Intelligent Video Surveillance System

Rutuja Rasam[1], Shubhangi Yadav[2], Sakshi Chavan[3], Komal Bijore[4]

B.E. Computer Engineering Department,

Prof. Shudhodhan Bokefode, Department of Computer Engineering.

TPCT's Terna Engineering College, Nerul, Navi Mumbai, Maharashtra 400706.

Email: shudhodhan358@gmail.com, rutujarasam@ternaengg.ac.in, shubhangiyadav@ternaengg.ac.in, sakshichavan@ternaengg.ac.in, komalbijore@ternaengg.ac.in

Abstract - *The frequent increase in number of vehicle crashes and accidental falls has increased over the past decade. Increasing the number of cars and excessive traffic in cities is a major problem at the current time. Statistics show that more and more accidents happen daily, and many of these could be avoided. Approximately 1.3 million people die each year as a result of road traffic crashes. Regardless of mindfulness cause, this issue is as yet expanding because of rider's poor practices, for example, drunk driving, speed driving, riding without adequate rest, riding with no cap insurance, and so forth. Accidental fall is also the most prominent factor that causes accidental death of people.*

This paper uses the YOLOv3 algorithm to extract features and complete the end-to-end prediction of the target frame and category at one time. In this paper, an already trained model of the coco dataset based on the YOLOv3 algorithm is used for fall and vehicle crash detection. The YOLO machine learning algorithm uses features learned by a deep convolutional neural network to detect an object. Experiments show that the YOLOv3 algorithm is a real-time object detection algorithm that offers excellent detection speed and detection accuracy.

Keywords - *YOLO, YOLOV3, Fall Detection, Vehicle Crash Detection, Coco trained model.*

I. Introduction

Vehicle Crash and fall accidents are the most common and major issues in our daily life. Due to many relative factors, the victims cannot acquire medical assistance in time, causing the best rescue time to be missed. Therefore, more and more attention has been paid to the research on fall and vehicle crash detection. Accurate and rapid detection of these accidents can effectively avoid casualties. Many technologies have been used to date in fall detection research and vehicle crash detection.

In this paper, the main aim of our proposed surveillance system is to provide immediate help to the victim after the detection of a Fall and Vehicle crash. Our system uses the YOLOv3 algorithm, a version of the YOLO(You Only Look Once) deep learning algorithm to perform the detection along with an already trained model based on the Coco dataset.

Fall Detection:

A fall in this study is defined as an event in which a person suddenly and inadvertently collapses from an upright position and stays in that position for a long time. Worldwide, falls are a leading cause of deaths or unintentional injuries in adults older than 65 years old, with 37.3 million falls requiring medical attention and 646,000 resulting in deaths annually. Seniors living alone are at high risk. Many common neurological problems result in falls: Peripheral neuropathy, manifesting with numbness and imbalance, spinal stenosis, resulting in pain and in-coordination, acute strokes, leading to sudden weakness, and Parkinson's disease, characterized by postural instability, etc. In addition, cardiovascular, musculoskeletal, and medication-induced problems often coexist. Orthostatic hypotension, knee arthritis, and iatrogenic dizziness are only a few examples. Even for

healthy seniors, activities such as climbing ladders, taking showers, going downstairs, and walking in snow could be dangerous. Falls are not exclusively problems for the elderly, but may also be a concern for the young. Postural Orthostatic Tachycardia Syndrome, seizures, anemia, pregnancy, and sports all can lead to unexpected falls. Without timely detection and treatment, complications such as bone fractures, intracranial hemorrhage, or nerve avulsion can result. Permanent disabilities and death are not unusual. One out of five falls gives rise to a serious injury such as a hip injury, broken bones, or a head injury. Close to 95% of hip injury is a consequence of a fall, usually falling sideways. Falls affect a large number and variety of people every year.

