International Journal of Scientific Research in Science, Engineering and Technology
Print ISSN: 2395-1990 | Online ISSN : 2394-4099 (www.ijsrset.com)
doi : https://doi.org/10.32628/IJSRSET23102120

Accident Prone System Using YOLO

Harsh Vyas[1], Samarth Sharma[2], Harshil Senghani[3], Dr. Ajaysinh Rathod[4], Dr. Avani Vasant[5]
[1,2,3]Sutdents, Computer Science and Engineering Department, BITS EDU Campus, Vadodara, Gujarat, India
[4]Professor, Computer Science and Engineering Department, BITS EDU Campus, Vadodara, Gujarat, India
[5]Head of Department, Computer Science and Engineering Department, BITS EDU Campus, Vadodara, Gujarat, India

ARTICLEINFO

Article History:

Accepted: 10 April 2023
Published: 06 May 2023

Publication Issue
Volume 10, Issue 3
May-June-2023

Page Number
09-16

ABSTRACT

Accident Prone System is an accident detection model with an object detection algorithm as its backbone. Object detection algorithms are an integral part of the deep learning. The proposed system aims for optimal automatic post-accident recovery, by deploying the latest open-source computational technology at hand, in the surveillance and dash cameras to detect accidents in real time. Attempts have been made previously where algorithms such as clustering, deep neural networks and Regional CNN have been used to create accident detection models but either they weren't able to achieve efficiency or real time detection speed or both. The proposed system uses the latest algorithm at hand and a comparative study is presented by implementing accident detection models with algorithms such as Single Shot Detector (SSD) and You Only Look Once (YOLO) which are way faster than traditional algorithms and also much efficient than its predecessors. Thus, the proposed system can be deployed for real time accident detection and help save life by faster post-accident recovery.
Keywords: Artificial Intelligence, Machine Learning, Object detection, Accident Prone System, Road Traffic Crash, Deep Learning

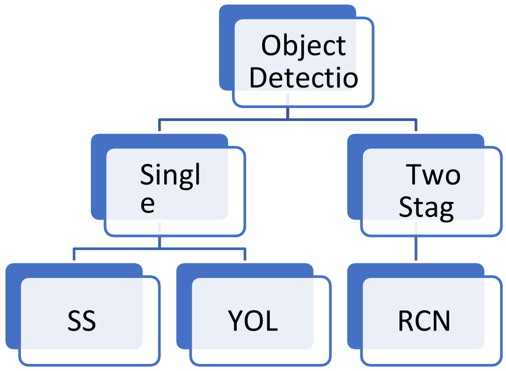
I. INTRODUCTION

The number of casualties due to road traffic accidents are increasing over years across the globe. Every year, lives of approximately 1.3 million people are cut short as a result of a road traffic crash. Between 20 and 50 million people suffer non-fatal injuries, with many incurring a disability as a result of their injury. Road traffic injuries cause considerable economic losses to individuals, their families, and to nations as a whole. Road traffic injuries are the leading cause of death for children and young adults aged 5-29 years. The prime objective of Accident-Prone System is based on detecting accidents and informing the emergency services like police, ambulance and fire brigade on the spot so that more and more families could be saved, especially when accident include serious injuries. Road accidents occur for several reasons including drink and drive and rash driving. Although many researches focused on preserving road accidents have been done using Artificial Intelligence and Machine Learning, today most of the cameras used haven't implemented the system for use in real-time. Only a few proposed solutions can be used in action due to their fast response time, however, they require a better grade camera and lack accuracy, such as a car may be predicted as truck and, in some circumstances, accidents may not be predicted. One of the highly dangerous acts would be if accidents are not predicted by the system. Few wrong predictions from the system may lead to heavy losses to family and nation as a whole. According to authors of [18], [19] and [2], the proneness of road accidents depends on 2 main factors. These factors are:

