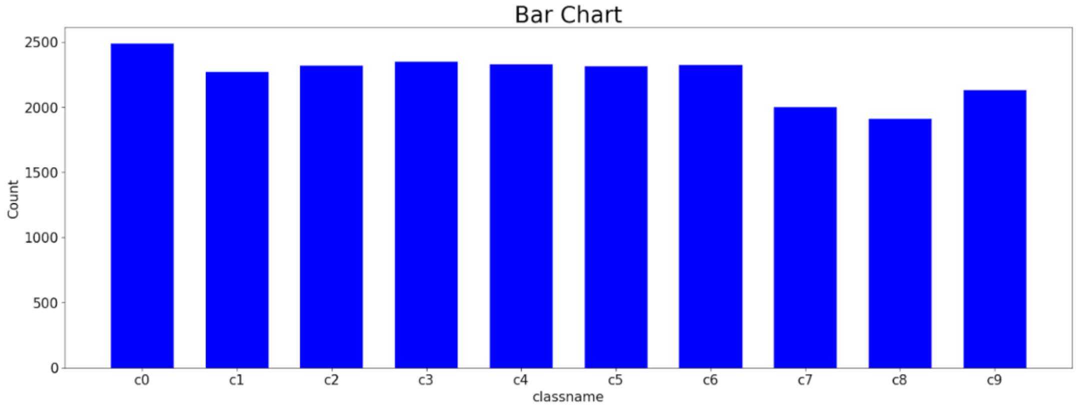

Figure - 3

The dataset has collected the training images from the State Farm. It consists of a variety of images distributed in 10 categories. The distribution is arranged in a way so that there will be very less possibility of duplicate images. All the images are of a driver sitting inside the car. There are a total of 22,424 images classified into 10 categories. The count of images in every category is shown in figure[3]. The categories are:

- C0: 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
- C9: talking to a passenger

5. RESULTS

Figure - 4

Figure - 5

With the hope of increasing accuracy, the convolutional neural network is being tested for the detection of a distracted driver. The database is split into two datasets, training, and testing. CNN determines whether a Driver is distracted or not, and if so, it also forecasts the type of distraction. The CNN model was trained using a 25-epoch. The performance of the CNN model on the testing dataset during training is shown in Figures 4 and 5. Figure 2 displays a sample confusion matrix.

6. CONCLUSIONS

The number of accidents involving vehicles is on the rise, and driver distraction is a major contributing factor. To combat this, a system that can identify and alert drivers to different activities, such as talking on the phone, texting while driving, eating, drinking, and conversing with passengers, is needed. To aid in such research, the StateFarm dataset was created and is publicly available through Kaggle. It contains 10 different classes, with 70% of the data used for training and the remaining 30% for testing and validation. The VGG architecture was used in this investigation to develop effective models based on the image attributes of the dataset, and the test photos were used to evaluate these models. With the help of this dataset, researchers can develop a system that can help reduce the number of accidents caused by driver distraction.

6. REFERENCES

[1]: Mangayarkarasi Ramaiah , Vanmathi Chandrasekara , Madhavesh Vishwakarma: A comparative study on Driver Distraction Detection using a Deep Learning model , pp 1-7 . (2022)

[2]: "Analysis On Driver Distraction Detection And Performance Monitoring System", International Journal of Emerging Technologies and Innovative Research (www.jetir.org), ISSN:2349-5162, Vol.9, Issue 5, page no.j101-j108, (2022).

[3]: Neslihan Kose, Okan Kopuklu, Alexander Unnervik, and Gerhard Rigol," Real-Time Driver State Monitoring Using a CNN-Based Spatio-Temporal Approach" Volume 8, Issue 5 pp. 328-333, 2022

[4]: Amal Ezzouhri, Zakaria Charouh,Mounir Ghogho,Zouhair Guennoun "Robust Deep Learning-Based Driver Distraction Detection and Classification".pp 1-8.(2021).

[5]: Inayat Khan, Sanam Shahla Rizvi, Shah Khusro, Shaukat Ali, Tae-Sun Chung, "Analysing Drivers' Distractions due to Smartphone Usage: Evidence from AutoLog Dataset", Mobile Information Systems, vol. 2021, Article ID 5802658 ,pp 1- 14 pages, 2021.

[6]:A. A. Kandeel, A. A. Elbery, H. M. Abbas and H. S. Hassanein, "Driver Distraction Impact on Road Safety: A Data-driven Simulation Approach," 2021 IEEE Global Communications Conference (GLOBECOM), Madrid, Spain, 2021, pp. 1-6

[7]:A. Kashevnik, R. Shchedrin, C. Kaiser and A. Stocker, "Driver Distraction Detection Methods: A Literature Review and Framework," in IEEE Access, vol. 9, pp. 60063-60076, 2021

[8]: K. Seshadri, F. Juefei-Xu, D. K. Pal, M. Savvides and C. P. Thor, "Driver cell phone usage detection on Strategic Highway Research Program (SHRP2) face view videos," in Proc. CVPR, Boston, MA, USA, pp. 35–43.(2020)

[9]: Prof. Pramila M. Chawan, Shreyas Satardekar, Dharmin Shah, Rohit Badugu, Abhishek Pawar." Distracted Driver Detection and Classification", Int. Journal of Engineering Research and Application ISSN: 2248-9622, Vol. 8,2019

[10]: G. Sikander and S. Anwar, ''Driver fatigue detection systems: A review,'' IEEE Trans. Intell. Transp. Syst., vol. 20, no. 6, pp. 2339–2352.(2018)

[11]: Vaishali1, Shilpi Singh2 1PG Scholar, CSE Department, Lingaya's Vidyapeeth, Faridabad, Haryana, India 2Assistant Professor, CSE Department, Lingaya's Vidyapeeth, Faridabad, Haryana, India. RealTime Object Detection System using Caffe Model. Volume: 06 Issue: 05 pp 5727- 5732.2018

[12]: R. P. A. S. Murtadha D Hssayeni, Sagar Saxena. Distracted driver detection: Deep learning vs handcrafted features. Volume 10 Issue 5 pp. 01-05. 2017

[13]: N. Das, E. Ohn-Bar, and M. M. Trivedi. On the performance evaluation of driver hand detection algorithms: Challenges, dataset, and metrics. pp. 2953-2958.(2015)

[14]: S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," pp. 91-99.2015

[15]: E. Ohn-Bar and M. Trivedi. In-vehicle hand activity recognition using the integration of regions. In IEEE Intelligent Vehicles Symposium (IV), pages 1034–1039.2014

[16]: N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. Dropout: A simple way to prevent neural networks from overfitting. J. Mach. Learn. Res The Journal Of machine learning research, volume 15, Issue 1, pp 1929-1958.2014

[17]: R. A. Berri, A. G. Silva, R. S. Parpinelli, E. Girardi and R. Arthur, "A pattern recognition system for detecting use of mobile phones while driving," in Proc. VISAPP, Portugal, pp. 1–8.(2014)

[18]: C. H. Zhao, B. L. Zhang, J. He and J. Lian, "Recognition of driving postures by contourlet transform and random forests," IET Intelligent Transport Systems, vol. 6, no. 2, pp. 161–168. 2012

[19]: X. Zhang, N. Zheng, F. Wang, and Y. He. Visual recognition of driver hand-held cell phone use based on hidden CRF pp. 248-251. 2011