Driver Distraction Detection using Transfer Learning

Manpreet Oberoi, Harsh Panchal, Yash Jain
Shah and Anchor Kutchhi Engineering College (BE-IT)

Abstract:- Number of road accidents is continuously increasing in last few years worldwide. As per the survey of National Highway Traffic Safety Administrator, nearly one in five motor vehicle crashes are caused by distracted driver. We attempt to develop an accurate and robust system for detecting distracted drivers. In this paper, we present a CNN based system that detects the distracted driver..VGG-16 architecture is modified for this particular task and various regularization techniques are implied in order to improve the performance. Experimental results show that our system achieves an accuracy of 82.5% and processes 240 images per second on GPU.

1. INTRODUCTION

According to the World Health Organization (WHO) survey, 1.3 million people worldwide die in traffic accidents each year, making them the eighth leading cause of death and an additional 20-50 millions are injured/ disabled. As per the report of National Crime Research Bureau (NCRB), Govt. of India, Indian roads account for the highest fatalities in the world. There has been a continuous increase in road crash deaths in India since 2006. The report also states that the total number of deaths have risen to 1.46 lakhs in 2015 and driver error is the most common cause behind these traffic accidents.

National Highway Traffic Safety Administrator of United States (NTHSA) describes distracted driving as "any activity that diverts attention of the driver from the task of driving." As per the definitions of Center for Disease Control and Prevention (CDC), cognitive distraction is basically "driver's mind is off the driving". In other words, even though the driver is in safe driving posture, he is mentally distracted from the task of driving. He might be lost in thoughts, daydreaming etc. Distraction because of inattention, sleepiness, fatigue or drowsiness falls into visual distraction class where "drivers's eyes are off the road". Manual distractions are concerned with various activities where "driver's hands are off the wheel". Such distractions include talking or texting using mobile phones, eating and drinking, talking to passengers in the vehicle, adjusting the radio, makeup etc.

Nowadays, Advanced Driver Assistance Systems (ADAS) are being developed to prevent accidents by offering technologies that alert the driver to potential problems and to keep the car's driver and occupants safe if an accident does occur. But even today's latest autonomous vehicles require the driver to be attentive and ready to take the control of the wheel back in case of emergency. Tesla autopilot's crash with the white truck-trailor in Williston, Florida in May 2016 was the first fatal crash in testing of autonomous vehicle. Recently in March 2018, Uber's self driving car with an emergency backup driver behind the wheel struck and killed a pedestrian in Arizona. In both of these fatalities, the safety driver could have avoided the crashes but evidences reveal that he was clearly distracted. This makes detection of distracted driver an essential part of the self driving cars as well.

In this paper, we focus on detecting manual distractions where driver is engaged in other activities than safe driving and also identify the cause of distraction. We present a Convolutional Neural Network based approach for this problem.

2. RELATED WORK

This section summarises review of some of the relevant and significant work from literature for distracted driver detection. Reference [1] provides a solution which consists of genetically weighted ensemble of convolutional neural networks. They have trained the CNN on raw images, skin segmented faces , face images, hand images and face+hand images. On these five image sources, they have trained and benchmarked an AlexNet network, an Inception V3 network, a ResNet networkhaving 50 layers and a VGG-16 network. They fine-tuned a pre-trained imageNet model for those networks. Then they evaluated a weighted sum of all the networks output yielding the final class distribution using a genetic algorithm. They obtained an accuracy of 76% using the VGG-16 network, an accuracy of 81% on the ResNet network and 90% on the Inception V3 network.

