Annexure-1 (A typical Specimen of Title Page)

ACCIDENT DETECTION SYSTEM

A Project Work

Submitted in the partial fulfillment for the award of the degree of

BACHELOR OF ENGINEERING

IN

Artificial Intelligence and Machine Learning

Submitted by:

AYUSH MUDGAL (18BCS6124)

SHUBHAM KARWAL (18BCS6129)

HIMANSHU KUNCHAL (18BCS6131)

SHIVAM (18BCS6138)

HRITHIK JOHRI (18BCS6145)

Under the Supervision of:

GURPREET SINGH PANESAR

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
APEX INSTITUTE OF TECHNOLOGY

CHANDIGARH UNIVERSITY, GHARUAN, MOHALI - 140413,

PUNJAB

APRIL-MAY & 2021

DECLARATION

We, student of **'Bachelor of Engineering in Artificial Intelligence and Machine Learning' of session 2018- 2022**, Department of Computer Science and Engineering, Apex Institute of Technology, Chandigarh University, Punjab, hereby declare that the work presented in this Project Work entitled '**Accident Detection System'** is the outcome of our own bona fide work and is correct to the best of our knowledge and this work has been undertaken taking care of Engineering Ethics. It contains no material previously published or written by another person nor material which has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgment has been made in the text.

AYUSH MUDGAL (18BCS6124)
SHUBHAM KARNWAL (18BCS6129)
HIMANSHU KUNCHAL (18BCS6131)
SHIVAM (18BCS6138)
HRITHIK JOHRI (18BCS6145)

Date :
25th,April, 2021
Place:
Chandigarh University

ACKNOWLEDGEMENT

We thank the almighty for giving us the courage and perseverance in completing the minor project. This project itself is acknowledgement for all those people who have given us their heartfelt co-operation in making this project a grand success. With extreme jubilance and deepest gratitude, we would like to thank our mentor, Mr.Gurpreet Singh Panesar, for his constant encouragement. We are profoundly grateful towards the unmatched services rendered by all my team members. Our special thanks to all the faculty of the Computer Science Department and peers for their valuable advice at every stage of this work. Last but not least, we would like to express our deep sense of gratitude and earnest thanks giving to our dear parents for their moral support and heartfelt cooperation in doing the minor-project.

Table of Content

Title Page

Declaration

Acknowledgement

Abstract

1. Introduction

1.1 Problem Definition

1.2 Project Overview / Specification

1.3 Software Specification

2. Literature Review

3. Methods

3.1. Infrastructure of Model

3.2. Data

3.3 Dataset

3.4 Preprocessing of the data

3.5 Making Model

3.6 Alert System

4.Research Objective

5. Problem Faced

6. Result

7. Conclusion

8. References

1. INTRODUCTION

According to the World Health Organization, there are about 1.35 million deaths and 20-50 million injuries as a result of the car accident globally every year. Especially, a certain proportion of deaths and injuries are due to untimely treatment and secondary accidents, which results from that rescue agency and vehicles around accidents cannot obtain quick response about the accident. Therefore, it is vital important to develop an efficient accident detection method, which can significantly reduce both the number of deaths and injuries as well as the impact and severity of accidents. Under this background, many fundamental projects and studies to develop efficient detection methods have been launched for developing and testing.

Although the exploitation of traffic camera views is promising, the number of datasets aimed at learning to detect and predict the accidents on those views is limited due to several unaffordable factors: (i) traffic accidents are rare events, thus, acquiring enough data by recording at a road intersection is infeasible because one may have to wait endlessly for the accident to happen; and (ii) the access to traffic camera data is legally difficult to obtain in practice. To this end, we propose an effective data collection process to exploit the edge-case data: YouTube videos of traffic accidents that have been uploaded by users over the world. We exploited the search engine of YouTube, and added our annotation processes using both internal annotators and outside workers to build a novel dataset, the Car Accident Detection and Prediction (CADP) dataset for multiple purposes: temporal segmentation, object detection, tracking, vehicle collision, accident detection and prediction. Our dataset contains 230 videos, each video containing at least one accident captured from fixed traffic camera views and 1,416 segments of traffic accidents.

(a) (b)

Figure 1: *(a) Can you depict* ***where*** *the accidents happen in the image plane?; (b) Can you* ***identify*** *and* ***forecast*** *the sequences containing accidents?* Best viewed in color.

1.1 Deep Learning:

Deep learning is an artificial intelligence function that imitates the workings of the human brain in processing data and creating patterns for use in decision making. Deep learning is a subset of machine learning in artificial intelligence (AI) that has networks capable of learning unsupervised from data that is unstructured or unlabeled, also known as deep neural learning or deep neural network.

Deep learning has evolved hand-in-hand with the digital era, which has brought about an explosion of data in all forms and from every region of the world. This data, known simply as big data, is drawn from sources like social media, internet search engines, e-commerce platforms, and online cinemas, among others. This enormous amount of data is readily accessible and can be shared through fintech applications like cloud computing. However, the data, which normally is unstructured, is so vast that it could take decades for humans to comprehend it and extract relevant information. Companies realize the incredible potential that can result from unravelling this wealth of information and are increasingly adapting to AI systems for automated support. Deep learning learns from vast amounts of unstructured data that would normally take humans decades to understand and process.