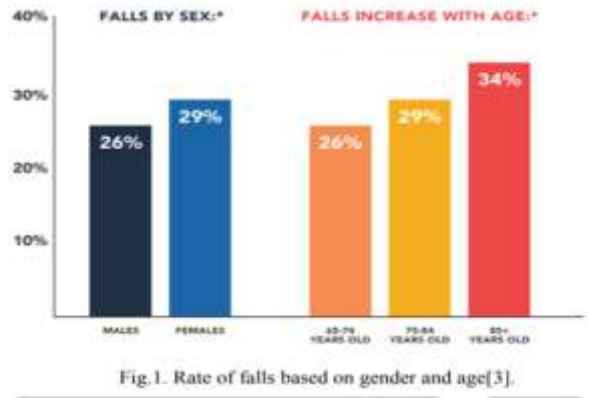

Fig. 1. Rate of Falls

Fall detection systems have different categories. In this project, we are going to design a camera-based fall detection system. Computer vision makes use of cameras to track the user's movements. A fall may be detected when the user is inactive for a long time. The advantage of the system is that it can detect multiple events simultaneously. The multiple-camera network reconstructs a 3D image, analyzes the volume distribution of the individual along the vertical axis, and triggers an alarm when most of the volume is near the floor for a predefined period.

Vehicle Crash Detection

Vehicle accidents have consistently had serious consequences, including loss of life and property. Vehicular Traffic has become a substantial part of people's lives today and it affects numerous human activities and services daily. Statistically, nearly 1.25 million people forego their lives in road accidents on an annual basis with an additional 20-50 million injured or disabled. Road traffic crashes ranked as the 9th leading cause of human loss and account for 2.2 percent of all casualties worldwide. In recent times, vehicular accident detection has become a prevalent field for utilizing computer vision to overcome this arduous task of providing first-aid services on time without the need for a human operator for monitoring such events.

In this project, we are proposing a video surveillance system that will be camera-based. Our approach is suitable for real-time accident conditions which may include daylight variations, weather changes, and so on. Our parameters ensure that we can determine discriminative features in vehicular accidents by detecting anomalies in vehicular motion that are detected by the framework. Additionally, we are adding immediate help facilities and other features. Here we are providing help to the victims by immediately informing nearby help stations.

Around the world, the deaths and injuries of people due to traffic accidents and falls can be minimized. An influential indicator of survival rates after detecting the accident is the time between the occurrence of the accident and the arrival of emergency responders to the scene. Reductions in this time, in turn, may affect the number of fatalities, and this is achieved through using automatic accident and falls detection and notification systems which are intelligent surveillance available on the roads and other places.

II. Literature Survey

Dhananjai Chand, Savyasachi Gupta, and Goutham K. proposed a system that is a computer vision-based framework to detect vehicular collisions. In this paper, a neoteric framework for the detection of road

accidents is proposed. The proposed framework capitalizes on Mask R-CNN for accurate object detection followed by an efficient centroid-based object tracking algorithm for surveillance footage. The probability of an accident is determined based on speed and trajectory anomalies in a vehicle after an overlap with other vehicles. The proposed framework provides a robust method to achieve a high Detection Rate and a low False Alarm Rate on general road-traffic CCTV surveillance footage. This framework was evaluated on diverse conditions such as broad daylight, low visibility, rain, hail, and snow using the proposed dataset. This framework was found effective and paves the way for the development of general-purpose vehicular accident detection algorithms in real-time.

Ying-Nong Chen, Chi-Hung Chuang, Hsin-Min Lee, Chih-Chang Yu, and Kuo-Chin Fan, state a novel fall detection mechanism based on a coarse-to-fine strategy in a dusky environment.

In this paper, the human body in a dusky environment can be successfully extracted using the thermal imager, and fragments inside the human body silhouette can also be significantly reduced as well. In the coarse stage, the optical flow algorithm is applied to thermal images. Most walk actions are filtered out by analyzing the downward flow features. In the fine stage, the projected MHI is used as the features followed by the NNFLE method to verify fall incidents. The proposed NNFLE method, which adopts a nearest neighbor selection strategy, is capable of alleviating extrapolation/interpolation inaccuracies, singular problems, and high computation complexity. Experimental results demonstrate that the proposed method outperforms the other state-of-the-art methods and can effectively detect fall incidents even when multiple subjects are moving together.

Bokefode Shudhodhan Balbhim and Harsh Mathur proposed a system that checks performance analysis of Image forgery detection using Transform Function and Machine Learning Algorithms.

This paper represents the study of image forgery detection based on transform methods and machine learning. The transform methods have great potential in image processing and pattern recognition. The various derivates of transform methods estimate the image forgery. The variants of transform include discrete wavelet transform, DCT, FFT, SIFT, and many more transforms. The applied transform methods have certain limitations and the detection of forged images is compromised. The machine learning algorithm increases the detection ratio of image forgery. The trends of machine learning algorithms focus on image forgery detection and improve the detection ratio. Machine learning provides various classification and clustering algorithms for image forgery detection. This paper analyzed the experimental performance of image forgery detection based on transform and machine learning algorithms. The analysis process uses MATLAB tools and a standard image forgery dataset for the detection ratio. The analysis of results suggests that the machine learning algorithm is very efficient instead to transform-based methods.

