CNN BASED ACCIDENT DETECTION, AN AI BASED SUPPORT SYSTEM

SAHITHI PRASANTHI M

(18B21D5820)

Under the esteemed guidance of

Mr. P. RAMAKRISHNA M. Tech, Ph.. D

Assistant Professor

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

KAKINADA INSTITUTE OF ENGINEERING & TECHNOLOGY

(Approved by AICTE & Affiliated to JNT University Kakinada)

Yanam Road, Korangi -533461 E.G.Dist (A.P).

ABSTRACT

The rapid growth of the civilization has made our lives easier and increase the auto vehicles usage to a great extent. With great increase in the usage of cars and automotive vehicles, chances of getting accidents is also very high. India needs to improve the way they respond to the road accidents this is a system that can help in the identifying the severity of the accident and detect the accident using deep learning and computer vision techniques. This paper aims to monitor the accident in cities and to reduce the death rates. Nowadays, road accident rates are very high. Early detection and timely medical aid will help a lot in these situations. Regular traffic systems are implemented with cameras and installed in most of the town to watch and control traffic. A Smart City with an AI traffic monitoring and reporting mechanism, a more superior traffic monitoring method may recognize and discover moving objects like automobiles and motorbikes in live camera supports. Furthermore, detect collision of those moving objects and helps to provide an accurate location to the nearby center about the accident to supply immediate medical care and also sends a message to the closest police headquarters.

CHAPTER 1

INTRODUCTION

1.1 Accidents

The increase in the usage of automotive vehicles is bound to have increased the accident rates during the last five years in India. The traffic accidents are not only restricted to developed geographies, even tier 2 and tier 3 cities in India also prone to severe accidents in the last three years. The percentage of causalities happens more in underdeveloped countries due to the inappropriate transport infrastructure and not proper signal controlling. A system that can detect the accidents and send an alert to the immediate nearby response team is very much need of the hour.

There are many reasons why an accident happens when a vehicle collides with other vehicles. Accidents happen due to road obstacle, pedestrian, line crossing, objects on the road. However, the most common kind of road accident is a vehicle collision, and 60 per cent of the accidents happen because of the road vehicle collision.

Object detection technology is actively used to determine the position plus size of targeting objects emerging on image or video. Object tracking is a different field in image processing to happen accomplished with novel testimony and tracing the points of recognized objects across periods. Nevertheless, to trace objects, it is essential to determine object class and status first in a firstly yielded static image with object detection. Hence, the completion of object tracking should remain intensely reliant upon the fulfillment of the object detection included.

CHAPTER 2

SYSTEM ANALYSIS

2.1 Existing Method

The current frameworks don’t train the models at granular features level, most of them works based on training a model based on the existing developed models. In this project we have developed a model that started learning from scratch from hundreds of the images and used that as for predicting. The models are deployed as APIS and the data was fetched through APIs for prediction purpose.

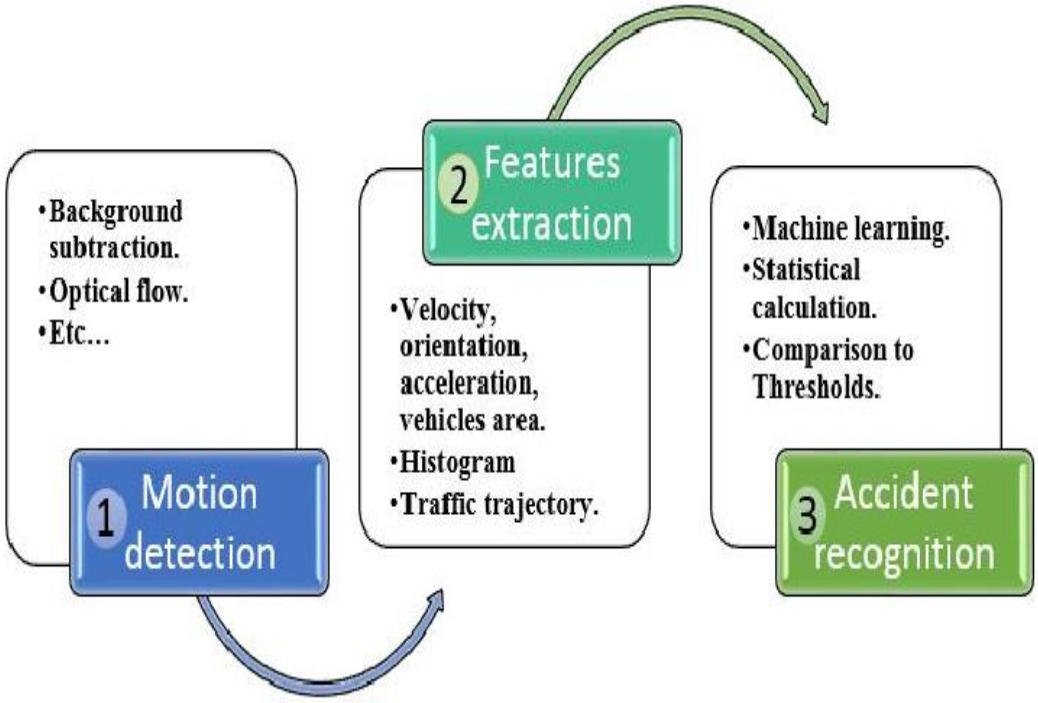

Figure 3.1: Existing System Architecture

2.1.1 Disadvantages of Existing System

- Models are used based on the pre trained datasets and doesn`t talks about object segmentation
- Level of accident severity was not captured using existing models
- Models use sensors to identify accident

2.2Proposed System

The proposed system mainly helps predict the type of accident and alert the police, and the nearest hospital so that the required first aid can be provided in time and save the life of the victim. We will be using Deep Learning based on Conventional Neural Network (CNN) to achieve the outcome of our proposed model. Here the accident image is captured and converted to the greyscale image.

The feature extracted to get different features to predict the severity-level of the accident. An alert to the nearest hospital, ambulance, police, and the traffic department can be alerted so that the traffic on the road can be handled. Proposed System covers the following phases

- Collection of data for training purpose
- Preprocessing the data
- Model building
- Model Training
- Model validation
- Deploy as API

- Use API for prediction

Figure 2.2: Proposed System Basic Architecture

2.3 Advantages

- The proposed architecture is a scalable solution to retrain based on the large number of iterations.
- In a single model itself we can predict weather the accident happened and severity of the accident.
- The moment accident occurs the accident will be taken as picture and shared with the relevant stake holders via email or other form of communication.

CHAPTER 3

SYSTEM DESIGN

System design is the process of the visualizing the entire architecture required for the product the process includes starting from designing the training datasets to creating convolutions and testing the models and validating the models it is the full process of the images to predictive values it has everything that is as part of the process.

3.1 System Architecture

System architecture talks about the high level components design that is required for the analysis we are going to have majorly three components of the system architecture preprocessing the data and validating the data and deploying the models for usage purpose.