1. Road factors, related to road conditions, traffic etc.
2. Human factors, related to characteristics of population.

Current technology can be utilized as a crucial aspect for these issues. YOLOV5 algorithms are major advancement in object detection algorithm and provide better accuracy in least amount of time as compared to other algorithms. YOLOV5 being advanced version of YOLO can make accident detection more reliable and faster which enables it to be used in real life cameras. As a consequence, it can also be deployed on areas with extensive traffic or where multiple accidents are encountered. In this paper we will present experimental comparisons of 4 different object detection algorithms (YOLOV5s, YOLOV5m, YOLOV4 and SSD) and results of YOLOV5m and YOLOV5salgorithms. The main goal is to use the manually made APS dataset which contains 510 images of both accidents and non-accidents and observe mAP (Mean Average Precision/Accuracy), training time and testing time of all 4 algorithms. In this way, the best performing algorithm can be utilized to reduce bottlenecks caused by road accidents.

II. LITERATURE REVIEW



Object Detection

Object recognition is a computer vision technique for finding instances of objects in images or videos. Object detection algorithms typically use machine learning or deep learning to produce meaningful results. The main purpose of object detection is to identify and identify one or more effective targets from still image or video data. It comprehensively covers various important technologies such as image processing, pattern recognition, artificial intelligence, and machine learning [5].

Object detection algorithms are of 2 types:

1. Single Stage Detectors and
2. Two Stage Detectors.

Single Stage Detectors

A one-stage object detection model refers to a class of object discovery models that skips the region suggestion phase of a two-step model and performs direct detection on densely populated samples [4].

There are 2 major algorithms using Single Stage Detectors:

1. You Only Look Once (YOLO)
2. Single Shot Detector (SSD)

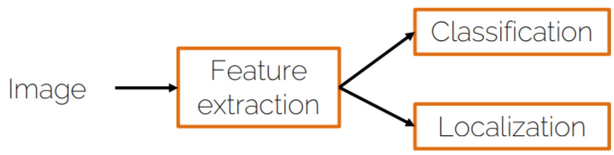


Figure 2: Single Stage Detector Working

Two Stage Detectors

The procedure is divided into two stages. In the first stage, features are extracted from the image and a region of interest (ROI) is suggested. The ROI consists of boxes that may allow you to place objects within the image. The second step uses features and ROI to calculate the final bounding box and class probability for each box [5].
One main algorithm using Two Stage Detector is Region-based Convolutional Neural Network (R-CNN).

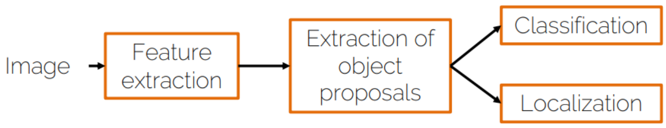


Figure 3: Two Stage Detector Working

SSD MobileNet
The Single Shot Detector (SSD) Object detection model utilizes Mobilenet as backbone. Hence it only takes single shot to detect multiple objects within the image. Thus, SSD is much faster compared to 2 shot (Regional Proposal Network) RPN- based approach [16]. The problems with this approach are as follows:
• The frameworks biggest drawback is that its performance is directly proportional to object size. This is because small objects may not contain useful information hence this approach lags in comparison to other approaches such as R-CNN.
• The next drawback is that the forward pass is controlled by the backbone network hence the slower inference time.

Faster R-CNN
R-CNN stands for Region-based Convolution Neural Network. Object detection primarily consists of two separate tasks that are classification and localization [15]. Now, in R-CNN the key concept is region proposals. They are used to localize objects within an image. The problems with this approach are:
• It takes huge amount of time to train as it classifies region proposals per image.
• It cannot be implemented in real-time system. Faster R-CNN is a Deep Convolution Network that is used for object detection. It uses region proposal method to create the sets of regions similar to R-CNN. In order to solve some drawbacks of the R-CNN, fast R- CNN feeds the input image to the CNN to generate a convolutional feature map instead of feeding the region proposals to the CNN.
The Fast R-CNN is faster than R-CNN because we don't have to feed the region proposals to the convolutional neural network every time. Instead, the convolution