Reference[2] provides a solution for drowsiness detection. They have taken 3 parameters for the drowsiness detection and have then subdivided each parameters. The 3 parameters are behavioural approach, vehicular approach and psychological approach. In behavioural approach, they try and detect the fatigue of the driver usng various parameters such as eye closure ratio, eye blinking, head position, facial expressions, and yawning. They use the Percentage of eye Closures ratio and based on the result of the ratio, the eyes are refered to as open or closed. Yawning based detection systems analyze the variations in the geometric shape of the mouth of drowsy driver such as wider opening of mouth, lip position, etc. The different vehicular based techniques used were lane detection, eye blinking period, Steering wheel angle, steering wheel behaviour. A single angle sensor was placed under the steering of the car used for detecting the drivers steering behaviors. The Physiological parameters-based techniques detected drowsiness based on drivers' physical conditions such as heart rate, pulse rate, breathing rate, respiratory rate and body temperature, etc. These biological parameters are more reliable and accurate in drowsiness detection as they

are concerned with what is happening with driver physically. The researcher used 3 different classification techniques namely support vector machine(SVM) which provided an accuracy of 98.4%, another method used was Hidden Markov Model and the last method used was the Convolutional neural network(CNN) providing an accuracy of 98%.

StateFarm's Distracted Driver Detection competition on Kaggle was the frst publicly available dataset for posture classifcation. In the competition.StateFarm defned ten postures to be detected: safe driving, texting using right hand, talking on the phone using right hand, texting using lef hand, talking on the phone using lef hand, operating the radio, drinking, reaching behind, doing hair and makeup,

and talking to passenger. Our work, in this paper, is mainly inspired by StateFarm's Distracted Driver's competition. While the usage of StateFarm's dataset is limited to the purposes of the competition, we designed a similar dataset that follows the same postures.

In 2017, Abouelnaga [4] created a new dataset similar to StateFarm's dataset for distracted driver detection. Authors preprocessed the images by applying skin, face and hand segmentation and proposed the solution using weighted ensemble of five different Convolutional Neural Networks. The system achieved good classification accuracy but is computationally too complex to be real time which is utmost important in autonomous driving

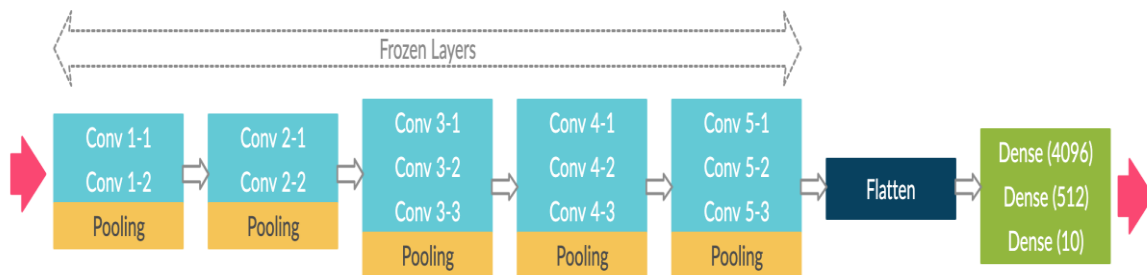Figure 1: 10 Classes of driver postures from the dataset
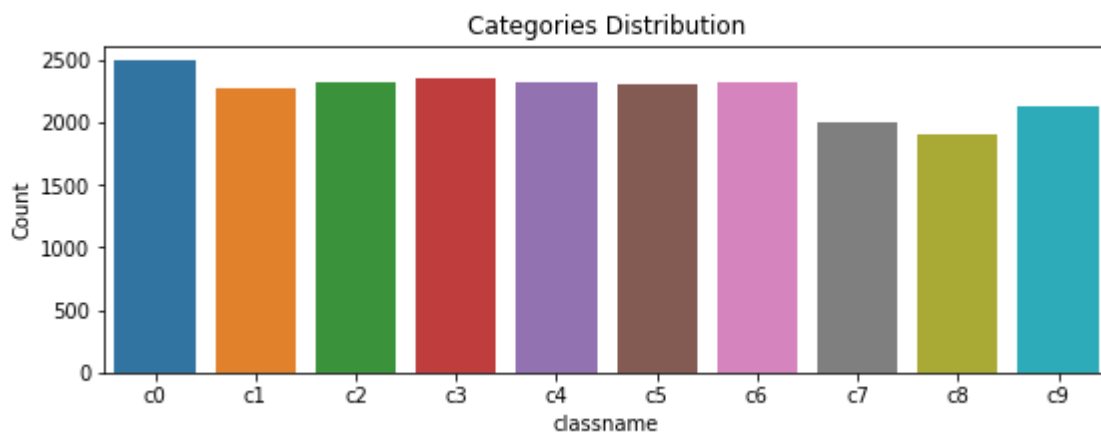

