Driver Distraction Detection using Neural Network

Patel Raj S[1], Patel Manthan R[2], Prof. Ajaykumar T. Shah[2]
Department of Computer Engineering
[1,2] Alpha College of Engineering & Technology, Gujarat, India
[3]H.O.D, Alpha College of Engineering & Technology, Gujarat, India

Abstract- *Distracted driving is known to be the leading cause of motor vehicle accidents. It is found that nearly one in five motor vehicle accidents were caused by distracted driving. Motivated to reduce these number of accidents, the purpose of this project is to accurately classify what drivers are doing and whether they are distracted or not and to alert them if they are distracted. This system leverages state of the art models and techniques of Deep Learning such as Convolutional Neural Network and Recurrent Neural Network. The input to the models are images of people driving. Each image belongs to one of the few classes of driver distraction.*

Keywords- Driver Distraction, Road safety, Deep learning, Convolutional Neural Network, Transfer learning.

I. INTRODUCTION

The problem of driver distraction is getting worse because of increasing use and dependence on the technology. Not only more people are now using vehicles, but also distractions like cell phones, smart phones, GPS system within them, feature loaded car dashboards etc. Nowadays car comes with interactive touch screens-based navigation systems and this requires interaction to operate. Smartphone notifications also play big role in distraction of the driver. Distracted driving, wherein a driver is engaged in another activity that takes their attention away from operating the vehicle, is identified to be a significant cause of fatal motor vehicle crashes and injuries. According to IndiaTimes report of 2017 nearly 2100 deaths across India last year happened because of phone usage while driving vehicle.

The WHO categorizes driver distraction as an important risk factor for road crash injuries. The United States Department of Transportation terms distracted driving as one of the most dangerous driver behaviours and an epidemic which has increased with the proliferation of mobile phones. Distracted driving is defined as any activity that diverts a person's attention from their primary task of driving. These types of activities include the usage of a mobile phone, eating and drinking, conversation with co-passengers, self-grooming, reading or watching videos, adjusting the radio or music player and even using a GPS system for navigating locations. Amongst these, mobile phone usage is said to be the most distracting factor. Every hour, 16 lives are lost to road crashes in India. In the last decade alone, India lost 1.3 million people to road crashes and another 5.3 million were disabled for life.

II. LITERATURE REVIEW

In the last decade alone, India lost 1.3 million people to road crashes and another 5.3 million have been seriously injured. India has the highest number of road crash fatalities, with a crash occurring every minute and one death every four minutes.

According to the Telecom Regulatory Authority of India (TRAI), India currently has 1 billion mobile phone subscribers. Due to the high penetration of mobile phones, their use on the road has also increased as proven by this study with 1 out of 2 respondents of this survey having used a phone while driving.

Using the image dataset obtained from kaggle.com we trained a convolutional neural network that can predict the driver's distraction while driving.

From various sources, we found that image data fed to the convolutional neural network gives better results than regular neural network.

III. SYSTEM OVERVIEW

First, we collected dataset of images of drivers from the Kaggle.com which is provided by State Farm. The figure 1 is example of some of the images of dataset.

Images are of resolution 640 X 480 pixels. Images of this high resolution are not suitable for the neural network both for training and testing purposes. So, in the training process we are going to reduce size of image. Dataset is a mixture of both male and the female drivers, therefore model can learn better.

Images divided into 10 classes of distraction while driving like texting on right hand, calling on left hand, talking with passengers, reaching behind, eating etc.

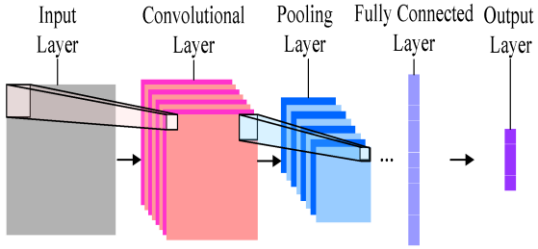
Steps of our system:

- Split dataset in TrainSet, DevSet, TestSet.
- Train a neural network
- Tune Hyper Parameters on DevSet
- Find accuracy sing TestSet.
- Predict given image

IV. NETWORK ARCHITECTURE

Simple Neural Network: In our first attempt to solve this challenge we tried to use a regular neural network. For that we reduced the image to 32X32 pixels greyscale, because of performance constraints of neural network we cannot feed a 640X480 pixel image. With this architecture our input layer was 32X32=1024 neurons, which is quite large number of units. Moreover, a 32X32 image lost a lot of features and information of image. The model was not giving more than 50% accuracy on test set. Which is not good for a system, So we decided to change the network.

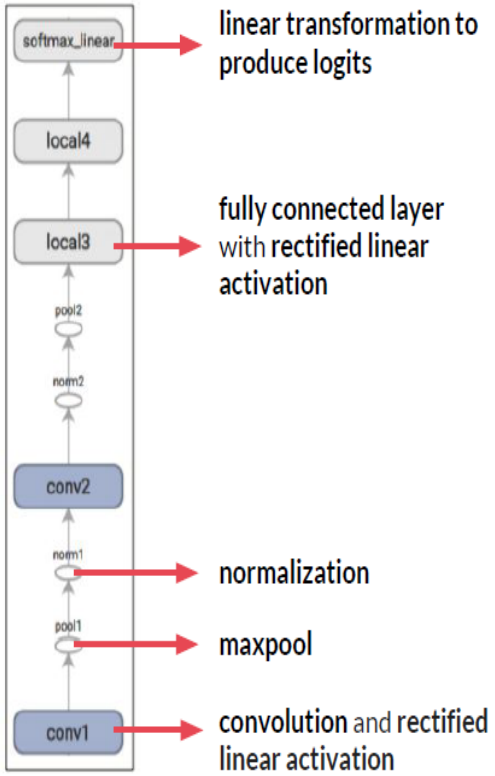
Convolutional Neural Network: CNN works better than regular neural networks on image data. For this network we reduced the image size only to 150X150 pixels RGB image. Which holds much more information of original image than the previous one. CNN consists of mainly Convolution operation, Pooling operation and fully connected layer. Different architecture of CNNs can have many of these operations. For the activation function we used ReLU(Rectified Linear Unit) and for the output layer we used the softmax function.