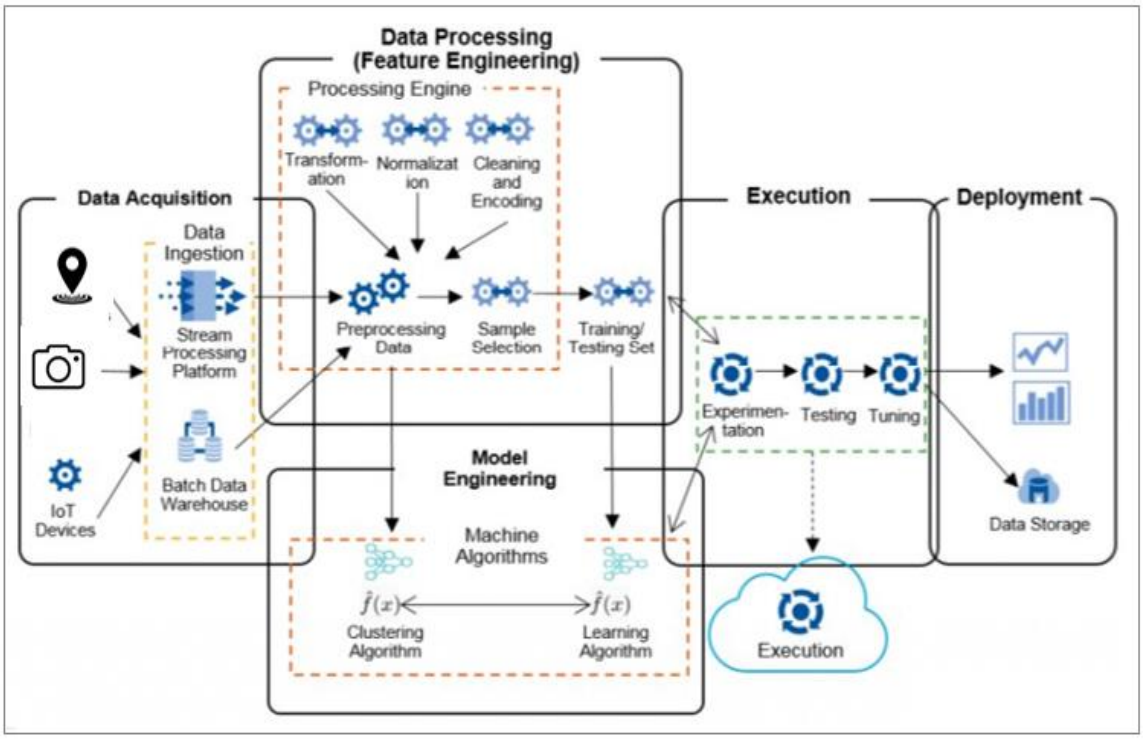

Figure 3.1: System Architecture

3.2 Sequence Diagram

This is the step-by-step procedure where, the data imported is validated, and the statistical analysis is made on the

data and the date is merged into sets, an error validation is also done and the algorithm is validated and the accuracy

is also tested.

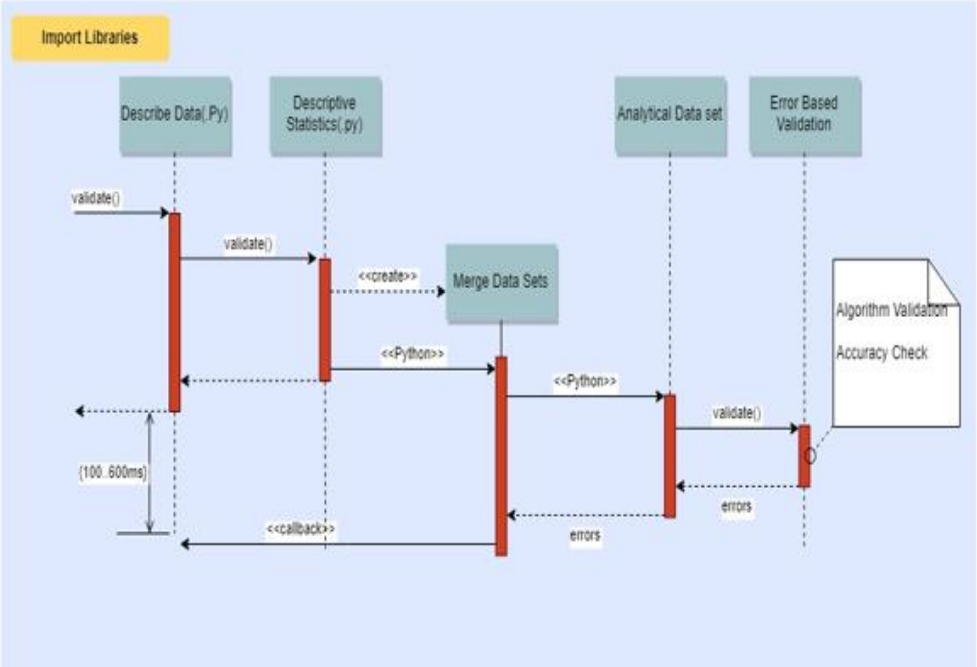

Fig 3.2 Sequence Diagram of accident severity

CHAPTER 4

IMPLEMENTATION

4.1 Hardware Requirement

- Processors : i3, i5, i7
- RAM : 4 GB
- Hard disk : 500 GB

4.2 Software Requirement

- Operating Systems : Windows 8,10 , Linux
- Language : Python-Stats Models, Sklean, Matplotlib, Seaborn Packages, keras, TensorFlow, open cv, pytorch
- IDE : Jupiter Notebook
- Backend : CSV,

4.3 Python Architecture

Python is an open source scripting language which was made by Guido van Rossum in 1989. It is a deciphered language with dynamic semantics and is exceptionally simple to learn. Python saves byte code (“.pyc” means compiled “.py” source) in heap memory for faster the execution time.

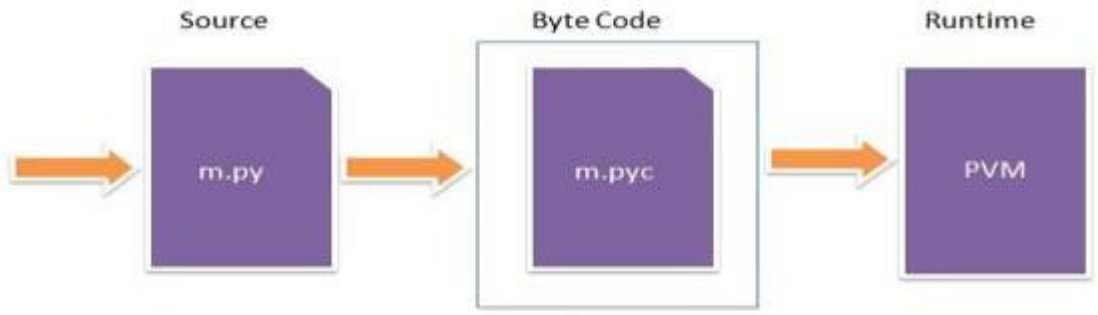

Figure 4.1: Python Architecture

4.4 COCO API

COCO (Common Objects in Context) is vast dataset of images created for segmentation, objects detections, persons keypoints detections, caption generation, and stuff segmentation. This package gives Python, Matlab, and Lua APIs that assist in parsing, loading, and visualize the annotation in COCO. The specific form of annotation is specified on COCO website. The Python and Matlab APIs are complete, and the Lua API gives only primary functionality.

Figure 4.2: Coco API Dataset Example

4.5 TensorFlow

TensorFlow stands an open-source end-to-end program for machine learning. It holds an extensive, libraries, extensible ecosystem of accessories and community sources that allow researchers to advance the state-of-the-arts in ML, and developer quickly builds and extend ML-power utilization.

4.6 OpenCV

Open Source Computer Vision Library is a repository of program function mostly targeting at real-time computers perception. Intels initially develops it, and it later backed by W Garage then Itseez. The libraries is free and cross-platform use under the open-source BSD license.