operation is done only once per image and a feature map is generated from it.
The drawback of this approach is that for the Region Proposal Network, all the anchor box are extracted from a single image. The network may take a lot of time to reach convergence as all the samples from a single image can have similar features and be correlated.
YOLO
YOLO is an algorithm that utilizes neural networks to provide real-time object detection. This algorithm's strength is its speed and accuracy [6]. It stands for You Only Look Once and has a simple architecture. It is trained to do classification and bounding box regression at the same time unlike the R-CNN and Faster R-CNN.
During the research, Single stage detector algorithms were used for implementation due to its advantages of speed and accuracy over two stage or multistage detector algorithms.

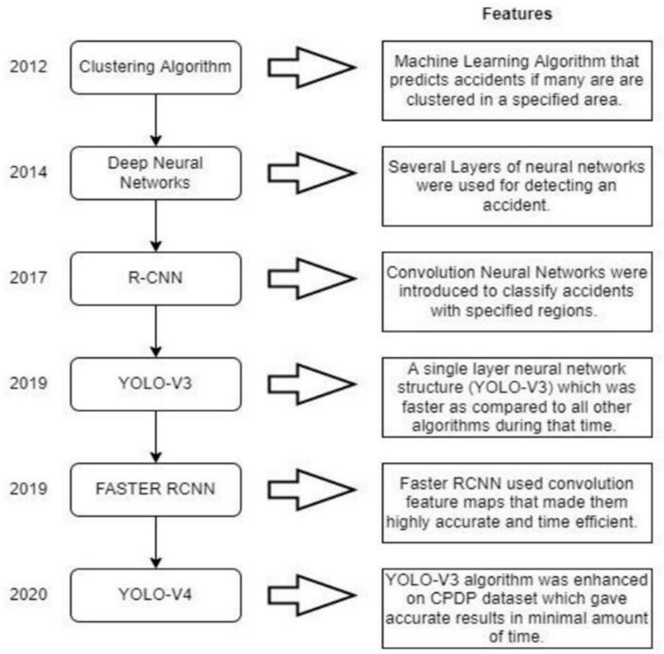


Figure 4: Evolution of technology related to the Accident- Prone System

Table 1: Comparisons of different Algorithms used for Accident proposed system.

Object Detection Algorithms	Number of stages	Accuracy	Speed/Frame (ms/frame)	Advancements	Disadvantages	Real-Time Usability
Clustering[10]	2	Low	Slow	The Initiative	High False Alarm Rate	✗
Deep Neural Network[22,11]	2	Low	Slow	Can perform image classification	Cannot be used for object detection	✗
RCNN[8,4]	2	77.4	2000	Step by step approach: finding the Higher probability region first and then classification	Very time consuming thus not suited for the real-time detection	✗
Faster RCNN[15,1,4]	2	69.4	200	10x faster than the RCNN	Accuracy, precision and recall dropped	✗
SSD[21,16]	1	56.4	39	5x faster than even the Faster RCNN	Accuracy, precision and recall dropped further	✓
YOLO v3[7,6]	1	60.6	29	Increased Speed of detection and accuracy	Metrics of the model thought better, not up to the mark for real-time systems	✓
YOLO v4[17,13]	1	83	34	Exemplary growth in accuracy with consistent speed	Lower Recall	✓

Table 1 elicits the comparison of 7 different object detection algorithms (Clustering, Deep neural network, RCNN, Faster RCNN, SSD, YOLO v3 and YOLO v4) with their respective speed, accuracy, enhanced features and drawbacks. The observations have been imitated from prior testing on COCO dataset.

Overall, only YOLO v3, YOLOv4 and SSD can be used in real time system because of their testing speed, however, it is also seen that all of these algorithms lack accuracy. Moreover, single stage detection algorithms outperformed multi stage detection algorithms in terms of speed.

III. PROBLEM STATEMENT

After the literature review, it was observed that on one hand there were models with good object classification accuracies but slower object detection speed and on the other hand there were algorithms which were faster but compromised on the object classification accuracies. So, the task was to build a model which can carry out object detection seamlessly and without compromising the classification accuracy. So, implementation of YOLOv5 for accident detection and classification was the task.

