Available Online at www.ijcsmc.com/journal

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X
IMPACT FACTOR: 7.056

IJCSMC, Vol. 11, Issue. 5, May 2022, pg.75 – 82

Vehicle Crash Detection using YOLO Algorithm

[1]Prof. L. K. Wani; [2]Md Maaz Momin; [3]Sharwari Bhosale; [4]Abhishek Yadav; [5]Manas Nili
Department of Electronics and Telecommunication, Bharati Vidyapeeth's College of Engineering, Lavale, Pune

DOI: https://doi.org/10.47760/ijcsmc.2022.v11i05.007

Abstract— Car accidents cause a large number of deaths and disabilities every day, a certain proportion of which result from untimely treatment and secondary accidents. To some extent, automatic car accident detection can shorten response time of rescue agencies and vehicles around accidents to improve rescue efficiency and traffic safety level. —Road Accidents, a very common reason of tragic deaths and many times the victim dies due to non-reporting of such accidents to the proper authority. Since the accident was not reported the lack of emergency medical care results in death. We live in an era of technology where we are moving towards making the city, A Smart City. A smart city with smart AI based traffic monitoring and reporting mechanism can help providing medical emergencies in real time and this would result in saving lots of life. Traditional Traffic systems are equipped with IP cameras and sensors, and are already installed in most part of the city to monitor and control traffic. These systems are able to generate traffic tickets automatically. In this paper we are proposing a more advanced traffic monitoring system which can identify and detect moving objects such are cars, bikes etc in live camera feeds and detect collision of these moving objects and immediately send emergency alerts to the nearby authority for them to take necessary actions.
Keywords— YOLO, CNN, neural network.

I. INTRODUCTION

More than one million people died in various traffic accidents around the world, and many people suffered minor injuries. Many studies have shown that many developing and underdeveloped countries have the highest road traffic accident death rates, even though these countries only account for half of the world's vehicles. According to the data available in India, there is an average of 13 deaths per hour, that is, 140,000 deaths per year. The main goal is to enable the system to detect accidents based on the video sequence transmitted by the camera. A tool that helps accident victims who need it by detecting the accident early and notifying the authorities from then on. The goal is to detect accidents in seconds by using advanced deep-learning algorithms that use convolutional neural networks (CNN or ConvNet) to analyze the frames from the video generated by the camera. This paper is motivated with the idea of implementing statistical method of machine learning to detect any kind of collision in a live feed with the application of convolution neural network.

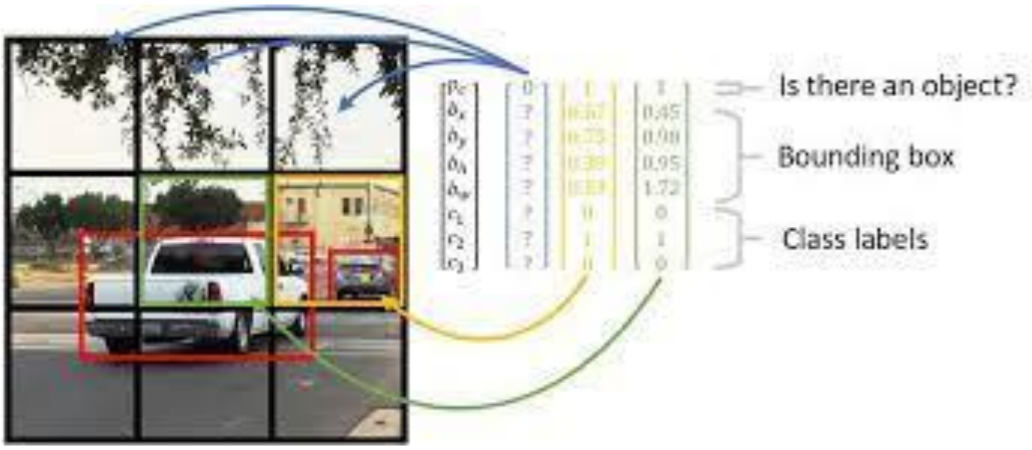

FIGURE 1: YOLO ALGORITHM

II. OBJECT IDENTIFICATION

Car Detection:

For car detection we used the YOLO net, it was proposed by, improved in and more recently, the last version was presented in. Despite other convolutional networks, it detects objects with a single pass creating a grid of SxS boxes, where each box has a logistic regression and a classification method, the regression method predict each box with five values: x, y, w, h and the confidence of the object being there. The classifier predicts C conditional class probabilities. At the final stage multiple bounding boxes appear around a single object, so non-maximum suppression is applied in order to keep the strongest detection around a single object. The architecture of YOLO for this research project is showed in the table I, obtained from Darknet library. The result in this step are the bounding boxes for each car, the YOLO performance is showed, here, this network has a very high accuracy and more important have a very low time processing against the rest.