OpenCV has base as C++, and its first interfaces is in C++, but it holds less excellent though comprehensive older C interface. The new algorithms and developments appears in the C++ interfaces. There are binding in Java, Python, and OCTAVE. The API for this interface are found during the documentations.

4.7 MODULES OF IMPLEMENTATION

Any supervised algorithms of machine learning will require a systematic flux of the steps below. They are simplified structures that can help to better identify the problem and to structurally implement the all phases of the project. We have the following modules:

- Convolution
- Kernel Filter
- Max Pooling
- Activation Function
- Model Validation
- Model Deployment

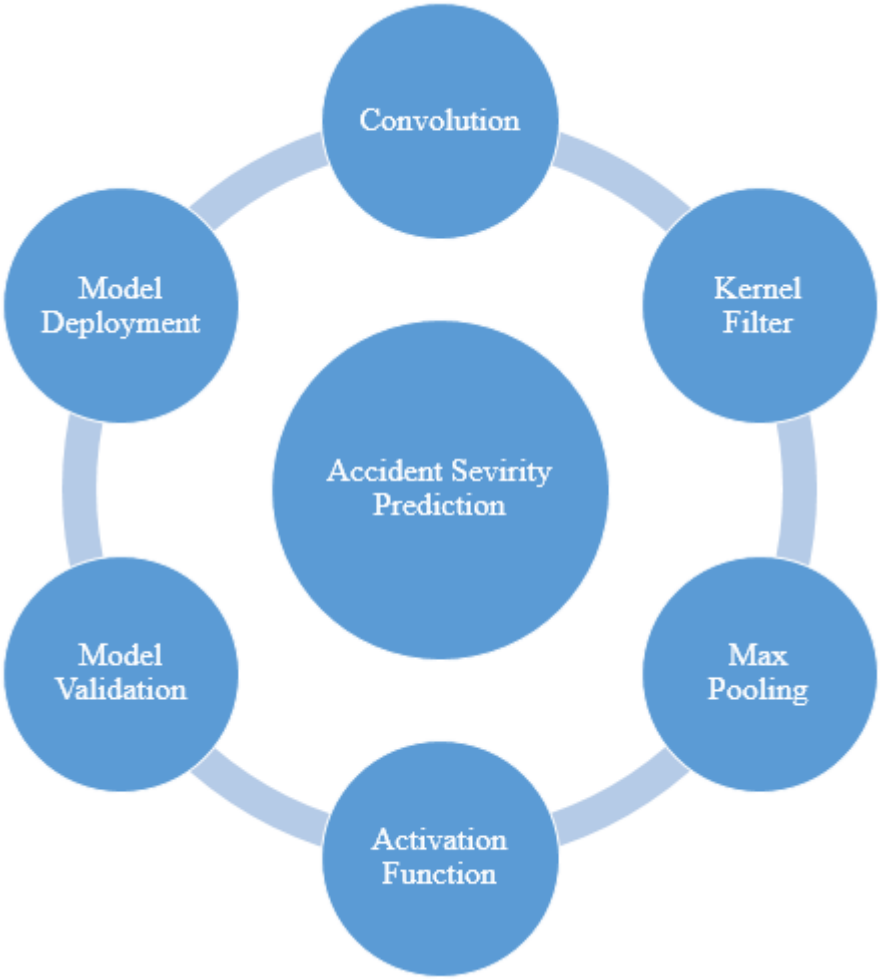

Figure 4.3: Modules used in accident severity prediction

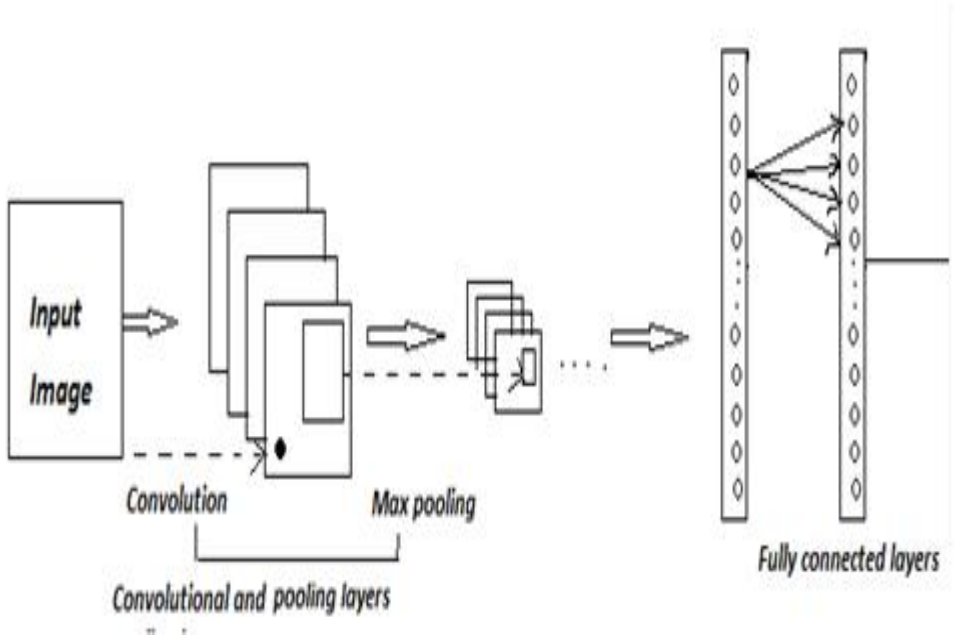

Figure 4.4: Different Modules in CNN

Module 1: Convolution

In mathematics, the term convolution is refereed as mathematical operations that applied upon a function with the help of the functions .it is defined as integral part of the product that is applied and

shifted over producing the convolution function. A CNN is Deep Learning algorithm that uses images as inputs and assigns weight and bias to different aspect of the images, and is capable of differentiating between them.

The time taken to preprocessing the data in convolutional neural networks is very less when compared to other machine algorithms the architectures of a convolutional neural networks is similar to the brain of the human where multiple neuron are connected over the cortex. It has the capacity to process the information through series of signals Images are generally made up of pixels the images size is typically represented in the form of rows, columns and RGB color formats. In case of a binary image that is a gray scale image the images are treated as binary values the black is represented as zero and white color is represented is white.

A ConvNet is can successfully capture the overlay parts of the images and an image will be processed through the pixel by pixel by taking a 3*3 or 5*5 convolution and the convolutions are processed through the entire images to create the new features after each convolution the images are processed through the weights and biases.

The below Figure, is an RGB image divided into three colors plane Red, Green, Blue. There are different such colour space in which image exists like RGB, Grayscale, HSV.

Processing an entire image is very time consuming and take lot of resources for computation and instead of this we are going to process the image by 5*5 or 3* 3 filters without losing features the image will be reconstructed to the image and process the image for analysis purpose.

