INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact factor: 4.295
(Volume 4, Issue 2)
Available online at: www.ijariit.com

Distracted Driver Detection using Convolutional Neural Network

Syed Thoufich
thoufich@gmail.com
Kammavari Sangha Institute of Technology, Bengaluru, Karnataka

Nikhil V Anand
nikhilvanand18@gmail.com
Kammavari Sangha Institute of Technology, Bengaluru, Karnataka

Rajaganapathi S
rajaganapathi96@gmail.com
Kammavari Sangha Institute of Technology, Bengaluru, Karnataka

Roopesh Kumar B N
roopeshkumarbn@gmail.com
Kammavari Sangha Institute of Technology, Bengaluru, Karnataka

ABSTRACT

In order to help in reducing possible accidents caused due to a distraction while driving. A model is proposed based on the convolution neural network. The Convolution Neural Network is commonly used the model in deep learning. Among various types of networks, convolution neural network resulted in high performance on image classification. Here we build and train the model with a dataset of dashboard camera images of the driver, which detects the pre-defined set of actions of the driver engaging in an unsound behavior. And provides a suitable notification indicating the driver of his/her hindrance.

Keywords: *Deep Learning, Convolution Neural Network, Image Classification, Distraction.*

1. INTRODUCTION

Feature recognition plays an important role in determining the actions of a driver in an automobile. This can be achieved by combining image classification, pose estimation, visual saliency detection concepts. The images obtained as input must be pixel processed swiftly to obtain immediate and accurate prediction percentages. The CNN deep learning concepts help in recognition and estimation process and provide results in a time constraint, suitable for real-life applications like human movement detection, object interference and so on. We strive to build a model called the D3 model for safer roads, which will help remind the driver to keep in an attentive state while driving by notifying his/her negligence while steering. The input for our model is the pictures of the drivers who are driving if in case a situation of the driver being distracted occurs the model is trained to analyze the situations like using the phone, back turned to road and so on and notify the driver with a suitable message. According to the Times Of India in India alone the road accidents caused in 2017 was 1,46,377, 47% of this staggering count was due to accidents caused by a distracted driver.

2. EXISTING SYSTEM

Bag of Words One of the two feature extraction technique used is BoW to extract features from each set of images. The bow can be defined as the "histogram representation based on independent features. BoW represents an image in three steps: feature detection, feature description, and codebook generation. The first step is to extract features from all the images. This is done by creating an image data store object which holds all the images along with their labels [3].

3. METHODOLOGY

Deep learning is a branch of machine learning, which is one of the major breakthroughs and research hotspots in machine learning in recent years. Image classification techniques fall mainly into two broad categories: unsupervised and supervised image classification techniques. In an unsupervised learning model, the system is first presented with a large amount of 'unlabeled' images. The system then builds a probabilistic model from the unlabeled images by finding patterns in them. On the other hand, in a supervised learning model, the system is first trained with multiple examples of images. A learning model is generated using the

training data. The model is then used to predict the features of an unknown image. Convolution neural network is a multi-layer artificial neural network specially designed to handle two-dimensional input data. [1] Each layer in the network is composed of multiple two-dimensional planes, and each plane consists of multiple independent neurons Composition, adjacent two layers of neurons connected to each other, and in the same layer of neurons are not connected between. CNN's are inspired by the early time-delay neural networks and TDNNs. The other feature extraction technique used is CNN features extraction. CNN has two main characteristics of CNN - it uses weight sharing and local receptive field. Ideally, CNN has three main types of layers convolution layer, pooling layer, and fully-connected layer. The convolution layers main purpose is to extract features and the pooling layer is responsible for feature mapping. The fully-connected layer is learning a function from the features extracted from the convolution layer. The convolutional layer is the core part of the Convolutional neural network, which has local connections and weights of shared characteristics. The aim of the Convolutional layer is to learn feature representations of the inputs. As shown in above.

A. Convolution layer

This is the main component phase of CNN and the main phase feature is to extract a feature from the given input image (features like sharp/straight edges, simple curves). A filter is the pixel representation of a feature which will be glided over the input image given from the top left corner to the bottom right corner in order to detect the image present in the image. The depth of the filter should be the same of that of the input image. In the Convolutional layer, there exist several feature maps. Each neuron present in the existing feature map is taken to extract the local characteristics of dissimilar positions in the previous layer, but for individual neurons, extraction is local characteristics of similar positions of former the different feature map. In constraint to obtain a brand new feature, the input feature maps are first coiled with a learned kernel and later the gained data is passed onto a non-directed activation function. The Two-dimensional convolution calculation is shown in the below Fig. 1. For example, for a raw input image with a resolution of 128x128, it is considered that 200 8 x 8 size feature particles of the picture have been gained by training in prior. Later, by taking these 200 feature particles, each 8 x 8 small block area in the authentic picture is convolved, and a feature particle can obtain a convolution feature map of 121x121, then at the end, the complete picture can be gained as 200x121x121 [2] convolution feature map:

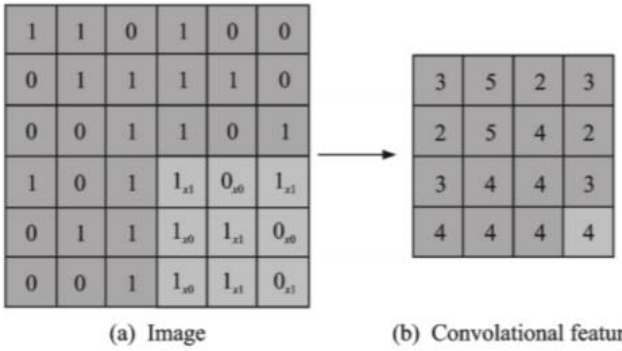

Fig. 1 Representation of two-dimensional convolution operation

B. ReLU (Rectified Linear Units) layer

After the completion of the process in the convolution layer, nonlinear or activation layer is applied immediately. The cause of this layer is to introduce nonlinearity into the system. In previous, functions like tanh and sigmoid were utilized. But after the better understanding by taking survey results ReLU is used because the training of the model became more swift and quick and it also helped in escalating and vanquishing the gradient problem where the hindrance in the lower layers was lessened exponentially.