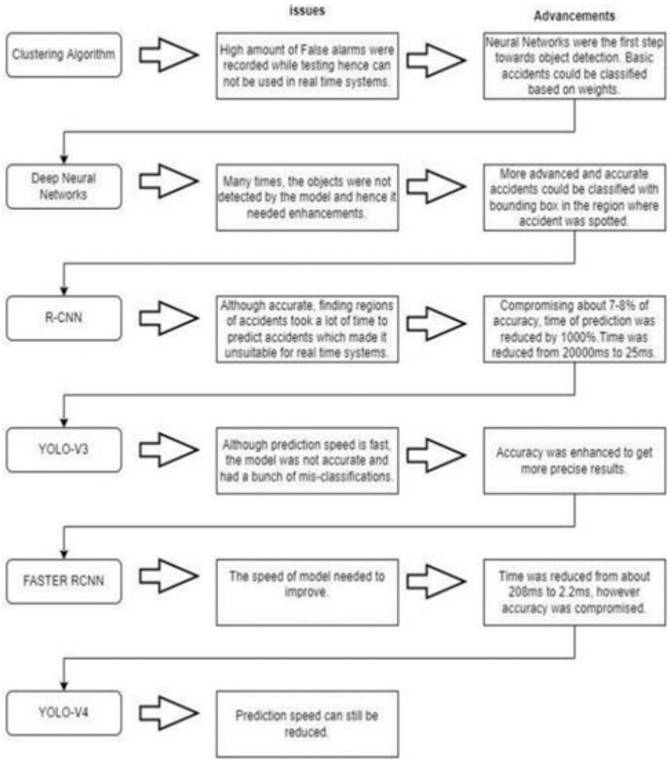


Figure 5 : Issues and improvements related to the Accident- Prone System

IV. PROPOSED SOLUTION

Although the project has been previously implemented using prior technologies, none of the given solution was used in real-time due to lack of either accuracy or detection time or both. Using the latest technology of YOLOv5, the model can predict in real-time system as well as gives high training and testing accuracy.

V. METHODOLOGY

There are two mainstream approaches that prevail for object detection:

1. Two pass approach - RCNN, Faster RCNN
2. Single pass approach - SSD, YOLO

In the two-pass approach, first pass is for determining the objectiveness regionally and the second pass is to detect object in the region with higher objectiveness with Convolution Network.

The single pass is fully convolution approach where objectiveness and detection are both done by convolutional network in a single pass.

While the algorithms with two pass approach perform well, the algorithms with single pass approach perform at par and faster. Thus, algorithms with single pass approach have an edge.

YOLO v5 Architecture:

The YOLO network consists of three main pieces:

1) Backbone: The role of Backbone is feature extraction. It is a CNN which extracts features at different granularities.

2) Neck: The role of neck is to generate feature pyramids which helps generalize the model better.

3) Head: The role of head is to do the final detection part with the input it received from the neck. Anchor box is used to predict the bounding box, objectiveness.

CPS – Cross Stage partial Networks are used as backbone for YOLOv5 models.

PANet- Path Aggregation Network is used to generate feature pyramids for the YOLOv5 models.

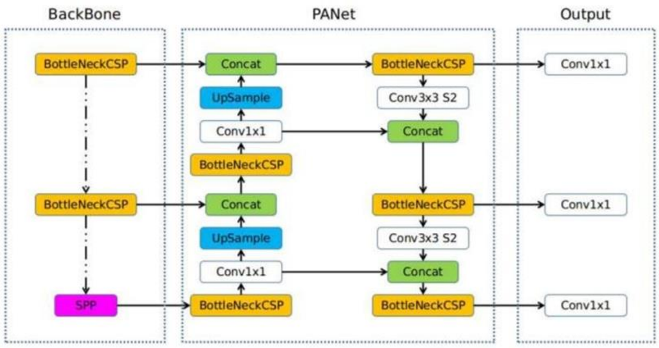


Figure 6: YOLOv5 Architecture [22]

VI. DATASET

A custom-made dataset was prepared from 34 videos of 5 seconds each. FFMPEG software was used during the training and testing phase to convert videos to images/frames. It took 'frames per seconds' and the 'path to image directory' as inputs and returned array of images corresponding to that video.