Convolution Layer — the Kernel

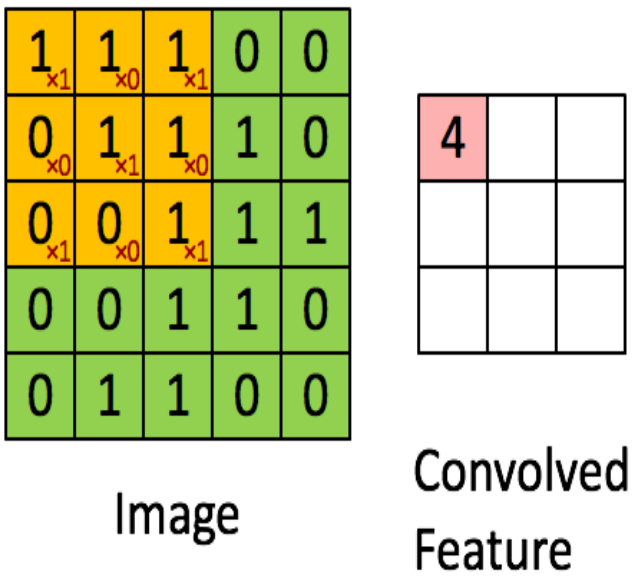

Figure 4.5: Convolution layer of the kernel

Original image dimensions are 32*32 and they were processed through a convolution filter of 3*3 which can help in generating an image of 28*28.

The visualization of 3*3 filters is in 3d space is visualized as below and the process leads to generate the multiple images of smaller sizes and the convolutional are used in creating feature maps the higher feature maps the higher the accuracy of the models can be possible over a period of time we are going to process the complete image for creating by applying a better kernel filters through the images

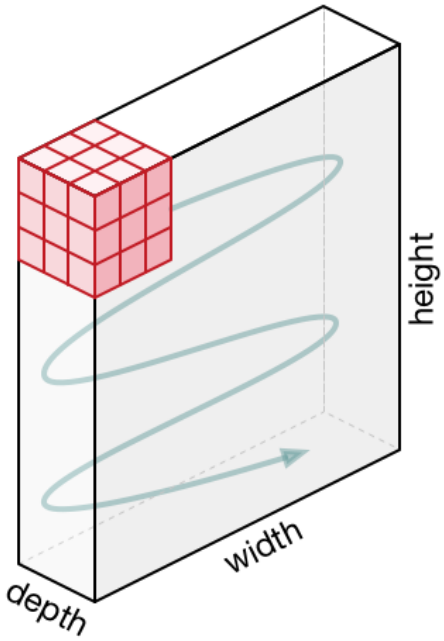

Figure 4.6: Visualization of 3x3 filters in 3d space

The feature maps are the actual features that goes as input to the convolutional neural networks the images will be processed through a predefined the activation functions and the activation functions are used to calculate the different weighted features and the process of the process the image is very fast. Before the fully connected layers the entire images processing happens through the convolutional neural networks and they are process with large scale and faster than the expected method. With different types of kernel s applied there is great chance of the getting bets features in the convolutional neural networks and the output is predicted the final outcome.

ACTIVATION MAPS

Creation of activation maps is one of the key feature in the deep learning algorithms. Below are the steps used in creating the convolutional neural networks?

1. Decide the size of the convolutional filter that needs to be applied on the original images
2. Start the kernel shift from the top and bottom to the entire image which results in the first activation map

3. Take another filter and another image and process this through the entire image that leads to the predicting the final outputs
4. A series of multiple filters leads to the final output and these are called as convolutional activation maps

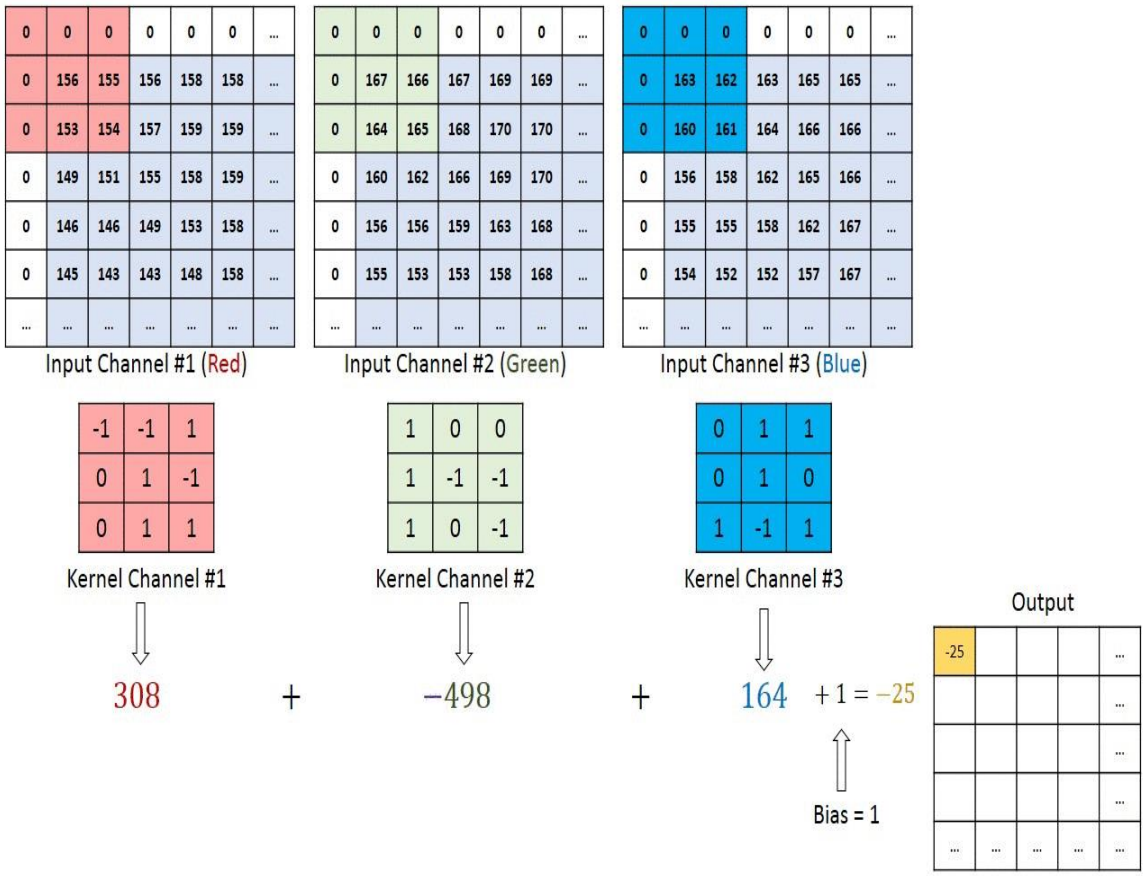

Figure 4.7: Activation maps

Module 2: Kernel Filter

The kernel filters are generally a matrix of operations that applied up on original matrix

1. **Convolutional 1D:** it has row level values the values that are randomly generated. it is used to multiple with the original kernel of the image
2. **Convolutional 2D:** These images are 2*2 matrixes. It has rows and columns as input to the image.
3. **Convolutional 3D:** These convolutional filters are 3*3 and they have been used to process the RGB images and the process of the images was happens through this multi-channel the images and the images was sided over the convolutional process and sided over the image to generate the multiple features maps this is standard filter we us whenever we want to process the images of the convolutions and created the features.

