Detecting Distracted Driving using Computer Vision

[1]Rohit Misra
[1]Student
[1]The International School Bangalore

Abstract **- The objective of this study is to investigate the usage of Artificial Intelligence and Computer Vision in the detection of distracted driving. The concept of supervised machine learning was utilized using Convolutional Neural Networks to develop an image classification model which analyzed and categorized images into 10 classes. Our algorithm accurately identifies the activity being performed in the image. A 5-layer convolutional neural network model was trained using a dataset of 22424 unique images. The model achieved an 81.5% accuracy.**

keywords **- Artificial Intelligence, Computer Vision, Neural Networks, Supervised learning, Image Processing.**

I. INTRODUCTION

In today's world, urban vehicular transportation is seeing drastic shifts to automation with technology interfacing with transportation. The increasing reliance on technology in automobiles raises demands for safer mobility. Automobile companies have adopted different approaches to introducing safer mobility, including improvements in vehicular safety standards, self-driving, and driver assistance. Leading automobile manufacturers such as Tesla Inc.[1]. are working to develop new methods making use of cameras and other sensors to detect factors, such as driver distraction, and drowsiness in order to alert the driver of the same and thereby reduce the likelihood of accidents related to careless driving. The onboard computers in such vehicles utilize Computer Vision Models to process the data received to provide accurate alerts.

Vehicles with such capabilities have successfully penetrated developed markets such as the USA, China, and the EU, due to the relative affordability of such automobiles. Unfortunately, India has not witnessed the mass adoption of such technology due to a lack of domestic innovation and unaffordability.

According to the Road Transport and Highways Ministry report on 'Road Accidents in India, 2018', there were 151,417 deaths due to road accidents in 2018. Most of these deaths were reported to be because of driver errors including over speeding, looking at the phone while driving, drowsiness, talking to the passenger while driving, texting while driving etc. [2] Due to a legacy of poor road conditions, poor driving standards and traffic, India is responsible for the highest number of road deaths in the world [3].

A United Nations study found that India's GDP takes a 3% hit every year due to road accidents, equivalent to over $58 billion in value terms [4]. The need of the hour is to mandate more stringent driving standards, potentially through the introduction of affordable technology.

This paper aims to explore driver assistance in the form of detection of Distracted Driving using Computer Vision in order to prevent accidents and save hundreds of thousands of lives.

II. APPROACH

The focus of the research is to develop an accurate Distracted Driving Detection model. To be effective at reducing traffic violations and preventing accidents, it is important that a distracted driving model serves as a pre-emptive and preventive solution. Hence, it needs to be able to accurately detect and classify a multitude of distractions. However, having too many classes reduces the speed of the model in a live scenario, and thus its effectiveness is affected.

A two-pronged approach is utilized to tackle this problem. First, an accurate image classification model was developed using Convolutional Neural Networks. Second, the model was fed frames at 1 second intervals from a live webcam feed to provide live classifications and appropriate alerts if required.

III. CONCEPTS

A. Supervised learning

Supervised learning is a machine learning technique which is based on training a model using the provided labelled data. The model is trained using training datasets until it can detect the relationships between data belonging to each specific class, yielding an acceptable accuracy. Once the model has been trained, it is used to classify or identify unlabelled data which are known as the testing datasets.

B. Convolutional Neural Networks

Artificial neural networks are a widely used classification technique, wherein the requirement to make assumptions and recognise relationships about the given data is given to layers of nodes that solve subproblems of the given input. The process replicates the way the human brain functions, utilizing nodes or 'neurons' which make up the hidden layers mentioned above. To each of its incoming connections, a node will assign a number known as a "weight." When the network is active, the node

receives a different data item — a different number — over each of its connections and multiplies it by the associated weight. It then adds the resulting products together, yielding a single number. If that number is below a threshold value, the node passes no data to the next layer. If the number exceeds the threshold value, the node "fires," which in today's neural nets generally means sending the number — the sum of the weighted inputs — along all its outgoing connections.[5]