1.2 Neural Networks

A neural network is a network or circuit of neurons, or in a modern sense, an artificial neural network, composed of artificial neurons or nodes. Thus, a neural network is either a biological neural network, made up of real biological neurons, or an artificial neural network, for solving artificial intelligence (AI) problems. The connections of the biological neuron are modelled as weights. A positive weight reflects an excitatory connection, while negative values mean inhibitory connections. All inputs are modified by a weight and summed. This activity is referred to as a linear combination. Finally, an activation function controls the amplitude of the output. For example, an acceptable range of output is usually between 0 and 1, or it could be 1 and 1.

These artificial networks may be used for predictive modelling, adaptive control and applications where they can be trained via a dataset. Self-learning resulting from experience can occur within networks, which can derive conclusions from a complex and seemingly unrelated set of information. Artificial intelligence, cognitive modelling, and neural networks are information processing paradigms inspired by the way biological neural systems process data. Artificial intelligence and cognitive modelling try to simulate some properties of biological neural networks. In the artificial intelligence field, artificial neural networks have been applied successfully to speech recognition, image analysis and adaptive control, in order to construct software agents (in computer and video games) or autonomous robots.

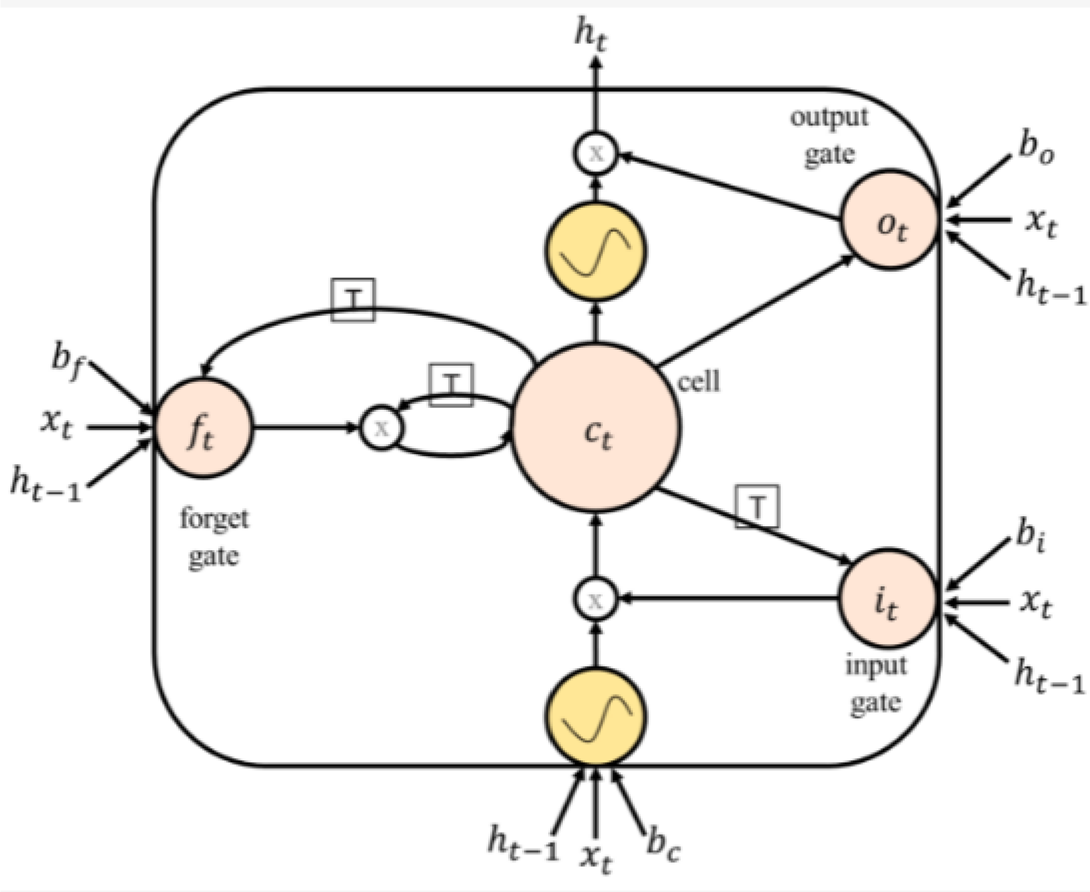
1.3 RNN

Recurrent Neural Networks (RNNs) are neural networks with feedback connections specifically designed to model sequences. They are computationally more powerful and biologically more reasonable than feed-forward networks (no internal states). The feedback connections provide a RNN the memory of past activations, which allows it to learn the temporal dynamics of sequential data. A RNN is powerful because it uses contextual information when mapping between input and output sequences. However, the traditional RNNs have a problem called vanishing or exploding gradient. To handle this problem, Hochreiter and Schmidhuber [23] proposed the Long Short-Term Memory (LSTM) algorithm.

1.4 LSTM

In LSTM, the hidden units are replaced by memory blocks, which contain one or more self-connected memory cells and three multiplicative units (input, output, forget gates). These gates allow writing, reading, and resetting operations within a memory block, and they control the overall behavior internally. Let ct be the sum of inputs at time step t, then LSTM updates for time step i at given inputs xt, ht−1, and ct−1

Figure 1. The structure of a memory cell in the Long Short-Term Memory–Recurrent Neural Network (LSTM–RNN).