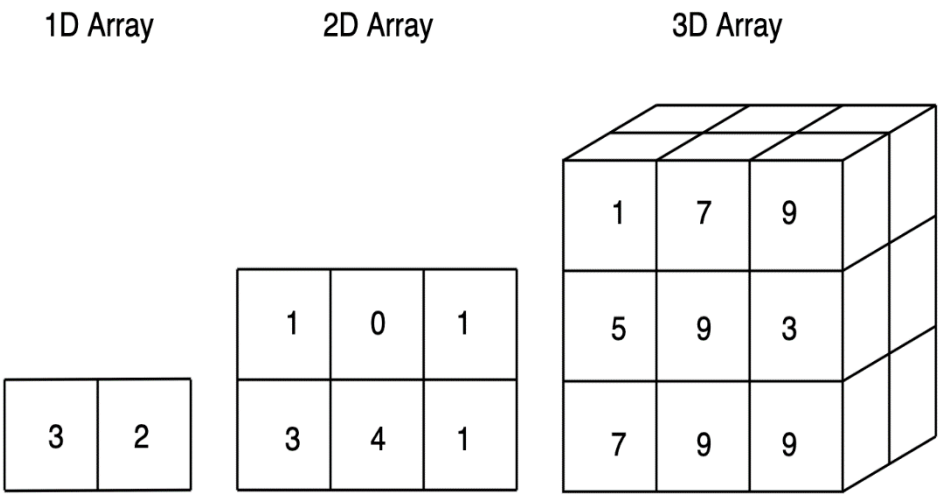

Figure 4.8: Kernel Filter

Module 3: Max Pooling

Convolutional neural networks are images that a filter applied systematically on the images to process images. The process of converting multiple feature maps can result in the create great feature maps we will be using a method called down sapling to process the images and the images processed through down sampling images are having better features. Convolutional neural networks are proven to be very effective the when we are stacking multiple layers.

A common approach to this problem is that down sampling can achieve through the process of taking a maximum of the convolutional neural network convolutions. A better approach is to use the pooling layer. The pooling layers is the method of using the maximum number of the image. It can also be called as max pooling.

The two types of pooling layers are the listed below

1. Average pooling : After applying the filter we take the average number out of all the processing
2. Maximum Pooling :
    a. After applying the kernel we will be applying the maximum number as the process for the analysis

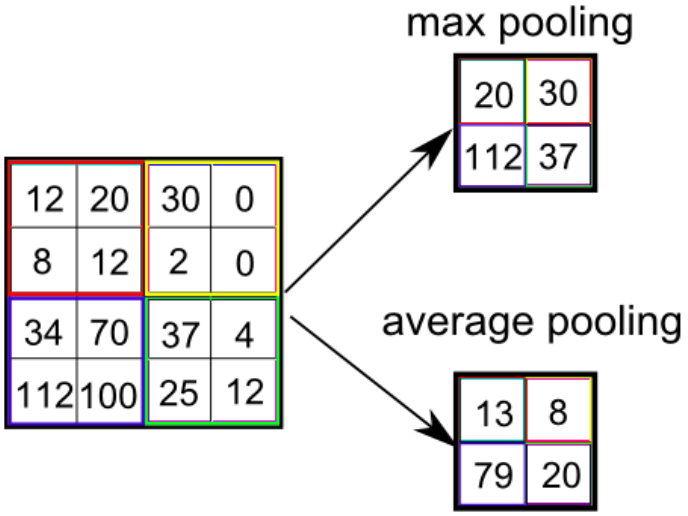

Figure 4.9: Max Pooling

Module 4: Activation Function

Activation functions are the key to convert the images to meaning full features that can be used as input to the convolutional neural network. Activation functions also can help in processing the images trough the final outputs of the models that can result in predicting the outcome of a variables.

Sigmoid Function: The sigmoid activation function is a binary predictive variable it takes the input of a value and try to squashes between the values that are in the range of 0 to 1 this particular activation function is used in the process of predicting the final variables of a binary outcome of the model. In general machine learning model perspective it is also called a sigmoid functions. Based on the input value the final outcome will be predicted in the range of 0 to 1 and the final output can be used to make informed decision about the models.

$$\sigma(x) = \frac{1}{1+e^{-x}}$$

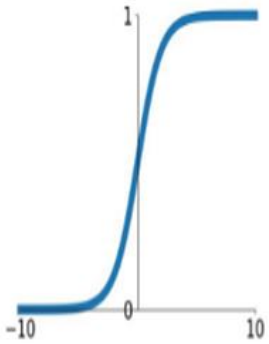

Figure 4.10: Sigmoid Function

TANH: This is a common activation functions when we want to predict a continuous variable. The continuous variables that can predict a range of values between – infinity to + infinity and this can leads to a great extension of the images that are possible to predict the final outcome of a continuous variables.

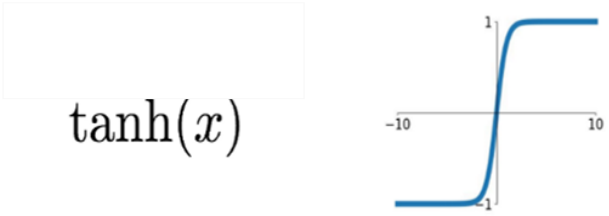

Figure 4.11: TANH Function

RELU: Rectified linear units is the activation function that is used to predict a continuous variables between 0 to infinity. The process of the predicting the output using the input variables is done via activation function. The Tanh is function that helps in predicting a positive number. It is very common practice for predicting sales, units, transactions …etc.

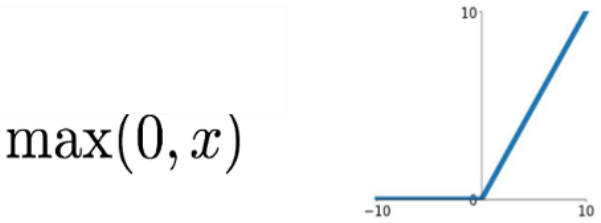

Figure 4.12: RELU Function

LEAKY RELU: It is one the activation function that can act like relu but with a deviation towards the negative values. The process of the prediction is where the value will range between a negative values to zero to positive number.

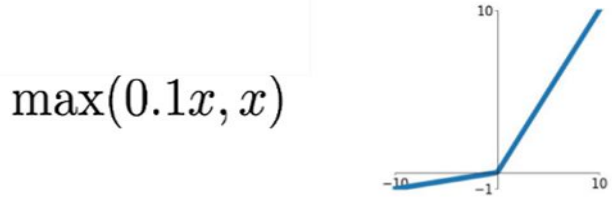

Figure 4.13: Leaky RELU Function

MAXOUT: The activation function's is a combination of the weight and biases that are processed through the weights and biases. There are multiple values that can be processed through the activation function method and they have been variously proceeded through the multiple layers of convolutional neural networks to get the final values of the output.

$$\max(w_1^T x + b_1, w_2^T x + b_2)$$

Figure 4.14: Maxout Function

ELU: Exponential linear units that are the process of the predicting a value through the function that has the exponential values. The exponential values that ranges between 0 and infinity. The process of the predicting the values lies between zeros to infinity.

$$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$$

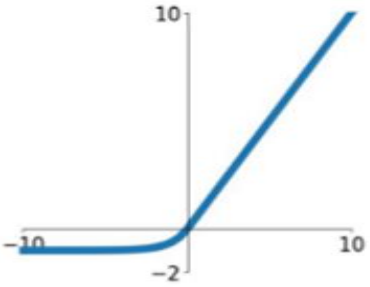

Figure 4.15: ELU Function

Module 5: MODEL VALIDATION

Model validation is the common process of the validating the predicted model through the multiple images and the multiple images are processed through the convolutional neural networks. A batch of 80 images is has been used to process through the convolutional neural networks and predicted the predicted values are compared against the original values and the accuracy of the model will be judged.