Dr. J. L. Mazher Iqbal and S. Heena Kousar proposed an Automatic Vehicle Accident Detection system that uses a Black Box for reporting.

The framework in the proposed system is placed in moving vehicles to detect accidents and report to In Case of Emergency (ICE). The first responder is getting notice through GSM and in addition to that the vehicle accident zone longitude and latitude information is obtained through GPS and GPRS. The proposed work can save lives as the ICE team reaches the spot on time, take care of the victim, and give medical treatment on time. The black box will record the voice of the victim after the accident occurs which will be used for further investigation.

Yanwei Yin, Liang Lei, Minghui Liang, Xiaobing Li, Yuanyuan He, and Lanyao Qin proposed a system based on the YOLO algorithm for fall detection for elderly people living alone.

In this paper, through self-built data sets, two deep learning detection algorithms of YOLOV4 and YOLOV5S are used, and network models are trained and tested on the computer by adjusting and optimizing network parameters. The experimental results show that although YOLOV4 has good characteristics after modifying the network parameters and the prior frame, compared with YOLOV5S, there are still shortcomings such as a bloated model, low frame rate and accuracy, and long training period. In general, the YOLO series algorithm is better than the general computer vision method, which can meet the real-time and accurate detection requirements and the model is easier to deploy. In the later stage, the network structure of YOLOV5 will be further optimized and better solutions will be explored, the best-trained model will be deployed on the web camera end, and the application will be promoted after verification. This is extremely far-reaching for creating a smarter, safer, and more humane home life.

III. Methodology

Agile is one of the most popular approaches to project management due to its flexibility, adaptability to change, and high level of customer input. Agile project management works on the basis that a project can be continuously improved upon throughout its life cycle, with changes being made quickly and responsively. Our project is a Machine Learning project so when we apply the Agile methodology we look to three key areas:

A. **Show progress**: In this, we keep focused on outcomes by treating the YOLO model as a product. And as a "product," it can be updated with new features and improvements. This could be easily documented and recorded using several tools for the best transparency.

B. **Show value**: Showing value also means building working software, and measuring the progress as the Agile Manifesto states. One attractive feature of Agile is that the sprint structure supports continuous delivery and can track the value added to the business. To do this, we tightly couple continuous testing and inspection throughout the project life cycle. Agile also promotes better clarity and understanding to any technical or non-technical stakeholder as it involves product backlog, sprint backlog, and release plan.

C. Pursue improvement and clarity: Applying an Agile mindset can deeply empower teams and cultivate responsibility. It provides clarity in interaction and communication. With a focus on continuous improvement and clarity, teams can greatly improve self-organization and take their skill sets to the next level. Feedback is something everyone needs to improve better. Agile allows decisions to be tested and rejected early with feedback loops, providing benefits not in any other model or methodology. These conversations are especially important in AI/ML projects. After all, we don't want to realize our model won't scale to the client's needs when we're halfway through a project.

IV. Proposed System

Algorithm: YOLOv3

YOLOv3 a version of the You Only Look Once algorithm, is a real-time object detection algorithm that identifies specific objects in videos, live feeds, or images. The YOLO machine learning algorithm uses features learned by a deep convolutional neural network to detect an object.

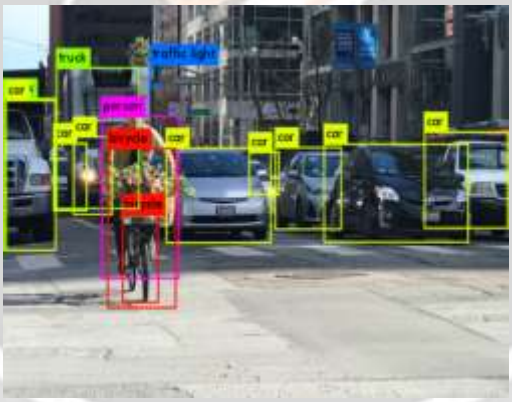

Fig.2. Image Detection

YOLO is a Convolutional Neural Network i.e. CNN that detects objects in real time. CNNs are classifier-based systems that can process input images as structured arrays of data and recognize patterns between them. YOLO does have the advantage of being significantly faster than other networks while maintaining accuracy. It enables the model to examine the entire image at test time, allowing its predictions to be informed by the image's global context. Regions are "managed to score" by YOLO and other convolutional neural network algorithms based on their similarity to predefined classes. High-scoring regions are noted as positive detection of the class to which they most closely identify. For example, in a live feed of traffic, YOLO can be used to detect different kinds of vehicles depending on which regions of the video score highly in comparison to predefined classes of vehicles.