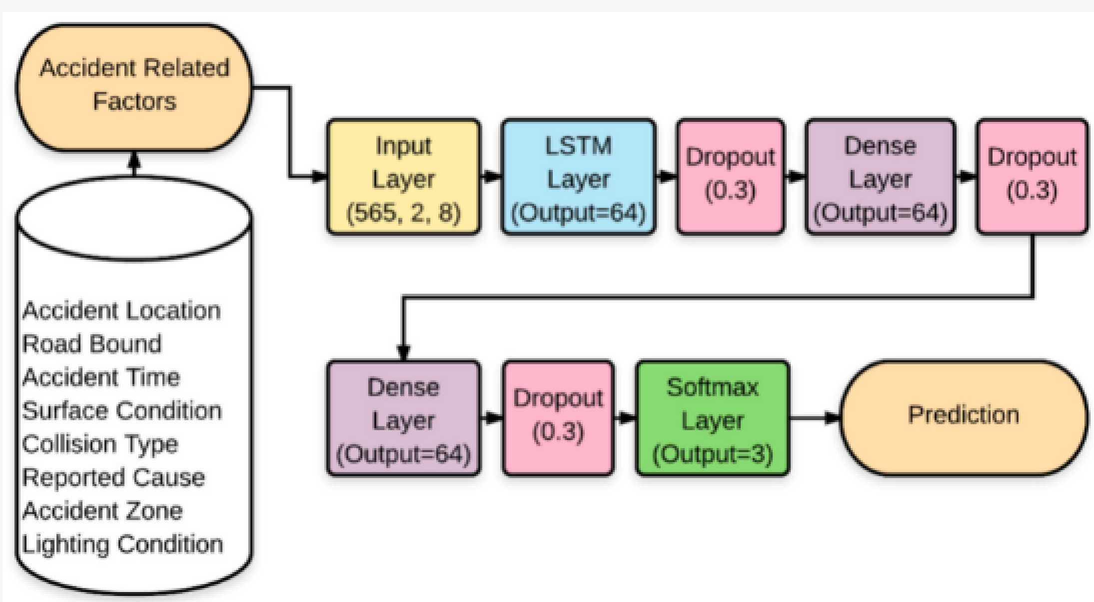


2 Literature Review Summary

Accidents and Car Modeling Dataset With the advancement of smart city and autonomous driving concepts, recent research has focused on traffic safety control using computer vision techniques. In addition, it offers information on related datasets for traffic control at road intersections: the DETRAC dataset for traffic camera events, for road traffic modeling, for single views at complex intersections dataset which simulated videos generated for traffic modeling that contains recording at a busy intersection and also consists of videos with fog, rain and snow to model traffic car behavior near intersections. On the CADP dataset, we perform experiments with Faster R-CNN to demonstrate detection efficiency. A new UCF-Crimes dataset for traffic accident videos contains 13 real-world anomalies such as Abuse, Accidents, and Shooting, and is based on understanding violent scenes in film. Both Dashboard and Traffic camera views, we believe, will provide crucial information for predicting accidents. In comparison to dashboard cameras, traffic cameras provide an overview of the entire road and can therefore monitor more vehicles. Based on the detection in the current picture, Pedestrian Detection predicts details about the pedestrian location. It offers a detailed description and reasons for using pedestrian monitoring to replace continuous detection and achieve real-time efficiency for pedestrian detection. It demonstrates that adding extra features, flow data, and background data are complementary additions that result in substantial gains over other strong detectors. It proposes a Faster R-CNN extension that uses contextual information and multi-level features to identify pedestrians in cluttered backgrounds by embedding information from a wider region beyond the original area of interest.

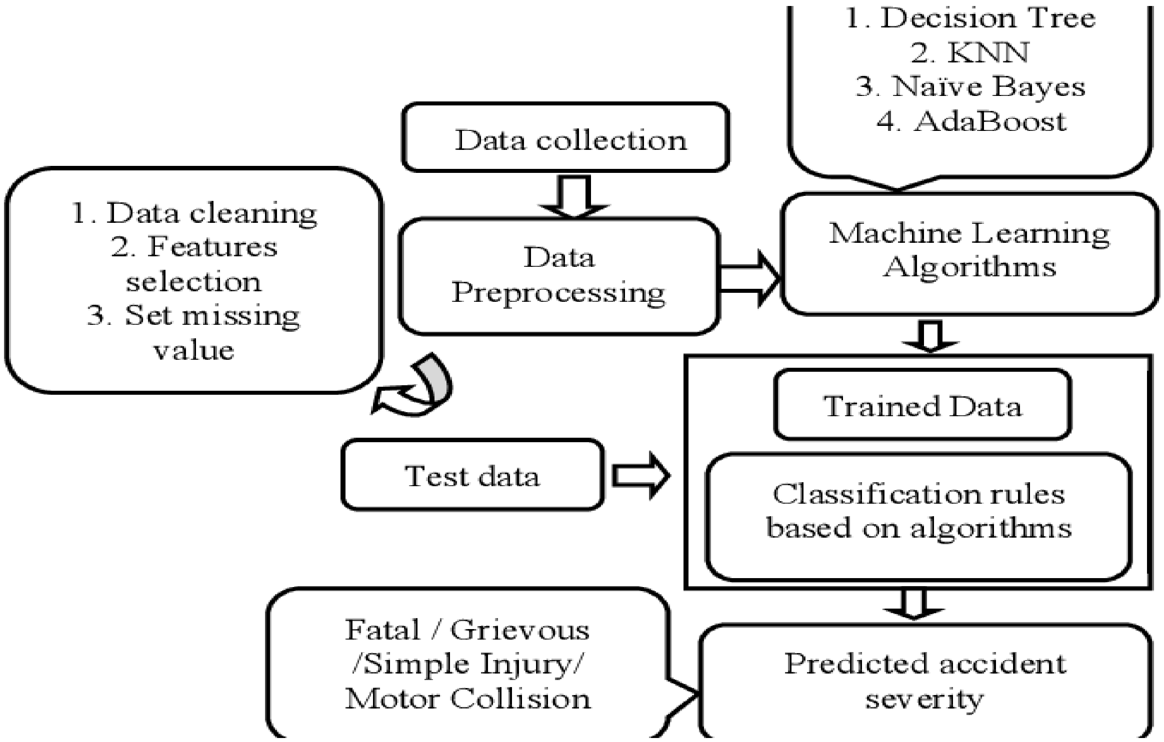
Detecting and Predicting Accidents Several studies on the use of cameras for accident prediction have been published in recent years. Dashboard cameras, for example, can be used to predict accidents. We agree that those datasets are critical for improving the response time of autonomous vehicles including self-driving cars and assisting with road surveillance.