As a research, we have narrowed down all the different techniques below as the reference based on the volume of the data we can select the technique that is used for predicting the images in the feature. A common practice is to use the image that are given for training purpose and divide them into two categories as they would be seeing as training and testing and a common ratio split is 80:20

Different techniques exist for the model validation and these are primarily used for validating the numerical data and however, can be leverage for the images as well these are the process that can help in predicting the models that are predicting are right or wrong

- Leave-p-out cross-validating.
- Leaving-one-out cross-validating.
- K-fold cross-validation.
- Holdout method.
- Returned random sub-sampling validating.

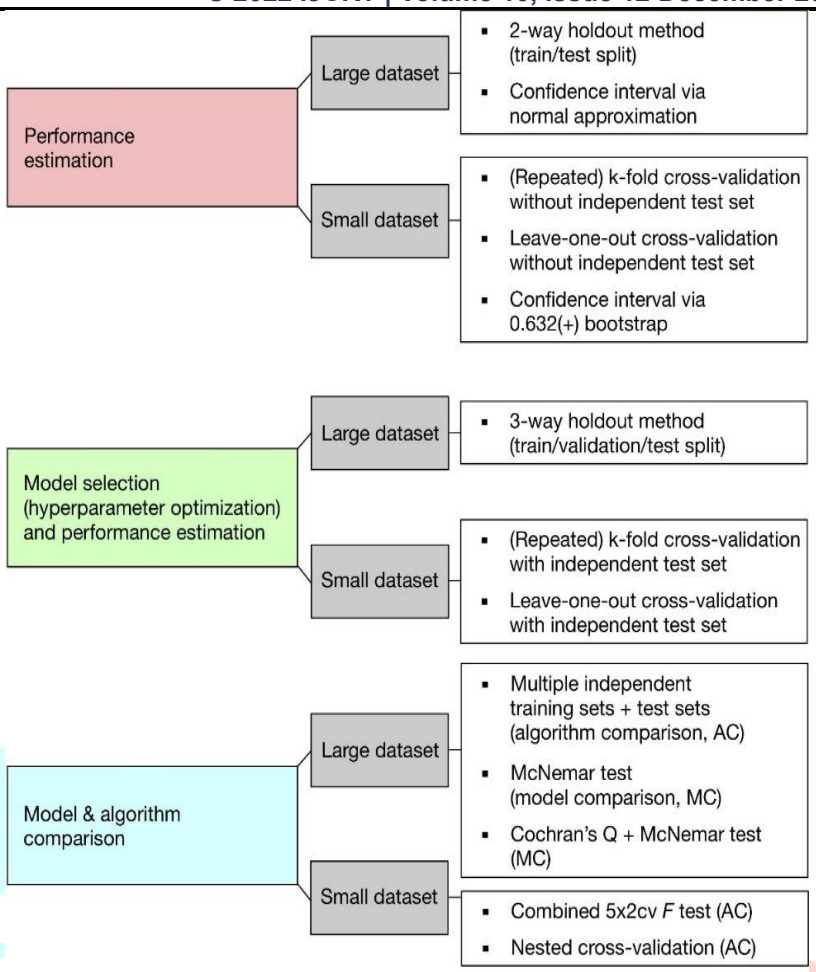

Figure 4.16: Model Evaluation Process

Module 6: MODEL DEPLOYMENT

Once the model is developed it needs to be hosted in an environment which can be easily available for everyone access there are many techniques available for hosting propose in this project we have selected Django and flask frameworks for predicting the final values. The model output will be created as API and the API will be used for the predicting the accident happening events and non-happening events. We have trained the models with hundreds of images and the hundreds of images re used for predicting the labelled data into the process of the predicting for the new variables.