YOLOv3 Architecture:

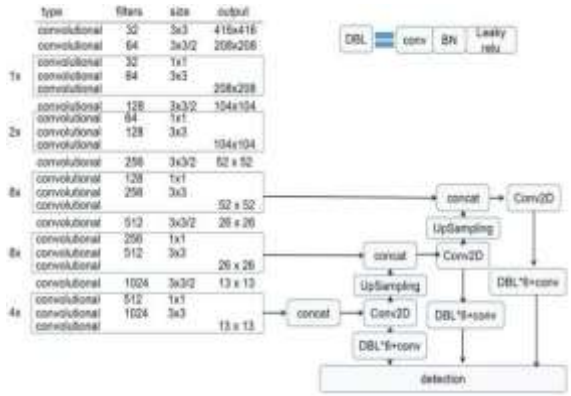

Fig. 3. YOLOv3 Architecture

Vehicle Crash Detection

Our preeminent goal is to provide a model which helps in solving the issue of traffic accidents by detecting the crashes that happen on roads and highways. Vehicle crash detection is done by using YOLO Version 3 Algorithm. The flow of the proposed system is shown below:

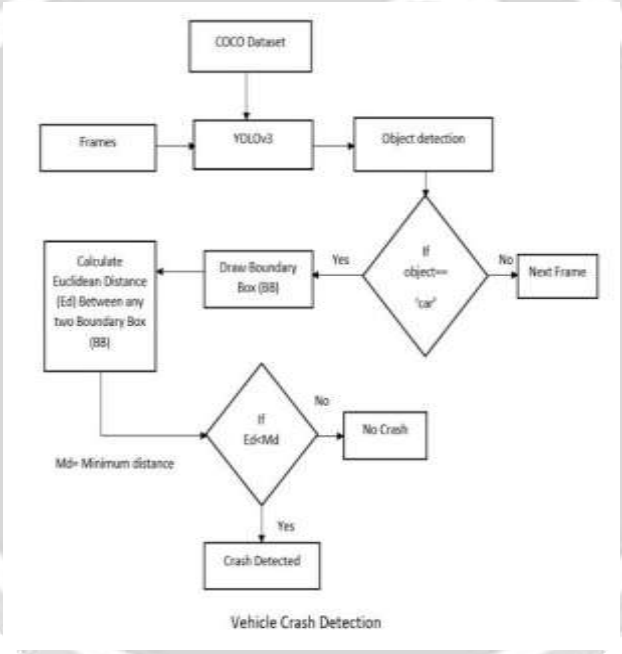

Fig. 4. Vehicle Crash Detection

(Dataflow Diagram)

The YOLO algorithm uses the trained model which is trained on the COCO dataset. For vehicle crash detection the frames are provided to the algorithm. based on training it detect the object in the object detection phase. in Vehicle crash detection specifically identifies the vehicles for example cars. If the object detected is a car then draw the boundary box around the object and calculate the Euclidean Distance between two boundary boxes of two vehicles.

$$d(x, y) = \sqrt{\sum_{i=1}^{n} (y_i - x_i)^2}$$

If the Euclidean distance is minimum then the minimum distance than a crash is detected otherwise no accident has occurred. If the object is not detected as a car or any vehicle then move to the next frame and detect the car in that frame.