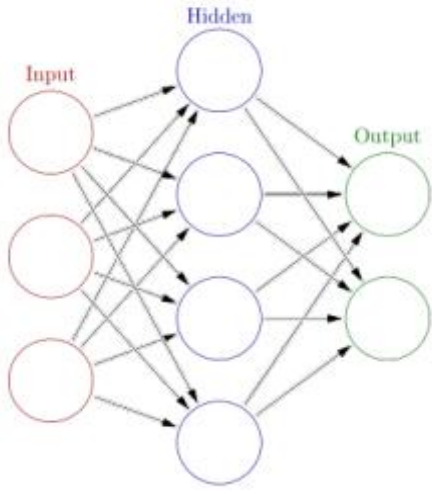

Figure 1 : A neural network model

In image processing, Convolutional Neural Networks models are used to train and test each input image. The CNN will pass the input images through a series of convolution layers with filters (kernels), pooling, fully connected layers (FC) and apply SoftMax function to classify an object with probabilistic values between 0 and 1. The below figure is a complete flow of CNN to process an input image and classifies the objects based on values. [6]

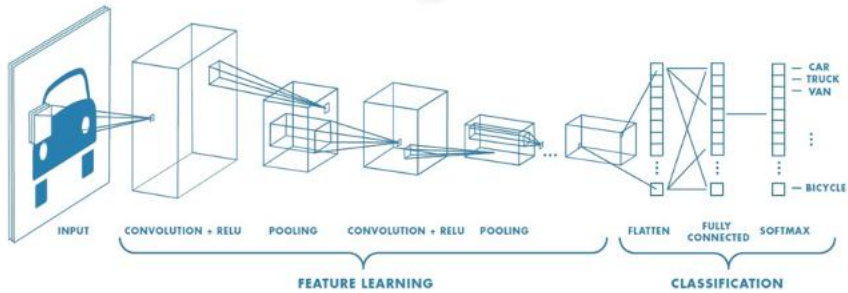

Figure 2 : A neural network model

IV. METHOD

The data set was obtained from "Kaggle" provided by State Farm. It consists of 2D dashboard camera images of drivers either driving responsibly or engaging in distracted behaviors.

Figure 3 : A sample image from each of the 10 classes provided by the dataset. From top left to bottom right : safe driving, texting with the right hand, talking on the phone with the right hand, talking on the phone with the right hand, texting with the left hand, talking on the phone with the left hand, operating the radio, drinking while driving, reaching behind, doing hair or makeup and talking to the passenger.

For this investigation, the concept of supervised machine learning was utilized using Convolutional Neural Networks. The training dataset was divided into classes which consisted of labelled images based on the event occurring in the image. The classes labels were:

'safe_driving', 'texting_right', 'talking_on_phone_right', 'texting_left', 'talking_on_phone_left', 'operating_radio', 'drinking', 'reaching_behind', 'doing_hair_makeup', 'talking_to_passanger'.

The model consists of a three-layer convolutional neural network which outputs 3D feature maps. Following this, utilizing a flatten function, the 3D feature maps were converted to 1D feature vectors. A dense layer of 64 units is added, passed through a rectified linear activation function. A dropout regularization at a rate of 0.5 was then added to reduce overfitting. Finally, a dense layer of 10 units is added and passed through a sigmoid activation function.

V. RESULTS

We noticed that thanks to the excellent database, we were able to accurately identify most distracted driving interactions. An 81.56% classification accuracy was observed on our testing database, and a 93.24% accuracy was observed on our validation database.

Figure 4 : Woman texting using her left hand

As seen above, it is clear that the model has accurately identified that the woman in the image is texting using her left hand.

Figure 5: Man drinking while driving

The man in the image has been correctly identified as 'drinking' and hence this is also unsafe driving.