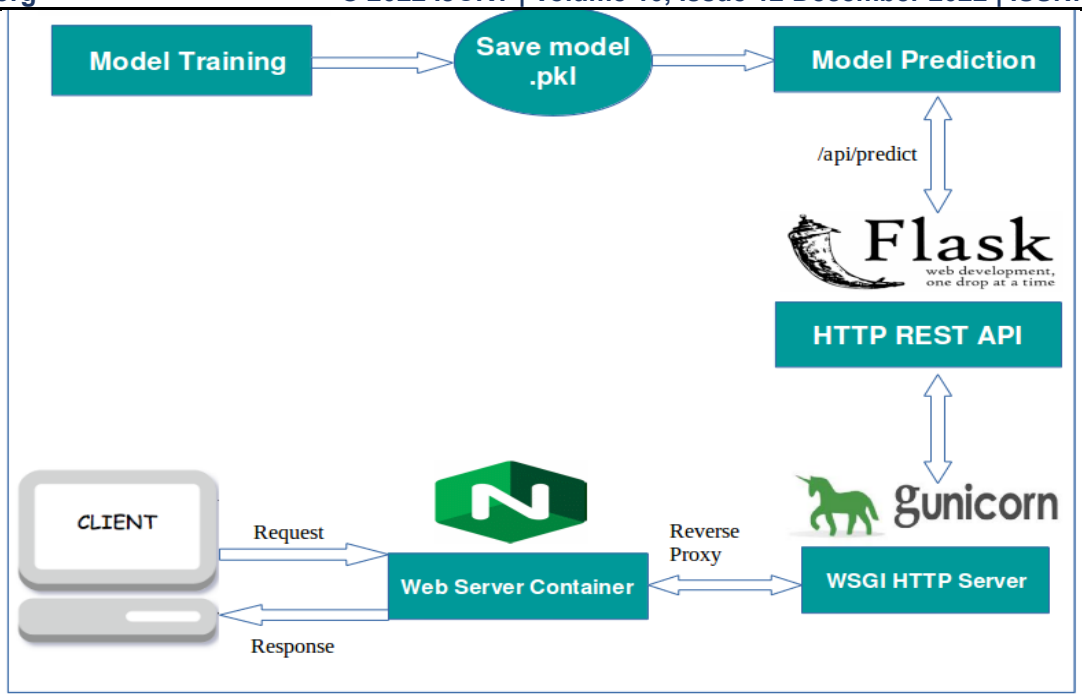

Figure 4.17: Deployment Model

CHAPTER 5

RESULTS AND SNAPSHOTS

5.1 RESULTS

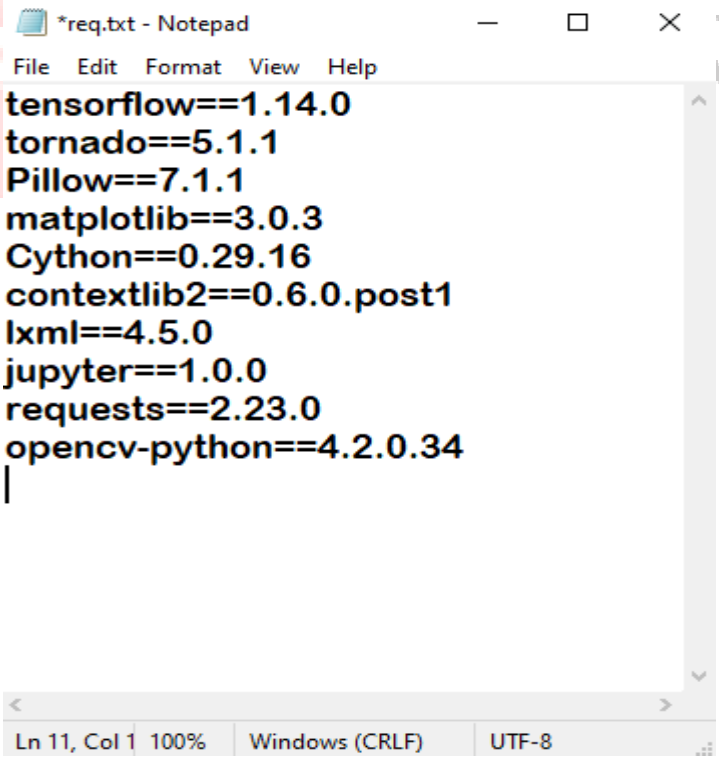

Figure 5.1: Dependencies Installed

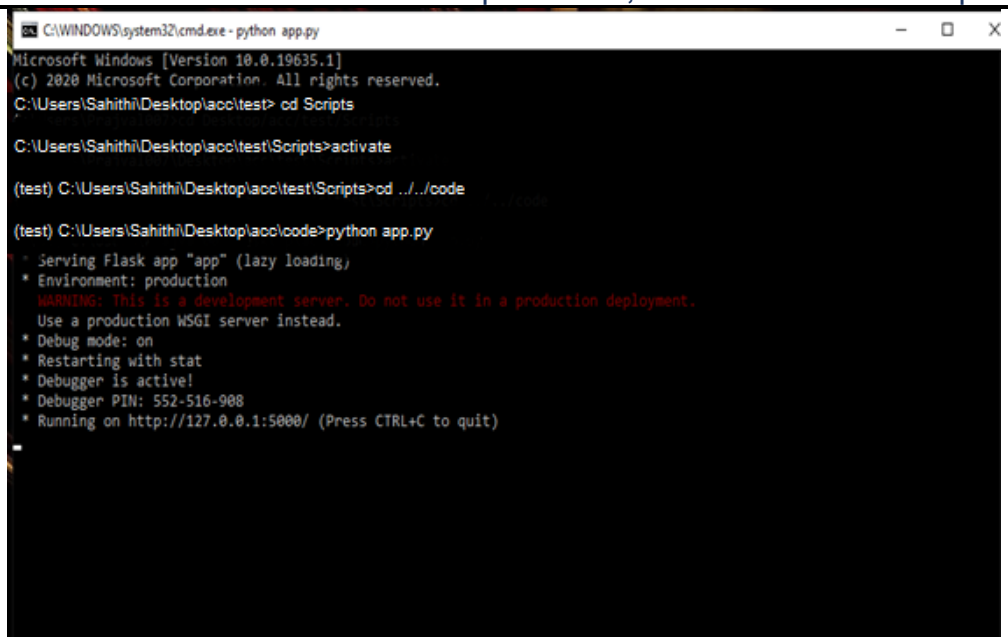

Figure 5.2: Starting the local server

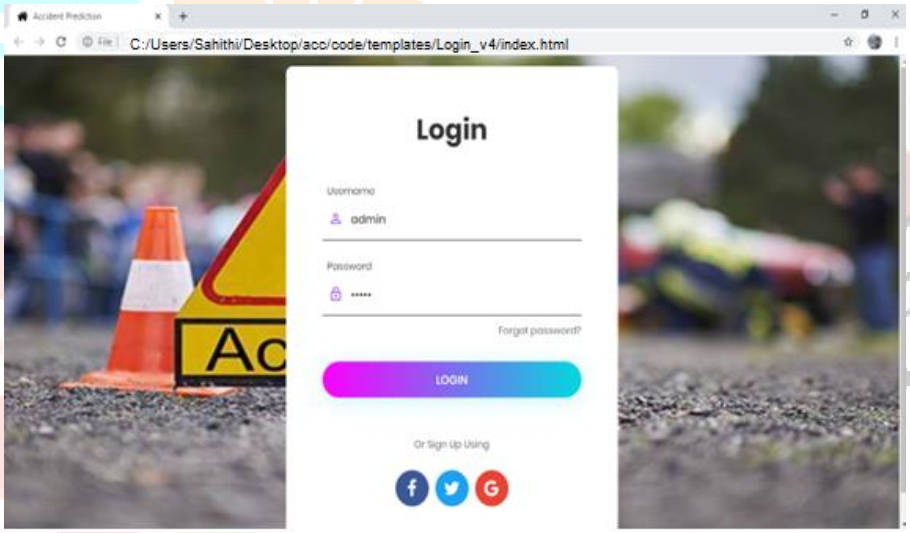

Figure 5.3: Login page

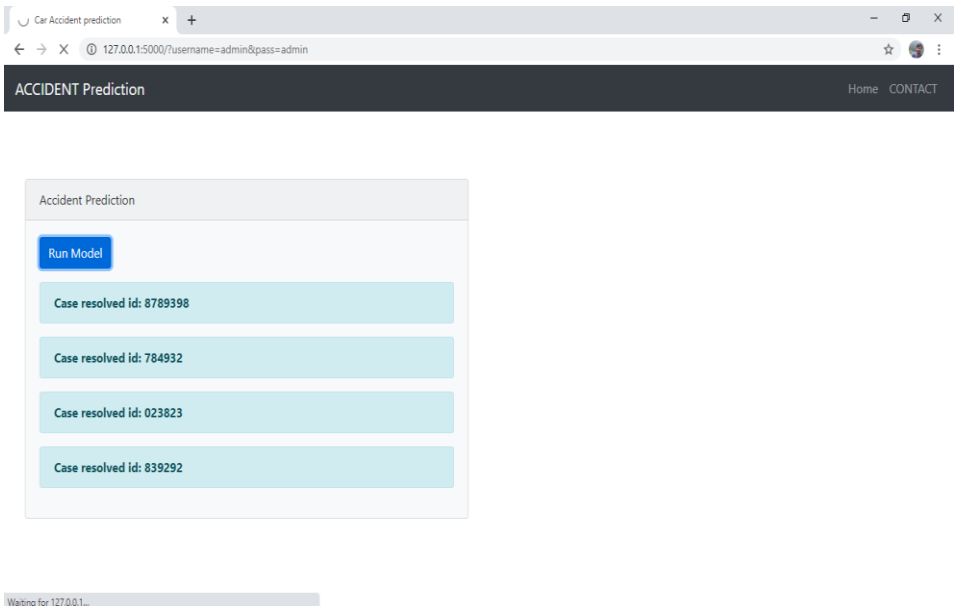

Figure 5.4: Homepage

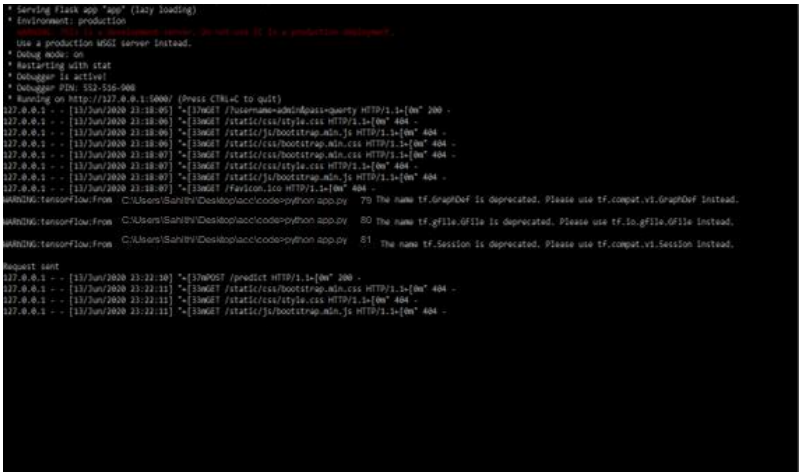

Figure 5.5: Command prompt showing details of backend

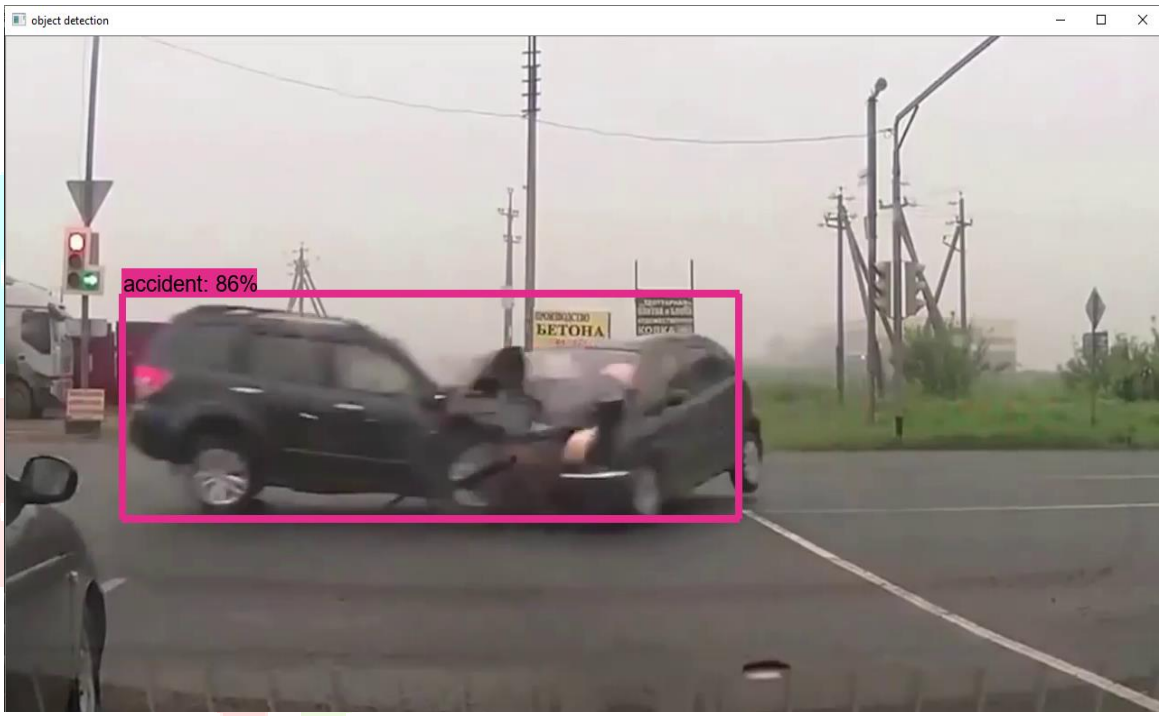

Figure 5.6: Accident detection video screen

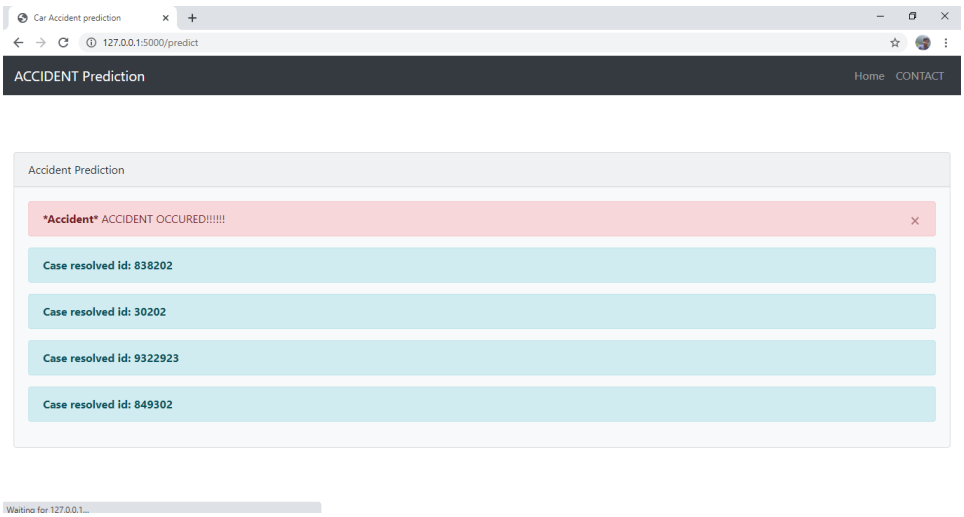

Figure 5.7: Alert message on homepage

Figure 5.8: Accident screenshot captured

CHAPTER 6

CONCLUSION AND FUTURE ENHANCEMENT

With the implementation of this project, we can provide the emergency service very fast and will be very useful. This will help the traffic department to manage the traffic and clear the accident spot. In case of a serious or major accident and there is nobody to call the ambulance or the police to inform about the accident, our project will be very much useful in such cases as automatically the police and the ambulance will be alerted with the severity of the accident and the accident image. To further enhancements and advanced models can be used to identify the accidents. An end-to-end product which can be used by the clients to identify how many people in the accident are injured this can be treated as enhancement for future work.

REFERENCES

[1] D Zhang and M. R. Kabuba, “Combining Weather Condition Data to Predict Traffic Flow: A GRU Based Deep Learning Approach”, No.978-1-5386-1956-8/17, 2017