Figure 2. The high-level architecture of the proposed RNN model used in this study.



3. Methodology

3.1 Basic Infrastructure of the Model



3.2 DATA

We used the CADP dataset for videos containing accidents and the DETRAC dataset which was originally for object detection of vehicles, as our videos did not contain accidents. To expand our dataset we have also downloaded YouTube videos that contain accidents. Over 380 videos were collected from the above mentioned sources.

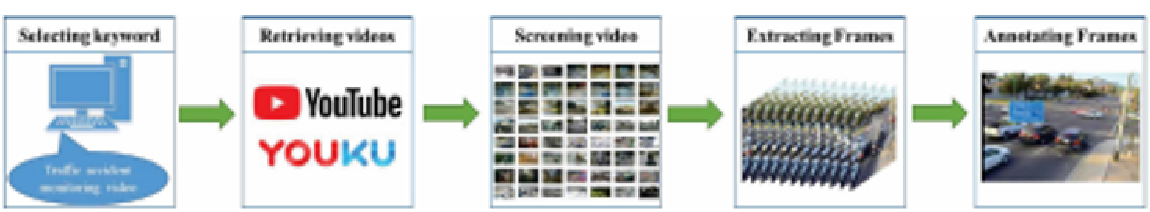
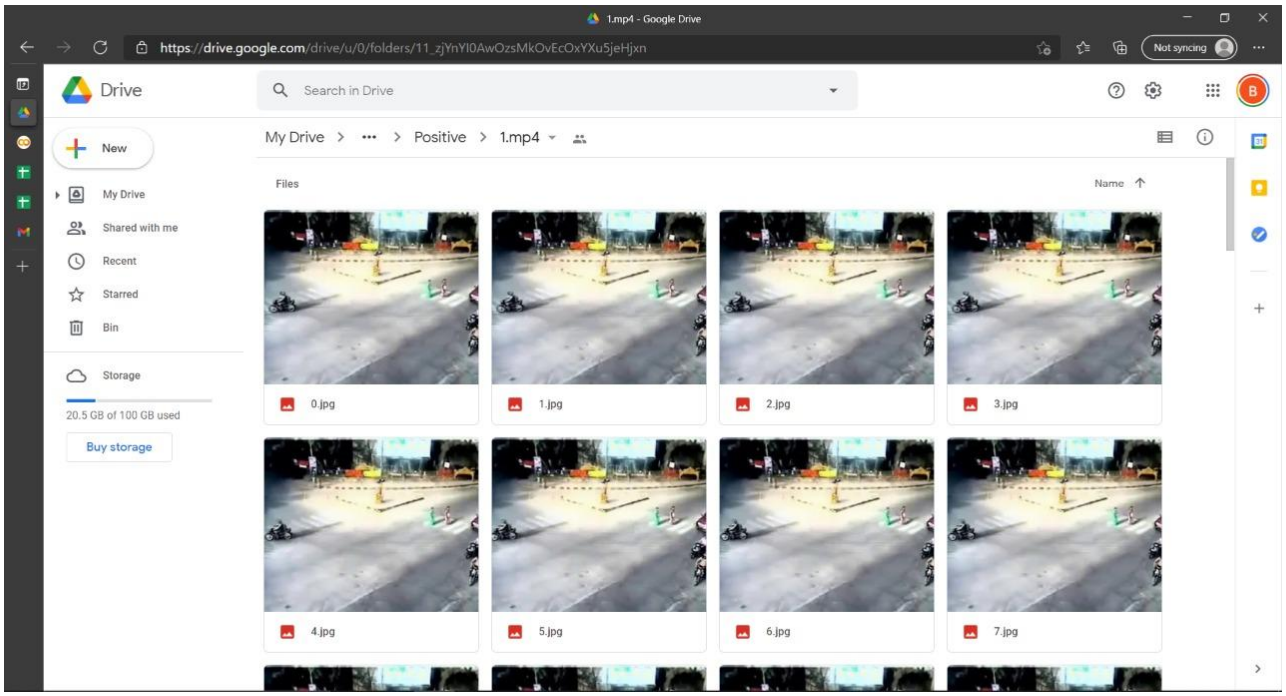