Figure 2: Model Architecture

Figure 3: Classification of the training data

3. DATASET DESCRIPTION

In this paper, we use the dataset that was taken from kaggle.The dataset consists of ten classes viz. safe driving, texting on mobile phones using right or left hand, talking on mobile phones using right or left hand, adjusting radio, eating or drinking, hair and makeup, reaching behind and talking to passenger. Sample images of each class from the dataset are shown in fig. 1. The dataset consists of a total of 102150 images. Out of these, 17939 images are taken for training the model. The training dataset is divided into 10 different classes according to the categories of the distracted driver and each class consists of almost 2200 images each. A total of 4485 images are taken for validation and 79726 images are taken for testing.
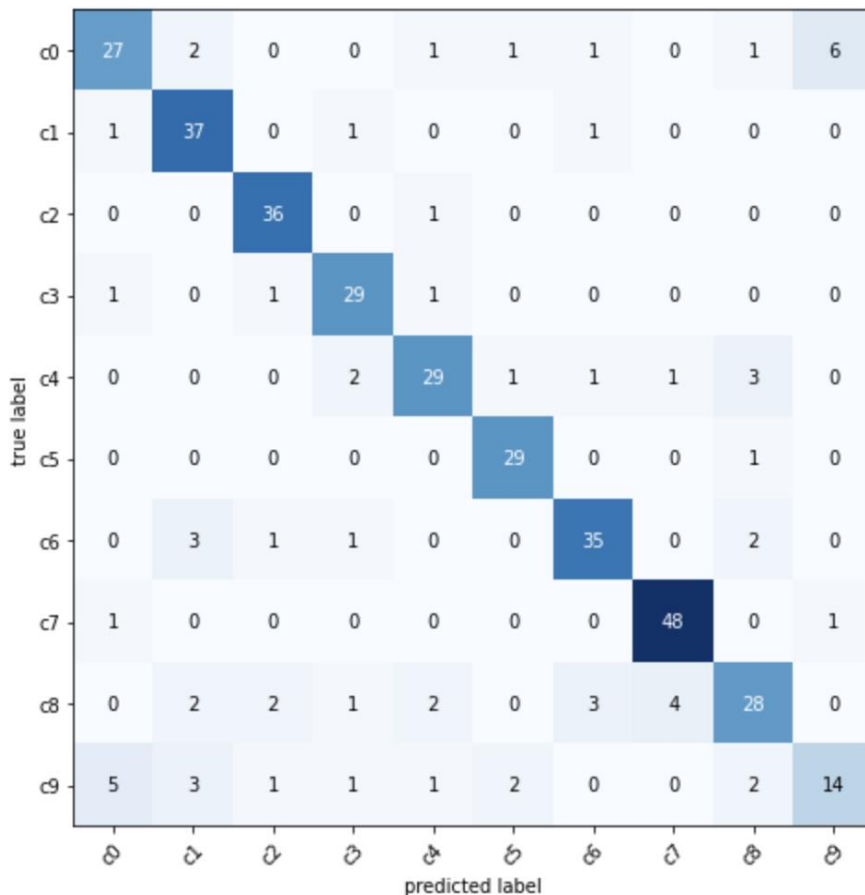
4. PROPOSED METHODOLOGY

Since last few years, CNNs have shown impressive progress in various tasks like image classification, object detection, action recognition, natural language processing and many more. The basic building blocks of a CNN based system include Convolutional filters/ layers, Activation functions, Pooling layer and Fully Connected (FC) layer. A CNN is basically formed by stacking these layers one after the other. Various architectures like AlexNet, ZFNet, VGGNet, GoogLeNet, ResNet have established benchmarks in computer vision. In this paper, for our project, we are using the VGG-16 architecture and modifying it for the task of distracted driver detection.

