Automatic Accident Detection Using Machine Learning

Suraj Pundlik Bhujbal[1] , Manish Kailas Gaikwad[2] , Yash Rajaram Nagare[3]

Department of Computer Engineering, Zeal College of Engineering & Research, Pune, Maharashtra, India.

---***---

Abstract: **In this fast-paced world, the number of deaths due to accident is growing at an expeditious rate. Major reasons for these accidents are rash driving, drowsiness, drunken driving, carelessness, etc. Drowsiness is one of the underlying causes of driving accidents, which contribute to many road fatalities annually. The number of fatal and disabling road accidents are increasing day by day and is a real public health challenge. Many times, in the road accidents, human lives will be lost due to delayed medical assistance. Drowsiness is one of the underlying causes of driving accidents, which contribute to many road fatalities annually. This paper focuses on designing web application-based accident detection system and avoidance system which detect accidents using image processing and inform the location of accident to nearby hospital using text SMS.**

Keywords- *Accident, Death, Drowsiness, Image Processing.*

---***---

I. INTRODUCTION

In India, more than 150,000 people die due to traffic accidents that means, about 400 fatalities a day! Surveys have been conducted and found that the source of majority deaths across the world is due to road accidents. Approximately 50 – 60 percent of the delays on urban freeways are associated with incidents, and on urban surface streets, a large percentage of traffic accidents and most delays occur at or near intersections [1]. In India 377 people die every day due to road accident which is four times more than the annual death toll from terrorism [2]. World Health Organization conducted a survey on different causes of death due to injury [3]. 93 percent of the world's fatalities on the roads occur in low- and middle-income countries, even though these countries have approximately 60 percent of the world's vehicles [4]. Most of the deaths arise due to the lack of immediate medical care provided to the victim at the time of an accident. Traffic hazard is one of the major issues to be dealt with, when it comes to transportation. Studies have shown that approximately four deaths are happening in every minute due to road accidents. Deaths occur mostly during the first hour of impact. Delay in reporting the accidents immediately to the control room is a serious issue. A lot of delay arises in each and every stage due to the human element involved [5]. Our government had implemented many strict rules to reduce the accidents that are increasing day by day. But unfortunately, they are not proven to be sufficient to reduce the number of accidents. Thus, only possible solution is to reduce the number of deaths caused due to these accidents.

2. RELATED WORK

State-of-the-art deep learning techniques have been recently applied to car crash detection using video data. Chan et al. (2017) offer a dynamic-spatial-attention (DSA) recurrent neural network (RNN) model to predict car accidents from dashboard cameras. Their model is trained to distribute soft-attention to object candidates gathering subtle cues dynamically as well as modeling the temporal dependencies of all cues to robustly predict accidents. Naidenov & Sysoev (2019) develop a car accident detecting system based on CNN using video capture recordings. Yao et al. (2019) present an unsupervised deep learning framework for traffic accident detection using dashboard cameras. In particular, their approach can detect traffic accidents by predicting traffic participant trajectories as well as their future locations.

While several existing studies based on the machine learning techniques above consider video data or audio data in order to detect car crashes, limited contributions have been made to consider both video data and audio data on roads for car crash detection. Such consideration is important to improve the performance of car crash detection, because one type of data may have information that the other type of data does not contain. The main contribution of this research is to provide an ensemble deep learning-

based car crash detection system that considers both video data and audio data from dashboard cameras in order to improve the performance of car crash detection.

We have come to notice that most accident detection systems make use of expensive sensors placed on the body of the vehicles or it makes use of existing sensors on a smartphone. This dependency of sensors makes this method expensive and less effective as compared to the proposed accident detection system.

3. METHODS & ALGORITHM

CNNs are used in modelling spatial data like images. CNNs have been successful in tasks like image classification, object detection, etc. CNN are used to model sequential data and make predictions based on them. LSTMs are widely used in areas of text classification, making language models, sequence generation, etc. Standard LSTMs can be used directly on sequential data where the input is spatial. Thus, to perform tasks which involve sequences of images or videos, a CNN architecture needs to be used.

The proposed model is a fusion of CNN layers for continuous video classification taken from a camera. The CNN part of the proposed model was mainly inspired by the Inception v3, but with certain tweaks it has fitted well to our training images.

A. Convolution Layer

The convolution layer is the first step toward extracting useful information from images. Convolution helps in preserving the relationship between the pixels of the image by learning image features using small squares of input data. This is done as a mathematical operation that takes two inputs, which are a part of the image as a matrix, and a kernel or filter.