Fig: Data collection and annotation for the CADP dataset

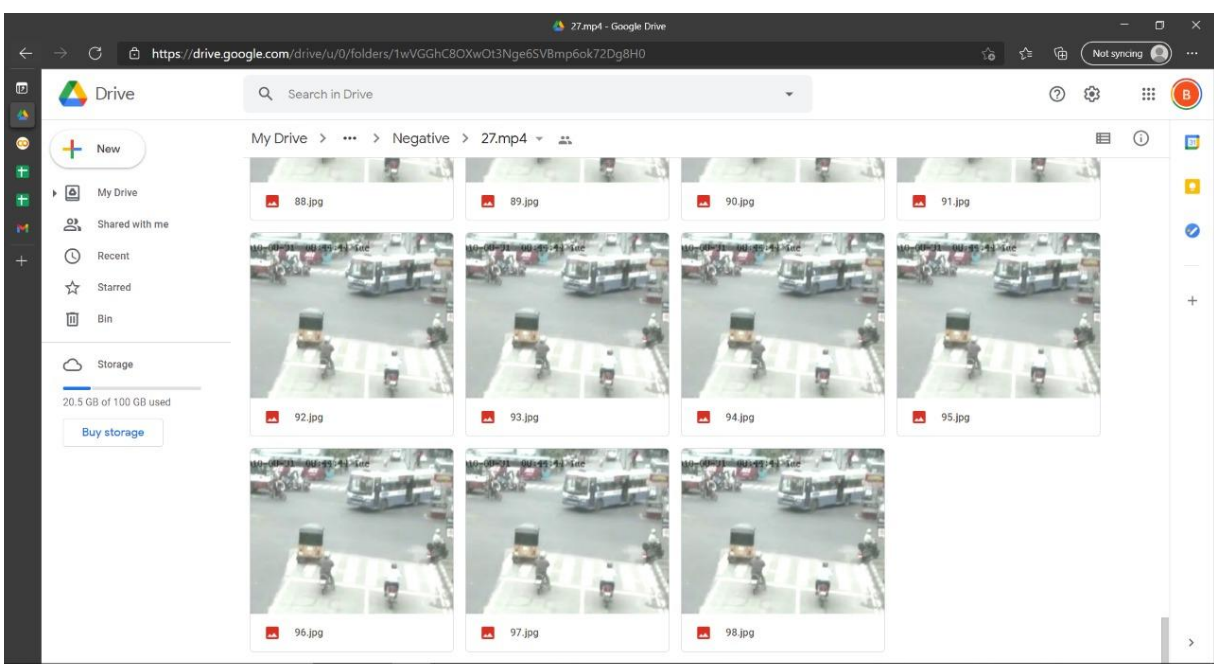
3.3 DATASET

For the final dataset, we had 232 (both positive and negative) videos with car, bus, bike etc. accidents recorded in the CCTV camera at the corners of the street. We also took the same number of negative cases (without accident) to maintain balanced classes . We store all our dataset and video frames in the Google drive .

Positive Frames :



Negative Frames :



3.4 PROCESSING THE DATA

As we see that each video is broken up into its individual frames to be analyzed separately. Each of these images is a two-dimensional array of pixels where each pixel has information about the red, green, and blue (RGB) color levels. To reduce the dimensionality at the individual image level, we convert the 3-D RGB color arrays to grayscale. Additionally, to make the computations more tractable on a CPU, we resize each image to (144, 256) - in effect reducing the size of each image to a 2-D array of 144x256.

Some of the functions used in processing of data :

- Horizontal Flip
- Frame Resize
- Conversion of image to standard form
- Make dataset function
- Frame Shaping