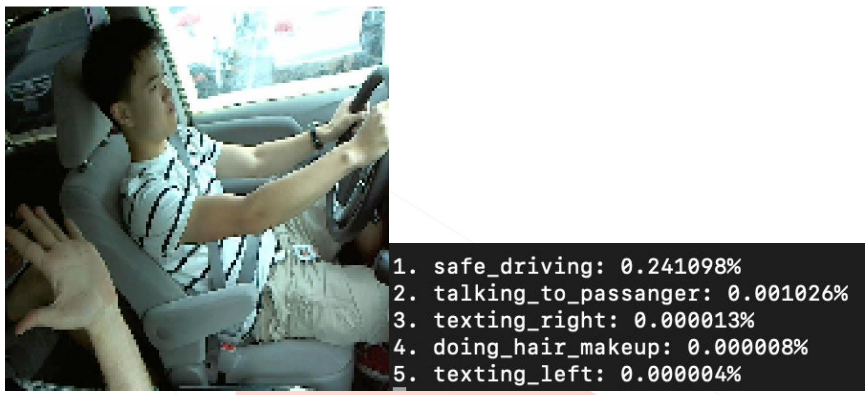

Figure 6: An image classified as safe driving

Similarly, safe driving has been identified with a high degree of certainty in the image above.

VI. Further Exploration

It is evident from the results above that our model is notably efficient in achieving high degrees of accuracy. Our model can be the first line in verification of drivers in India. We imagine a system where, when a certain threshold of distracted driving is crossed, the driver's license is revoked for a certain period of time. For example, a driver could routinely be speaking on the phone while driving. After the 5th violation in a day, the driver gets an official warning. And should the violations continue, the driver's license is revoked for a month. Our system stores the frames, where distracted driving is observed. These frames can be verified by a human reviewer, to ensure fairness. The frames can also be shared with drivers to provide them with a way to improve their driving.

We imagine our system could also be used to reform India's broken driving license distribution strategy. Drivers could be given a temporary driving license for a month. During the month, should the driver engage in continuous driving violations, the license can be revoked, and safe driving can be encouraged.

With the rise of Ola and Uber, and other ride hailing apps, road accidents in India have dramatically increased. Ola and Uber drivers normally drive for more than 12 hours a day, leading to drowsiness and distracted driving. The driver's dashcam can be used to monitor the driver's performance and both the rider and the company can be alerted if the driver is driving rashly.

It has also been acknowledged that the dataset solely consists of images taken during the daytime. However, many accidents and instances of distracted driving occur during the night wherein in-car visibility of the driver is limited. To bypass this issue, training a separate model consisting of a dataset with thermal images is hypothesized to produce accurate results when lighting is minimal.

We also believe that we can increase the accuracy of our model by adjusting the number of layers, the number of nodes per layer, and the weights.

References

[1] Tesla Inc. 2020. *Autopilot*. Retrieved 27 October 2020, from: https://www.tesla.com/autopilot.

[2] Ministry of Road Transport & Highways, Government of India, (2018). *Road Accidents in India*. Ministry of Road Transport & Highways, p.3.

[3] Athrady A. (2018). *India tops world in deaths due to road mishaps: WHO*. Deccan Herald. Retrieved 27 October 2020, from https://www.deccanherald.com/national/india-tops-world-deaths-due-707066.html.

[4] PTI (2016). *India loses $58 billion annually due to road accidents: UN study*. Mint. Retrieved 27 October 2020, from https://www.livemint.com/Politics/F9ljlqoWYdxxgJZ4razuiI/India-loses-58-billion-annually-due-to-road-accidents-UN-s.html.

[5] Hardesty L. (2017). *Explained: Neural networks*. MIT News. Massachusetts Institute of Technology. Retrieved 27 October 2020, from https://news.mit.edu/2017/explained-neural-networks-deep-learning-0414.

[6] Prabhu, R. (2020). *Understanding of Convolutional Neural Network (CNN) — Deep Learning*. Medium. Retrieved 27 October 2020, from https://medium.com/@RaghavPrabhu/understanding-of-convolutional-neural-network-cnn-deep-learning-99760835f148.