Fall Detection

Our proposed model is to provide a service by detecting the fall of an elder person. The Fall detection is also done by using YOLO Version 3 Algorithm. The flow of the proposed system is shown below:

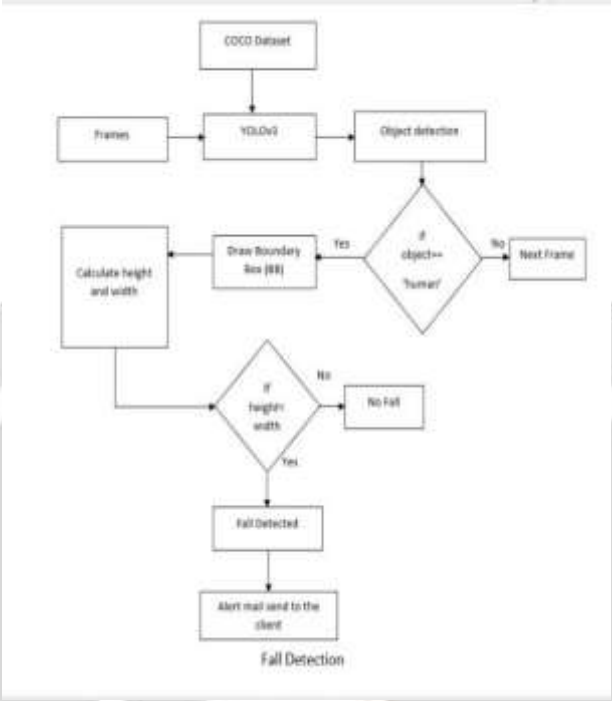

Fig. 5. Fall Detection

(Dataflow Diagram)

The YOLO algorithm uses the trained model which is trained on the COCO dataset. For fall detection the frames are provided to the algorithm. based on training it detect the object in the object detection phase. Fall detection specifically identifies the human. If the object detected is a human being then draw the boundary box around the object and calculate the height and width of that boundary box. If the height of the box in ten continuous frames get decrease from its width then detect it as a fall and send the alert mail to the client. If the height is always greater than the width then the fall is not detected and moves to the next frame.

V. Performance Measure

TABLE I. YOLOv3 PERFORMANCE

Model	Model Size	Train	FPS (Frames per second)	Accuracy
YOLOv 3-320	247 MB	COCO trainval	45	78%

VI. Conclusion

In this paper, through an already trained Coco dataset model, based on a deep learning detection algorithm, YOLOv3, Fall detection, and vehicle crash detection is proposed. The YOLOv3 algorithm extracts features and completes the end-to-end prediction of the target frame and category at one time.

The experimental results show that although YOLOv3 has good characteristics and is efficiently faster still there are shortcomings compared to YOLOv4 and Yolov5, such as these algorithms outperform YOLOv3 in terms of accuracy.

In general, the YOLO series algorithms are better than the general computer vision method, which can meet the real-time and accurate detection requirements and the model is easier to deploy.

VII. Limitations

Our proposed system detects the fall of a person and vehicle crashes using the YOLO version 3 Algorithm which gives fast and accurate results but it also has some limitations, those are:

In-Vehicle crash detection, when two vehicles are passing very close to each other, there will be minor space between them at that time. The model may detect it as a crash has occurred between two vehicles as the bounding boxes around the vehicles will be collided.

In fall detection if a person intentionally tries to fall then also the model will detect it as a fall and send the alert mail to the client. Each frame of fall detection takes 2 seconds to check if the fall is detected or not, thus the model results in an accuracy of 78%. The reason for the decrease in accuracy is that: some of the similar postures were confusing for the model to predict the falling. It can be fixed with multiple cameras to see the postures of the people more clearly.

VIII. Acknowledgment

This work was supported by TPCT's Terna Engineering College, Nerul, Navi Mumbai, Maharashtra 400706.

IX. References

[1] Fall detection in dusky environment, Springer, Fall Detection in Dusky Environment | SpringerLink

[2] Computer Vision-based Accident Detection in Traffic Surveillance, 1911.10037.pdf (arxiv.org)

[3] Car Accident Detection and Notification System Using Smartphone, International Journal of Computer Science and Mobile Computing. https://www.researchgate.net/publication/291356742_Car_Accident_Detection_and_Notification_System_Using_Smartphone

[4] Automatic Vehicle Accident Detection and Reporting with Black Box, International Journal of Applied Engineering Research, ISSN 0973-4562 Vol. 10 No.87 (2015), https://www.researchgate.net/publication/290061841_Automatic_Vehicle_Accident_Detection_and_Reporting_With_Black_Box

[5] An eight-camera fall detection system using human fall pattern recognition via machine learning by a low-cost android box, https://www.nature.com/articles/s41598-021-81115-9

[6] Performance Analysis of Image Forgery Detection using Transform Function and Machine Learning Algorithms, https://turcomat.org/index.php/turkbilmat/article/view/12075