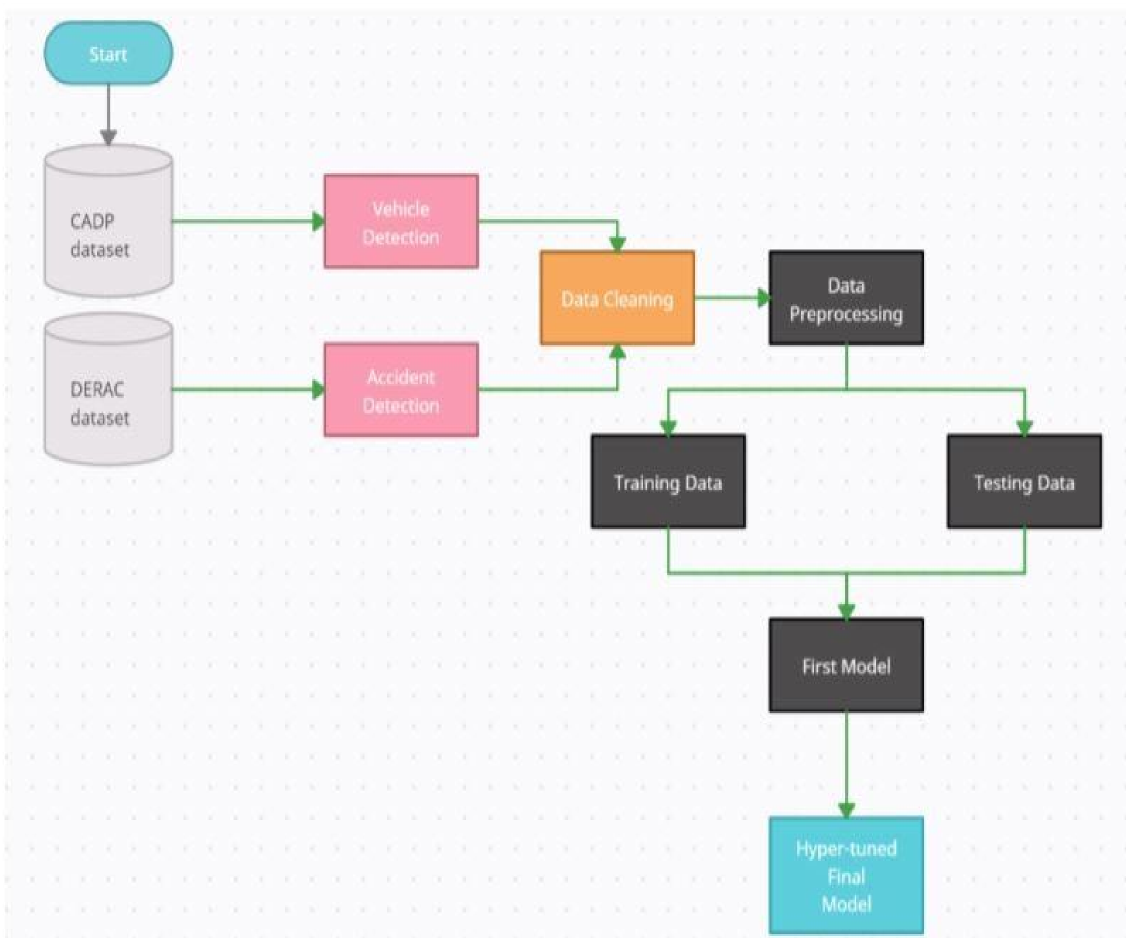
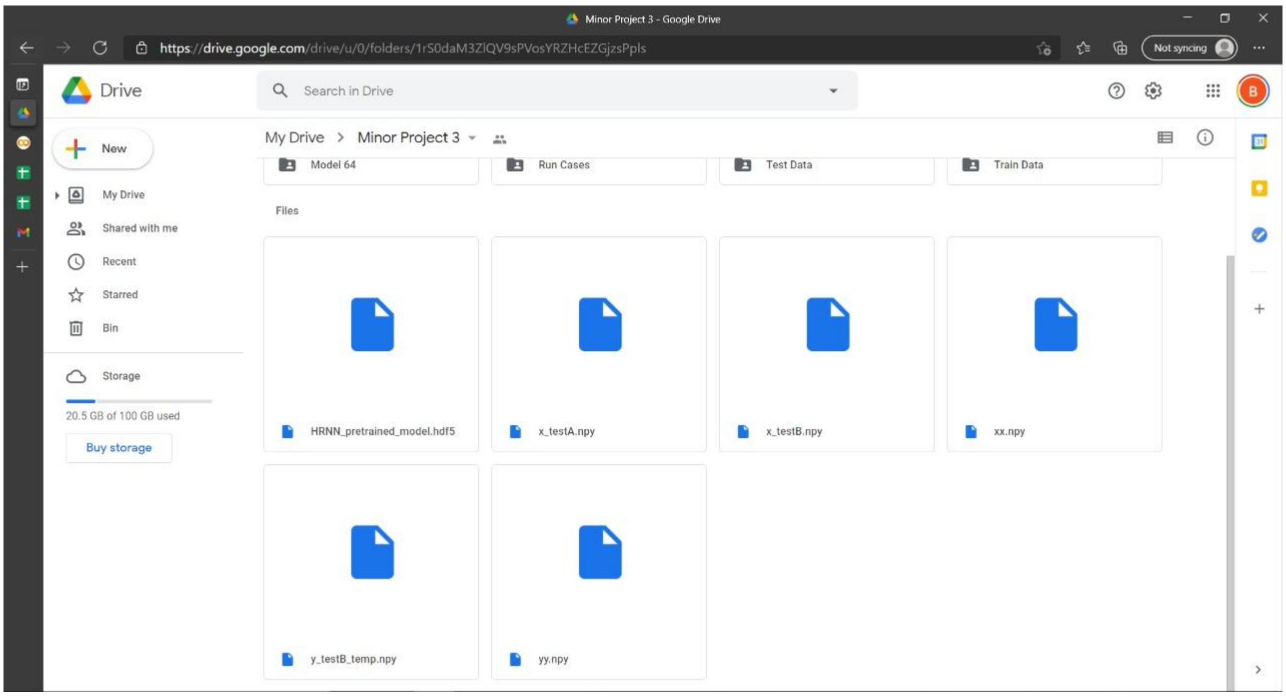


Fig: Flowchart showing Development Steps

3.5 Making Model

We used various models to obtain the best accuracy but after analysis we are good to go with HRNN.