V. TRAINING THE NETWORK

Neural network requires a loss function to measure how well and how close the predicted value is to ground truth labels or class. We have used the cross-entropy log loss function to measure loss. Which is given below.

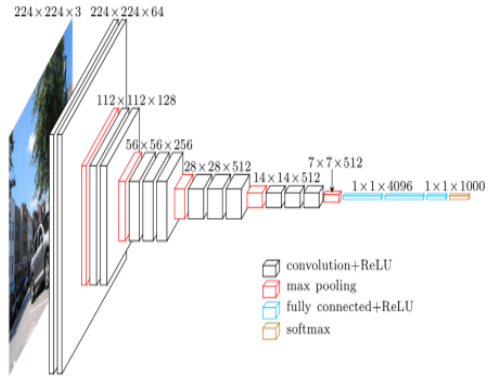
$$L = -\sum_{j=1}^{K}[t^j log(p^j) + (1 - t^j) log(1 - p^j)]$$



For training the network we used the back-propagation method. Back propagation is used to compute the gradients and to update weight matrix to learn the features from the images. For the efficiency, we had to store activations calculated in the forward pass in the cache. For sanity check of Back propagation implementation, we can use numerical gradient computation and compare those gradients with the gradients calculated by the back propagation. If the difference is of order 10e-7 we can ensure that implementation is correct.

Transfer Learning: For further improvement in the accuracy we can use the concept of transfer learning in this task. In transfer learning we use the pretrained model and update/train last few layers on our dataset. Some example of pretrained networks are VGG16, VGG19 and Inception Network. These networks are trained on the data of millions of images.

VGG16 network is characterized by its simplicity, using only 3×3 convolutional layers stacked on top of each other in increasing depth. Reducing volume size is handled by max pooling. Two fully-connected layers, each with 4,096 nodes are then followed by a SoftMax classifier. The figure below describes VGG16 network architecture. Model is trained on ~1.2 million training images with another 50,000 images for validation and 100,000 images for testing of ImageNet challenge.



Testing: For testing the model TestSet images are used. If the predictions are wrong, we can make a confusion matrix that can show us which two classes of output are overlapping and we can train our model to further improve for those classes.

VI. FUTURE ENHANCEMENT

We aim to build a system that can reduce road accidents by preventing and alerting users from being distracted. This project is our first step towards this goal. We have found that we can add numerous features in this system to help in safe driving. Some of the features that we have identified are as follows. First, we can make the system to work in night mode also. Right now, images taken from night are not getting predicted correctly. Second, the system is running on low fps, we are aiming to make it work in higher fps. Moreover, As the neural networks are computationally expensive we are planning to build a server-based classification system in which the user device just have to send the image data captured from the camera and send it by API calls no processing of data is required on user side

VII. CONCLUSTION

This project is intended to increase road safety by alerting a driver while driving vehicles. As the smartphones and gadgets are essential part of one's day to day life, it is hard to live without them. This work has looked at solving the detection of distracted drivers through images obtained from the State Farm Distracted Driver Detection. The model trained from the training data using convolutional neural network. The model can classify the given input image into ten classes of distraction (including safe driving position). Overall, the system has been proven to be effective and we hope that it reduces the injuries and deaths resulting from distraction while driving.

ACKNOWLEDGMENT

We express our sincere thanks to Prof. Ajaykumar T. Shah Head of Department of Computer Engineering, Alpha College of Engineering and Technology for their Support and guidance for this project and care taken by them in helping us to complete the project work successfully.

REFERENCES

[1] Deep Learning for Computer Vision: Expert techniques to train advanced neural networks using TensorFlow and Keras.
[2] SaveLife foundation survey http://savelifefoundation.org/wpcontent/uploads/2017/04/Distracted-Driving-in-India_A-Study-on-Mobile-Phone-UsagePattern-and-Behaviour.pdf.
[3] Statistics on road accidents because of smartphone by World Health Organization https://www.who.int/gho/road_safety/mortality/en/
[4] TensorFlow Official Website https://www.tensorflow.org/
[5] Keras Official Website https://keras.io/
[6] Kaggle Inc. State farm distracted driver detection dataset.
[7] PY Simard, D Steinkraus, JC Platt, 'Best Practices for convolutional neural networks applied to visual document analysis', 2003

[8] ImageNet: VGGNet, ResNet, Inception, and Xception with Keras (Tutorial for implementing transfer learning in keras by pyimagesearch): https://www.pyimagesearch.com/2017/03/20/imagenet-vggnet-resnet-inception-xception-keras/
[9] VGG16 network implementation in tensorflow: https://www.cs.toronto.edu/~frossard/post/vgg16/
[10] ImageNet challenge http://image-net.org/