C. Pooling layer

Completing ReLU layers, coders apply pooling layer. There are many layer options and among these max-pooling is the most popular which is shown in Fig. 2. This considers a filter in the default size of 2x2 and a strand of the same length. Later, applies the gained input value and outputs the maximum in every sub parts that the filter convolves around. Pooling layer exceedingly decreases the space dimension of input image volume. It makes the input representation smaller and manageable by reducing the number of parameters and computations. This phase resolves to overfit.

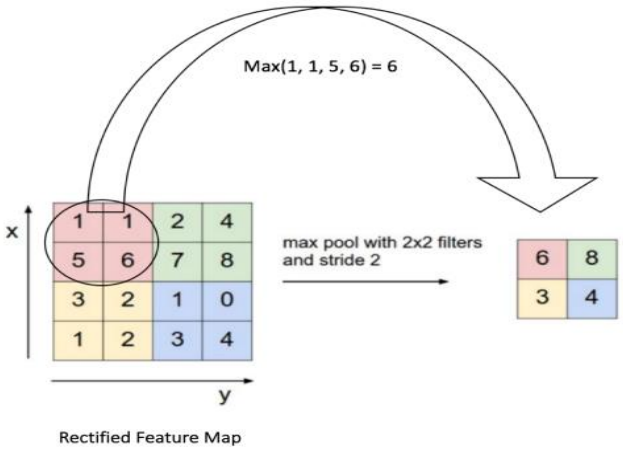

Fig. 2 Representation of Max pooling layer

D. Fully Connected layer

The term "Fully Connected" implies that every neuron in the previous layer is connected to every neuron on the next layer [4]. The output from the convolution and pooling layers represent high-level features of the input image. The purpose of the Fully Connected layer is to use these features for classifying the input image into various classes based on the training dataset. For example, if there is a recognized image has a feature prediction 0.1, 0.75, 0 the model shows the response output of the 0.1 since there is a 10% chance of being the determined feature which was the highest prediction value obtained when processed. The output will be the correct probabilities for different cases.

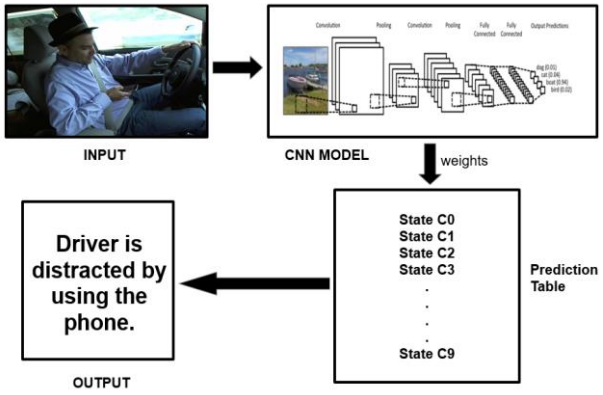

Fig. 3 Design of two-dimensional convolution operation

Considering the above-explained accurate CNN concepts we have come up with a model shown in Fig. 3 which is trained for 10 cases from C0 to C9 [cases being driver using the phone, the driver turning back]. The D3 model is trained to recognize these case features with highly accurate percentage predictions that show cause for driver distraction and front-end system shows the suitable notification to indicate the driver to remain attentive. The images of the drivers are the input to our model and these images are pixel processed through several layers as explained above. And the gained output is a feature detection which is high in percentage and highly reliable.

The data for the prediction is collected from CPBC. The Central Pollution Control Board (CPCB), is an organisation, which was founded in September 1974 under the Water (Prevention and Control of Pollution) Act, 1974. CPCB was handed the functions under the Air Act, 1981(Control and Prevention of Pollution of Air).CPCB has important functions given in Water Act, 1974(Control and Prevention of Pollution of Water), and Air Act, 1981(Control and Prevention of Pollution of Air) to increase the air quality of air and to control, prevent or remove air pollution in the country.

4. CONCLUSION

In this paper, we propose a model which can be trained to recognize the actions of a driver and relay suitable notifications to the driver if he/she is being distracted while driving. The approach predicts a high percentage results indicating features of normal and harm and may help in saving a life in an automobile. The long-term vision of the model is to reduce the staggering statistics and death counts caused by road accidents and bring awareness to the person steering the automobile. To improve road sense and the prevention rate of road accidents caused by the hindrance of drivers by notifying them in prior of possible dangers.

5. REFERENCES

[1] Tianmei Guo, Jiwen Dong Henjian Li˙Yunxing Gao, "Simple Convolutional Neural Network on Image Classification", Computing University of Jinan, Jinan.
[2] Ahmed Ali Mohammed Al-Saffar, Hai Tao, Mohammed Ahmed Talab, "Review of Deep Convolution Neural Network in Image Classification", Faculty of Computer Systems and Software Engineering Universiti Malaysia Pahang Pahang, Malaysia haitao@ump.edu.my.
[3] Abdullah "An Application of pre-Trained CNN for Image Classification", BRAC University Dhaka,Bangladesh and Mohammad S. Hasan School of Computing and Digital Technologies Staffordshire University Stoke-on-Trent, UK.
[4]Saad Albawi Tareq Abed Mohammed,"Understanding of a Convolutional Neural Network" ,Kemerburgaz University Istanbul, Turkey, Saad AL-ZAWI Diyala University Diyala , Iraq.