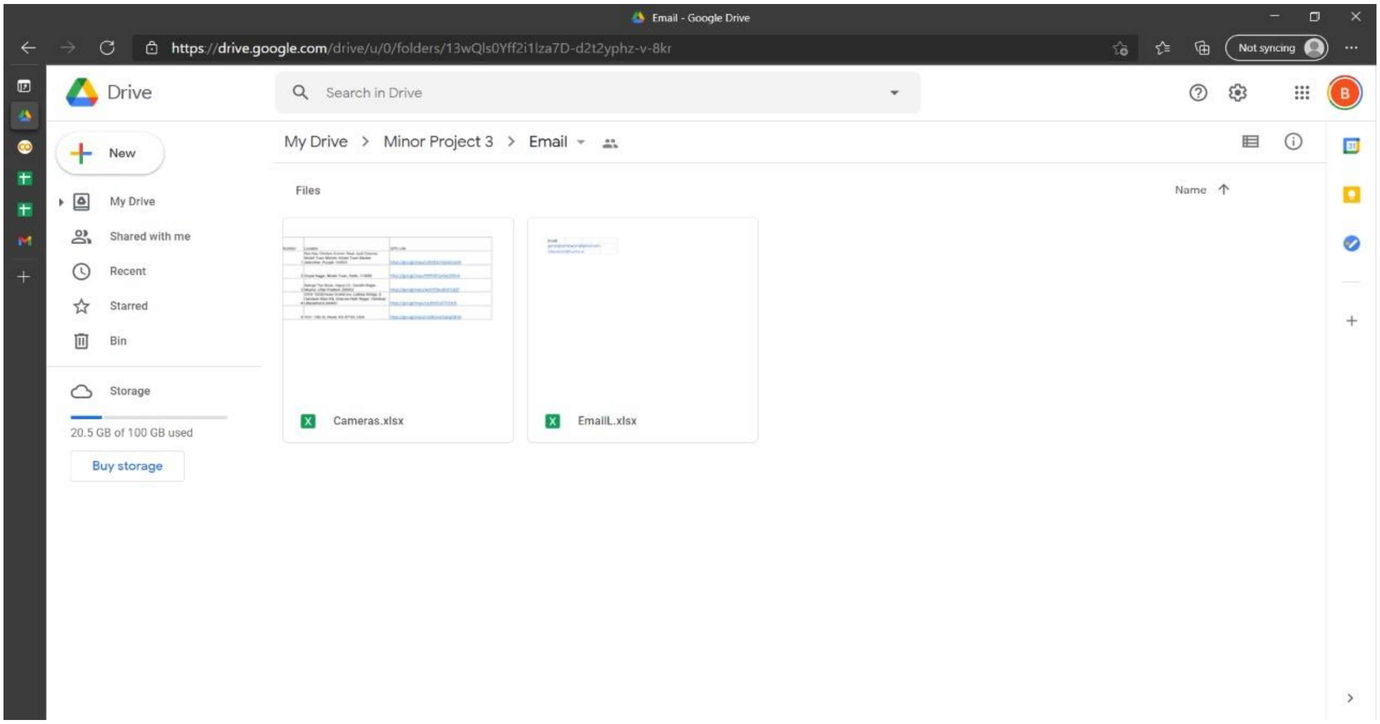
We split our data into testing and training parts .



3.6 ALERT SYSTEM

If the accident has occurred we will send the email and the coordinates of the location to the authorities via Email.

We maintained the separate database for the registered Email Id through excel file .



4. RESEARCH OBJECTIVES

The identification and reporting of automobile accidents is a difficult problem that has drawn a lot of attention from researchers. They've proposed and tested a number of approaches for detecting car accidents. In general, there are two types of automobile accident detection methods: based on the state of the vehicle and the features of the accident video

The CADP dataset has an average video length of 366 frames per video, which is 3.66 times longer than the dataset. There are 554 frames in the longest film. Road accidents with long videos have their own category in UCF-Crimes but only temporal annotations are given. The CADP dataset contains a collection of videos that have been completely annotated in terms of space and time. The number of positive videos in our dataset for just traffic accidents (1416 videos) is significantly higher than in UCF-Crimes (151 videos of road accidents) and DAD (about 600 videos). It's worth noting that there are videos in CADP that contain many accidents. We did not compile videos of negative incidents because our dataset is dedicated to traffic accidents (positive events). Negative segments may be essential for learning which is present in the other datasets DETRAC and can easily reveal the existence of negative events.

5. PROBLEM FACED

- Finding the appropriate videos for the data set
- Reshaping the frames
- Disc Storage
- Crashing of RAM while training the model .

6. Result

From CCTV Traffic Camera footage, we made the Car Accident Detection and Prediction (CADP) Dataset. The paper provides a comprehensive account of the challenges encountered during the dataset's construction, including data collection and access to traffic camera footage. On our dataset, we presented the findings of cutting-edge object detection and accident prediction models. We highlighted the strengths and shortcomings of these baseline models, and by incorporating context mining or enhanced context mining, we were able to outperform the initial findings. Finally, we demonstrated that enhanced context mining does not increase the object detection score obtained with incremental context mining. We show the final accident forecasting model, which can predict accidents about 2 seconds before they happen with an accuracy of 80%.

Info on how to do it the Tensor flow framework is used to implement our system1. Detecting artefacts at various scales of images improves efficiency during research (multi-scale testing). Until convergence, we fine-tune all object detectors in the CADP trainval range. We use the information outlined in the previous section to forecast accidents.