The VGG-16 model that we are using has 13 convolutional layers and there are 3x3 filters in all the 13 convolutional layers, along with the ReLU Activation function, 2x2 max pooling with stride-2 and categorical cross-entropy loss function. We have used Adam Optimizer in our model. The fully connected layers (FC) which have 4096 channels each have been replaced by convolutional layers having 512 channels each. The modified network

a preprocessing step, all the images are resized to 224×224 and per channel mean of RGB planes is subtracted from each pixel of the image. Initial layers of the CNN act as feature extractor and the last layer is softmax classifier which classifies the images into one of the predefined categories. However the original model has 1000 output channels corresponding to 1000 object classes of ImageNet. Hence the last layer is popped and is replaced with dense layer with 10 classes corresponding to the 10 classes present in our dataset. Here, the cross entropy loss function is used for performance evaluation.

```
(<Figure size 576x576 with 1 Axes>,
 <matplotlib.axes._subplots.AxesSubplot at 0x7fc742ef7550>)
```

IJERTV9IS050862
www.ijert.org
1344
(This work is licensed under a Creative Commons Attribution 4.0 International License.)

```
(<Figure size 576x576 with 1 Axes>,
 <matplotlib.axes._subplots.AxesSubplot at 0x7fc742ef7550>)
```
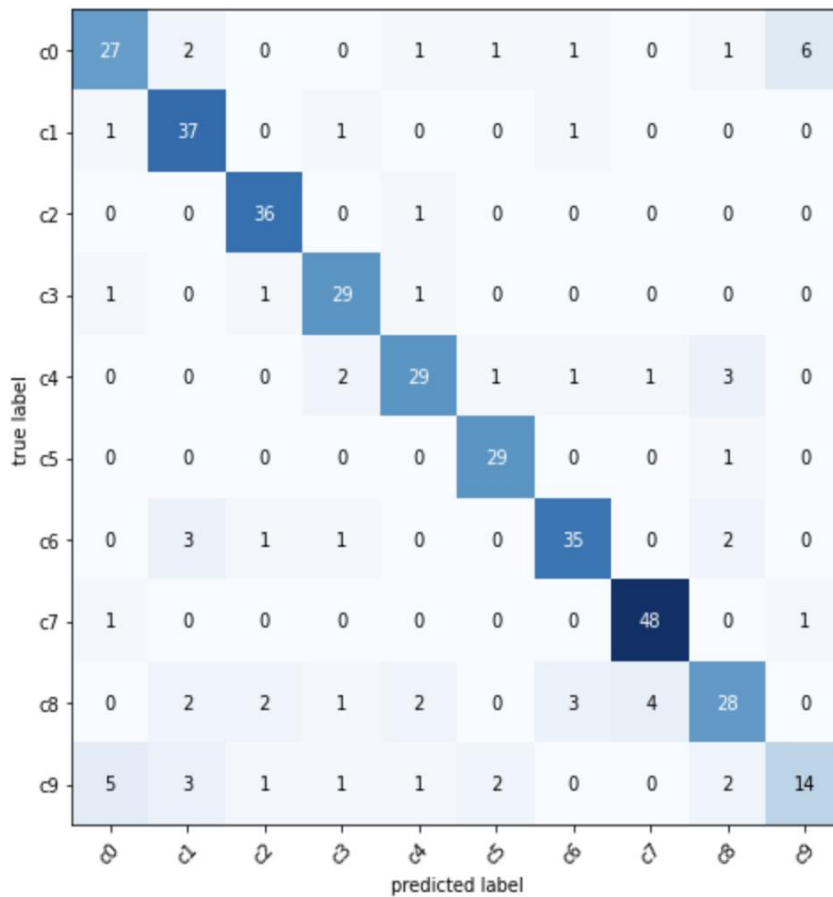