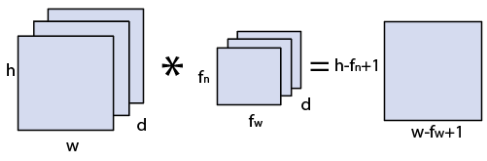

Fig. 1. Image matrix multiplies kernel or filter matrix

It is a picture of a 5 x 5 image matrix convoluted by a 3 x 3 filter matrix. The resultant matrix is a 3 x 3 matrix, called the feature map. Different filters applied to the same image can extract different information from the image. It may be used to extract spatial information such as edge detection, blur, etc.

B. Pooling Layer

The main job of pooling layers is to reduce the number of parameters when the given image is too large. Spatial pooling, also known as subsampling or down sampling, diminishes the dimensionality of each Feature Map but maintains the most relevant information.

Pooling layers are generally of 3 major types:

- Average Pooling
- Max Pooling
- Sum Pooling \

C. Fully Connected Layer

In this layer, we smoothen our grid into a vector and give it into a layer like a primitive neural network.

4. PROPOSED WORK

Since there was no dataset available, a dataset was created that includes accident and non-accident images. If an accident occurs, an alert message will be sent to the nearby hospital. We trained the system with the created dataset. The trained system is then incorporated with the cameras so as to capture the video of the vehicles on the road. By calculating the probability, the system predicts whether an accident happened or not. In case of an accident, an alert is sent to the control rooms using the API module. Fig. 1 is the flowchart depicting working of the system. The camera module records the video of vehicles in the road. The camera is placed at fixed locations, mostly in accident-prone areas. Whenever an accident occurs, it is predicted using our deep learning model and followed by sending alert message to the nearby hospital.

A. Dataset Creation

Our main focus is on accidents occurring in India. But the dataset was not available, so the dataset was created initially. It consists of five thousand images which includes 100 accident images and 100 non-accident images. The dataset consists of the

images of accident that happen at different parts of India. The images are downloaded from google images. All images are then converted into the same format and same size.

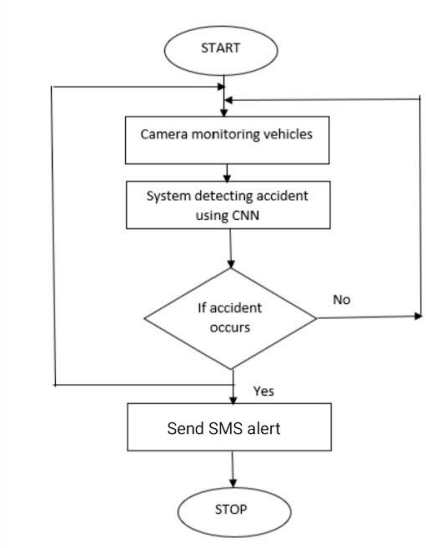

Fig. 2. Working of the System

B. Training and Validation

200 images belonging to 2 classes were used for training purpose. These images are passed through the different layers of convolutional neural network. It consists of convolutional and max pooling layers. The images are resized before passing through the convolutional layers. This happens when the model which we use fits so well to the training set. The dataset consists of accident and non-accident images. For validation, we use 200 images belonging to two classes which are accident and non-accident. The trained model has to be evaluated with some other images. This is important since it helps to evaluate the accuracy of our system. Dataset was trained for 30 epochs to get more accuracy. We got both training and validation accuracy and loss to understand how well the images have been learned by the model. Fig. 3 shows the accuracy gained for training and validation of images, the training and validation curve indicates that the system is trained well. Fig. 4 shows the loss occurred during training and validation. The curve here also indicates the loss incurred is also less as the number of epochs increased.

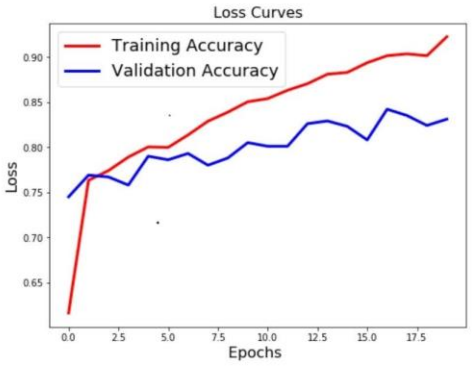

Fig. 3. Training and Validation accuracy.