We chose Faster HRNN as the benchmark for further experiments based on our observations about the efficiency of these two detectors. shows the efficiency of Faster HRNN over three sampled folds. In the CADP trainval series, we can see that this detector performs consistently. However, in the "Person" group, we can see that the performances deteriorate and become dysfunctional.

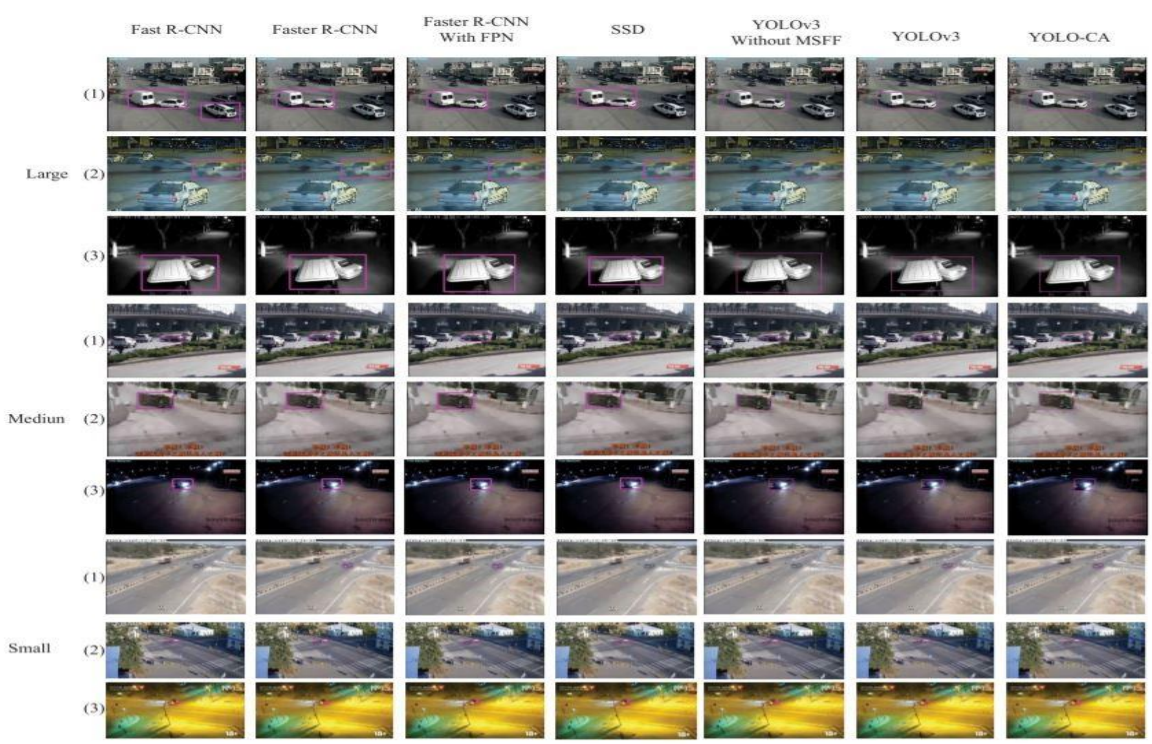


FIGURE 10. **Some visual results of the seven models among different scales of objects.**

7. Conclusion

Validation by cross-validation we took a random sample of 103 videos to train object detectors and accident predictors. The remaining 102 videos were used to put the forecasters to the test. Our decision was based on the need to build a stable model that we needed to test on a large number of samples, so we split the train and test sets 50:50 (train and test sets), with identical set statistics in terms of the number of objects in each set. To compute the accuracy of object detection, we randomly sample three folds (train/test split) from the frames of the 103 videos in the trainval series.

8. REFERENCES

1 R. Benenson, M. Omran, J. Hosang, and B. Schiele. Ten Years of Pedestrian Detection, What Have We Learned? In ECCV 2014 Workshops, pages 613–627, 2014.

2 S. R. E. Datondji, Y. Dupuis, P. Subirats, and P. Vasseur. A Survey of Vision-Based Traffic Monitoring of Road Intersections. IEEE Transactions on Intelligent

3 R. Girshick. Fast R-CNN. In IEEE International Conference on Computer Vision, 2015.

4 . Girshick, J. Donahue, T. Darrell, and J. Malik. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pages 580–587. IEEE, jun 2014.

5 The National Safety Council. The state of safety - a state-by-state report, 2017.

6 C. Vondrick, D. Patterson, and D. Ramanan. Efficiently Scaling up Crowdsourced Video Annotation. International Journal of Computer Vision, 101(1):184–204, jan 2013.

7 C. L. Zitnick and P. Dollar. Edge Boxes: Locating Object ´ Proposals from Edges. In European Conference on Computer Vision, pages 391–405, 2014.

8 S. Lyu, M.-C. Chang, D. Du, L. Wen, H. Qi, Y. Li, Y. Wei, L. Ke, T. Hu, M. Del Coco, et al. Ua-detrac 2017: Report of avss2017 & iwt4s challenge on advanced traffic monitoring. In Advanced Video and Signal Based Surveillance (AVSS), 2017 14th IEEE International Conference on, pages 1–7. IEEE, 2017.

9 Y. A. LeCun, Y. Bengio, and G. E. Hinton. Deep learning. Nature, 2015.