Figure 4: Confusion Matrix for the given model

5. RESULTS AND DISCUSSIONS

We design a Convolutional Neural Network based system for distracted driver detection. The pre-trained ImageNet model is used for weight initialisation and concept of transfer learning is

applied. Weights of all the layers of network are updated with respect to the dataset. The batch size is set to 40 and the number of epochs are set to 30. The dataset consists of a total of 22,420 images. Out of these images,17939 images Are used for training, 4481 images are used for validation and a further 387 images are used for testing.

Table 1 depicts the classwise accuracies for each of the 10 classes from the dataset. Figure 3 shows the confusion matrix for the given model. We have achieved an accuracy of 82.5% for our model

Class	Total Samples	Correct Predictions	Incorrect Predictions	Accuracy
Safe Driving	39	27	12	70
Texting-right hand	40	37	3	92.5
Talking on phone-right hand	37	36	1	97.2
Texting-left hand	32	29	3	90.6
Talking on phone-left hand	37	29	8	78.3
Adjusting Radio	30	29	1	96.6
Drinking	42	35	7	83.3
Reaching Behind	50	48	2	96
Hair & Makeup	42	28	14	66.6
Talking to passenger	29	14	15	48.27

Table 1: Class wise accuracies for the given model

6. CONCLUSION AND FUTURE WORK

Driver distraction is a serious problem leading to large number of road crashes worldwide. Hence detection of distracted driver becomes an essential system component in everyday life today. Here, we present a robust Convolutional Neural Network based system to detect distracted driver. We modify the VGG-16 architecture for this particular task and apply regularization techniques namely the ReLU Activation function and Adam optimizer to prevent overfitting to the training data. Experimental results show that the model predicts the distracted drivers with an accuracy of 82.5%.

As an extension of this work, we can lower the number of parameters and computation time. Incorporating temporal context may help in reducing misclassification errors and thereby increasing the accuracy. Also, in future, we wish to develop a system that will detect visual and cognitive distractions as well along with manual distractions. We can also use other architectures as well as genetic algorithms to futher improve the accuracy.

7. REFERENCES

[1] Centerfor disease control and prevention. https://www.cdc.gov/motorvehiclesafety/distracteddriving/.

[2] National highway traffic safety administration traffic safety facts : https://www.nhtsa.gov/risky-driving/distracted-driving/.

[3] State Farm Distracted Driver Detection :https://www.kaggle.com/c/state-farm-distracted-driver-detection.

[4] Y. Abouelnaga, H. M. Eraqi, and M. N. Moustafa. Real-time distracted driver posture classification. CoRR, abs/1706.09498, 2017.

[5] E. Ohn-Bar, S. Martin, and M. M. Trivedi. Driver hand ac-tivity analysis in naturalistic driving studies: challenges, al-gorithms, and experimental studies. J. Electronic Imaging, 22(4):041119, 2013.

[6] Center for disease control and prevention. https://www.cdc.gov/motorvehiclesafety/distracted_driving/

[7] National highway traffic safety administration traffic safety facts : https://www.nhtsa.gov/risky-driving/distracted-driving

[8] Murtadha D Hssayeni, Sagar Saxena, Raymond Ptucha, Andreas Savakis; Rochester Institute of Technology, Rochester, NY US : Distracted Driver Detection: Deep Learning vs Handcrafted Features https://www.ingentaconnect.com/contentone/ist/ei/2017/00002017/00000010/art00004?crawler=true&mimetype=application/pdf

[9] C. Yan, F. Coenen, and B. Zhang. Driving posture recognition by convolutional neural networks. IET Computer Vision, 10(2) , Pages :103–114, 2016. https://www.researchgate.net/publication/283433254_Driving_posture_recognition_by_convolutional_neural_networks

[10] 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, 6(2) , Pages :161–168, June 2012.https://ieeexplore.ieee.org/document/6205797