[2] Daxin Tian, Chuang Zhang, Xuting Duan, and Xixian Wang “An Automatic Car Accident Detection Method Based on Cooperative Vehicle Infrastructure Systems” No.2019.2939532, Vol.7, 2019

[3] Fang Zong, Hongguo Xu, and Huiyong Zhang, “Prediction for Traffic Accident Severity: Comparing the Bayesian Network and Regression Models”, Hindawi Publishing Corporation Mathematical Problems in Engineering Vol.2013, Article ID 475194, 2013

[4] Kyu Beom Lee, Hyu Soung "An application of a deep learning algorithm for automatic detection of unexpected accidents under bad CCTV monitoring conditions in tunnels" Shin International Conference on Deep Learning and Machine Learning Emerging Applications, 2019

[5] Liujuan Cao, Qilin Jiang, Ming Cheng, Cheng Wang "Robust Vehicle Detection by Combining Deep Features with Exemplar Classification", PP: S0925-2312(16)30644-0, 2016

[6] Liwei Wang, Yin Li, Svetlana Lazebnik, "Learning Deep Structure-Preserving Image-Text Embedding", IEEE Conference on Computer Vision and Pattern Recognition, PP: 1063-6919/16, 2016

[7] Marco D'Ambros, Michele Lanza, Faculty of Informatics University of Lugano, Switzerland*: "*An Extensive Comparison of Flaw Prediction Approaches".in International Conference on Software Engineering. 2010.

[8] Sahithi Prasanthi M,Subrahmanya Sarma M,Abhiram M,Dr Y Srinivas," KNN File Classification for Securing cloud infrastructure", proceedings of RTEICT IEEE International conference May 2017

[9] Sahithi Prasanthi M,Subrahmanya Sarma M,Abhiram M,Dr Y Srinivas,"Insider Threat Detection With Face Recognition And KNN User Classification", proceedings of CCEM IEEE International conference November 2017

[10] Sahithi Prasanthi M,T Srikara Krishna Perraju,Subrahmanya Sarma M,Abhiram M,Dr Y Srinivas,"IOT Based Street Light Management", proceedings of CCEM IEEE International conference May 2018.