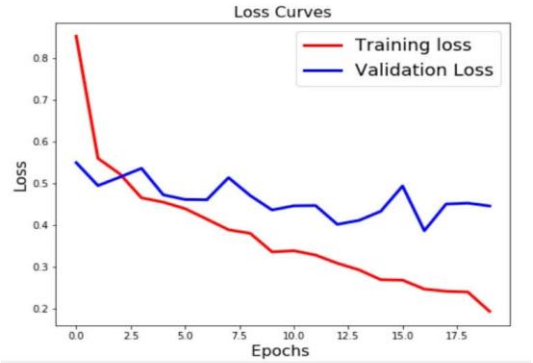

Fig. 4. Training and Validation loss

C. Video Classification

After training, model predicts accidents with an accuracy of 85 percent. For video classification, deque was used for performing the rolling prediction averaging. The deque is initialized to the default value 128. Then it starts to loop over frames of the video which is given as the input. If there are no frames to grab it means that it has reached the end of the video. The frame dimension was set to 64. This is done because the dataset with which we trained are images with 64*64 dimensions. The next step is the mean subtraction. After that, the results of predictions of each frame were obtained. These results are then stored into a deque. Using the probability of frames, it was able to predict whether it is an accident or not. The model will detect when an accident occurs and message will be sent to hospital.

All the layers as shown in table 1 were trained using the Image Data Generator in keras which allows image augmentation. Image

Augmentation is a process by which a set of pictures can be used to artificially expand the dataset. The most common features of image augmentation include zoom, pre-processing function, shear etc. These parameters basically apply random sheer, rotation, and zoom on the existing pictures thus giving the model more data to learn on. The values used for the proposed model are as follows:

Layer (type)	**Output Shape**	**Param #**
conv2d (Conv2D)	(None, 222, 222, 32)	896
activation (Activation)	(None, 222, 222, 32)	0
max_pooling2d (MaxPooling2D)	(None, 111, 111, 32)	0
conv2d_l (Conv2D)	(None, 109, 109, 32)	9248
activation(Activation)	(None, 109, 109, 32)	0
max_pooling2d_l (MaxPooling)	(None, 54, 54, 32)	0
conv2d 2 (Conv2D)	(None, 52, 52, 64)	18496
activation_2(Activation)	(None, 52, 52, 64)	0
max_pooling2d_2 (MaxPooling2D)	(None, 26, 26,64)	0
flatten (Flatten)	(None,43264)	0
dense (Dense)	(None,64)	2768960
activation_3(Activation)	(None,64	0
dropout (Dropout)	(None,64	0
dense 1 (Dense)	(None, 2)	130
activation_4(Activation)	(None, 2)	0

Table 1. CNN Trained Model

Total params: 2,797,730
Trainable params:2,797,730
Non-trainable params: 0

5. RESULTS

Once the system starts running it takes into account each frame of the video that it is capturing from the Road side camera and runs it through the proposed model and when it detects an accident the system immediately sends a message using the API module. It also sends the frame at which it detected an accident and what percentage of accident it is. It also shows the time stamp as to when the accident was detected. Fig. 5. shows an accident frame along with the details.

Fig. 5. Accident Detection on an Image

The below frame is of a non-Accident situation where only No Accident detected is shown.

Fig. 6. Non-Accident Detection on an Image

6. CONCLUSION

The proposed system is used to detect road accidents. When an accident is detected, an alert message is sent to nearby control rooms using the API module. This system is more reliable and economical when compared to existing systems. It can detect accidents with high level of accuracy as the model architecture is trained using the created dataset. With the help of this system, immediate action can be taken by sending alert to the officials and will help the medical teams to reach the accident spot in time and save the valuable human lives. Thus, the proposed system will play an important role in the society where road accidents have nowadays become a major threat.

REFERENCES

[1] Gokul Rajesh, Amitha Rossy Benny, Harikrishnan A, James Jacob Abraham and Nithin Prince John “A Deep Learning based Accident Detection System” IEEE 2020

[2] Sreyan Ghosh, Sherwin Joseph Sunny, Rohan Roney "Accident Detection using Convolutional Neural Networks" IEEE 2019

[3] Hankil Kim, Jinyoung Kim, and Heakyung Jung "Convolutional Neural Network Based Image Processing ystem" JICCE 2018

[4] Vargas-Cuentas, Natalia I., and Avid Roman-Gonzalez. "Facial image processing for sleepiness estimation." Bio-engineering for Smart Technologies (Bio SMART), 2017 2nd International Conference on. IEEE, 2017.

[5] Melissa Yauri-Machaca, Design of a Vehicle Driver Drowsiness Detection System through Image Processing using Matlab.

[6] Mugila.G, muthulakshmi.M, santhiya.K, prof.Dhivya.P- Smart helmet system using alcohol detection for vehicle Protection [International Journal of Innovative Research in Science Engineering and Technology (IJIRTSE) 2016

[7] Dhivya M and Kathiravan S, Dept. of ECE, Kalaignar Karunanidhi Institute of Technology- Accident- Avoidance System for Vehicles.