For images in the dataset, a 5 second video was tripped at 3 frames per second using the ffmpeg software resulting in an array of 15 frames per video.

The frames were then annotated using the LabelImg tool. The format of annotations of image for YOLO is .txt.

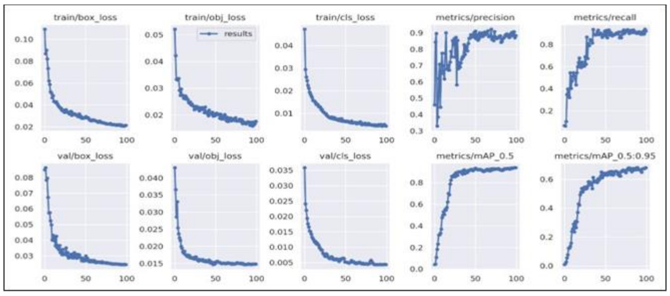


Figure 7: YOLO .txt format

Figure 7 shows .txt format file used for YOLO- V5 algorithm. The file contains class name (represented by unique number for every class) followed by bounding box dimensions or anchors (bx: x co- ordinate, by: y co-ordinate, bw: width of annotation, bh: height of annotation)

After annotating the images of dataset manually, the dataset containing images with their corresponding annotation files were uploaded to Roboflow platform which helped in splitting dataset into train, test, and validation dataset (in ratio 80:10:10 respectively). Furthermore, it made importing the dataset on Google Collaboratory (cloud platform) simple, easily accessible and gave structured hierarchical format to the data.

VII. IMPLEMENTATION

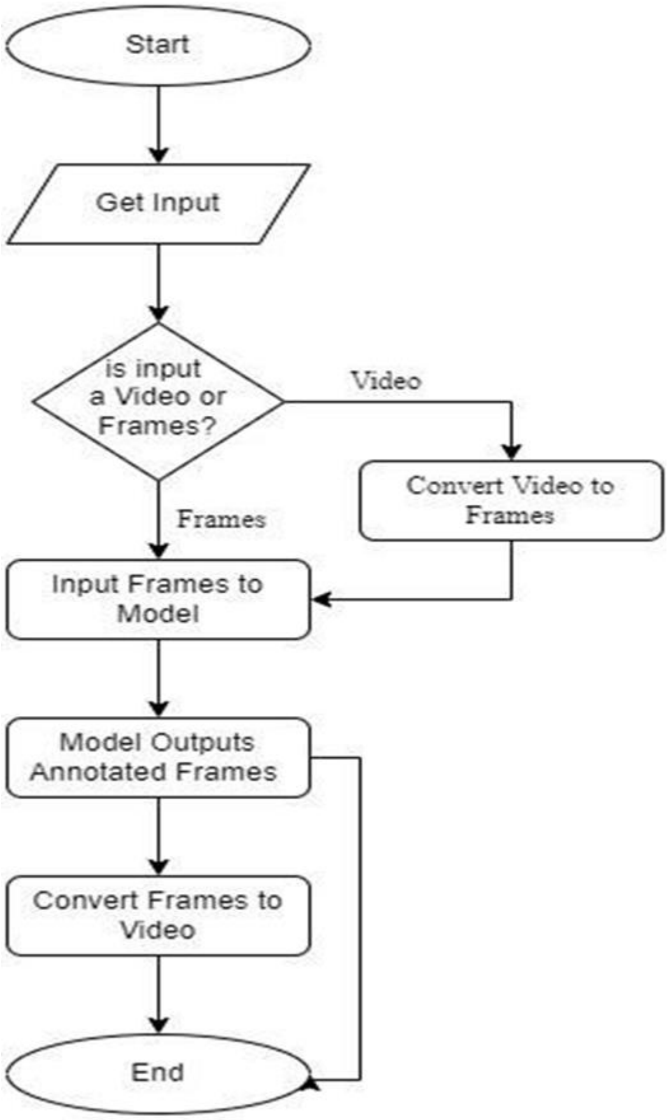


Figure 8: System Flowchart

The figure 8 contains system flowchart which shows the working of final model. The code of model has two ways of taking input:

1. Live video from camera and
2. Take video or frames stored in local memory.

User can run code block which they want to run. If video is chosen then it automatically gets converted into frames using ffmpeg tool (for local memory). Then the frames are given to the model to process. The resulting annotates frames are provided as a result to the end user. The frames can also be converted to video by running alternate code-block provided.

YOLOv5s and YOLOv5m were implemented using the custom-made dataset on Google Colab. Also, YOLOv4 and SSD were implemented to do a comparative study.

VIII. RESULTS

YOLOv5s:

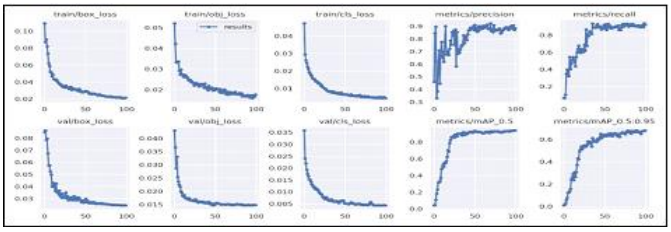


Figure 9: Performance of YOLOv5s

Overall, all the training and validation loss were reduced tremendously which affected training output of the model. On the contrary, values of precision, recall and accuracy reached peak by end of the training.

YOLOv5m:

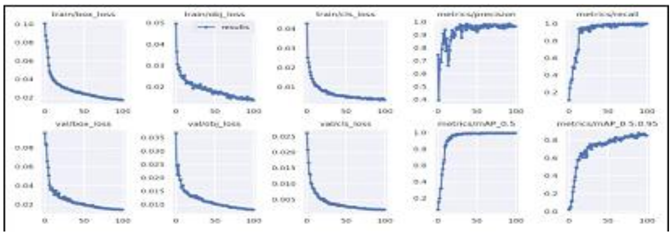


Figure 10: Performance of YOLOv5m

Overall, the output did not favour expectations for about initial 10 epochs however the expected performance started to occur. The values of precision, recall and accuracy increased gently.

Below is the summarization of the validation/testing process, which will help us visualize the difference quantitatively.

Table 2: Validation and Training accuracy of different object detection algorithms

Model Name	Validation Accuracy (%)	Training Accuracy (%)
SSD	69.69	76.73
YOLO-V4	69.9	97.1
YOLO-V5s	68	93.9
YOLO-V5m	85.2	99.4

Table 3: precision and recall of different object detection algorithms

Model Name	Precision (%)	Recall (%)
SSD	76.7	79.6
YOLO-V4	65.4	99.9
YOLO-V5s	88.2	92.4
YOLO-V5m	96.4	99.9

The table 2 and table 3 illustrate the observations about Precision, Recall, training time and Accuracy (Mean Average Precision), recorded during the training phase of 4 different algorithms (SSD, Faster, YOLO-V4, YOLO-V5s and YOLO-V5m).

It was observed that YOLO-V5 algorithm outperforms other object detection algorithms in terms of precision, recall and accuracy. SSD had the least accuracy however had best training time.

IX. CONCLUSIONS

YOLOv5s and YOLOv5m were successfully implemented and they outperformed their predecessors. YOLOv5m gave the optimal results and thus a model was created which is not only accurate but faster than any of its predecessors and thus can be used in real time situations.

The models only limitation is the obstruction of the camera and the detection of faraway accidents that cannot be properly seen or observed through the available cameras. The system can be enhanced through the design of an alert system to mitigate nearby emergency services on the detection of the accident and by utilizing better and bigger dataset to get more concise results.

X. REFERENCES

[1]. Ankit Shah, Jean Baptiste Lamare, Tuan Nguyen Anh, Alexander Hauptmann, “CADP: A Novel Dataset for CCTV Traffic Camera based Accident Analysis”, Nov 2018

[2]. Feasibility Study using Improved Image Segmentation, Machine Learning and Sensors”, 2019

[3]. Government of India, “Road accidents in India”, https://morth.nic.in/sites/default/files/Road_Ac cid ednt.pdf, 2018.

[4]. https://github.com/ultralytics/yolov5/issues/280

[5]. https://machinethink.net/blog/object-detection

[6]. https://medium.com/codex/a-guide-to-two-stage- object-detection-r-cnn-fpn-mask-r-cnn-and-more- 54c2e168438c

[7]. https://pjreddie.com/media/files/papers/YOLO9 0 00.pdf

[8]. Joseph Redmon, "You Only Look Once: Unified, Real- Time Object Detection", 2016

[9]. Ming Liang, Xiaolin Hu, "Recurrent Convolutional Neural Network for Object Recognition", TNList, 2015

[10]. Muhammad Babar Ali Rabbani, Muhammad Ali Musarat, Wesam Salah Alaloul, Ahsen Maqsoom, Hamna Bukhari, Waqas Rafiq, “Road Traffic Accident Data Analysis and Its Visualization”, National University of Sciences and Technology, Islamabad, Pakistan, July 19, 2021

[11]. Nejdet Dogru, Abdulhamit Subasi, “Traffic Accident Detection by Using Machine Learning Methods”, International Burch University, Sarajevo, Bosnia and Herzegovina, March, 2012

[12]. Renu, Durgesh Kumar Yadav, Iftisham Anjum, Ankita. "Accident Detection using Deep Learning: A Brief Survey", IJECCE, 2020
[13]. Road traffic accidents in India 2019, GOI
[14]. Rohit Hooda, Rishabh Kawediya, Krish Sonigara, Snehal Andhare, "Accident Detection System using Surveillance Camera's", IRJET, Apr 2020
[15]. Sanjay Kumar Singh," Road Traffic Accidents in India: Issues and Challenges", World Conference on Transport Research - WCTR 2016Shanghai. 10-15 July, 2016.
[16]. Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks", Microsoft Research, 2017
[17]. Shifeng Zhang, Longyin Wen, "Single-Shot Refinement Neural Network for Object Detection", 2018
[18]. Shilpa Jahagirdar, Sanjay Koli, "Automatic Accident Detection Techniques using CCTV Surveillance Videos: Methods, Data sets and Learning Strategies", IJEAT, Feb 2020
[19]. Sodikov, Jamshid, "Road Traffic Accident Data Analysis and Visualization in R", International Journal of Computer Science Engineering and Information Technology Research (IJCSEITR), 2018
[20]. The World Health Organization Global Status Report on Road Safety,The World Health Organization; Geneva, Switzerland, 2018
[21]. Vipul Gaurav, Sanyam Kumar Singh, Avikant Srivastava, "Accident Detection, Severity Prediction, Identification of Accident-Prone Areas in India and
[22]. Wei Liu, Dragomir Anguelov, "SSD: Single Shot MultiBox Detector", 2016
[23]. Zhong-Qiu Zhao, Peng Zheng, Shou-tao Xu, and Xindong Wu,"Object Detection with Deep Learning: A Review", IEEETransactions on Neural Networks and Learning Systems, April 16, 2019

Cite this article as :

Harsh Vyas, Samarth Sharma, Harshil Senghani, Dr. Ajaysinh Rathod, Dr. Avani Vasant, "Accident Prone System using YOLO", International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Online ISSN : 2394-4099, Print ISSN : 2395-1990, Volume 10 Issue 3, pp. 09-16, May-June 2023. Available at doi : https://doi.org/10.32628/IJSRSET23102120
Journal URL : https://ijsrset.com/IJSRSET23102120