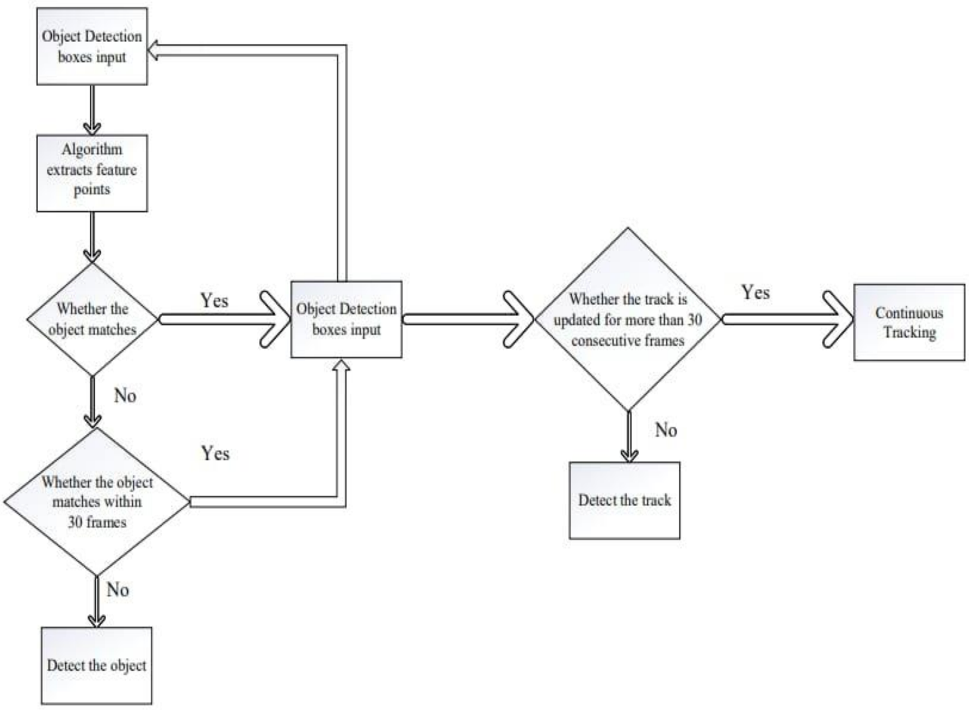

FIGURE 2: OBJECT DETECTION FLOWCHART

Car tracking:

We apply a car tracking algorithm each 30 frames, generating one tracker per car and saving the tracker position. Then we extract small videos, one per car. In the right of the Figure 5 we show an example, where three small videos have been extracted because three cars were detected in the first stage of the proposed model. The car tracker used in our model is based in a visual. Objects algorithm with correlation filters, proposed by Danelljan et al. The algorithm consists in the translation and scale estimation of the object through a series of correlations over the Fourier domain, where each correlation filter is submitted to an online learning process.. The procedure of correlation filters since the position of the visual object, according to Chen et al, can be summarized as follows. In each frame the region whose position was predicted in the previous frame is extracted for detection. Then the features HOG are extracted since the intensity level of each pixel of the image region, a cosine window is applied for smoothing a boundary effects. Next, a series of correlation operations are performed with element-wise multiplications using the Discrete Fourier Transform (DFT), it is computated with an efficient Fast Fourier Transform (FFT) algorithm, the answer is a spatial confidence map obtained with the inverse of FFT (IFFT). The position with the maximum value located on the map is predicted as the new position of the target. Then, the features of the estimated position region are extracted again for the training and updating of the correlation filter, being part of the online learning, the IFFT is applied again to obtain the response map and another prediction, this process continuous for each frame.

Car Crash Detection:

In order to detect a car crash scene, we are going to use the ViF descriptor because of the very low cost and acceptable accuracy. The ViF descriptor regards the statistics of magnitude changes of flow vectors over time. In order to get these vectors, used the optical flow algorithm proposed by named Iterative Reweighted Least Squares (IRLS), but in this context, we used the ViF descriptor with Horn-Schunck as optical flow algorithm proposed by.

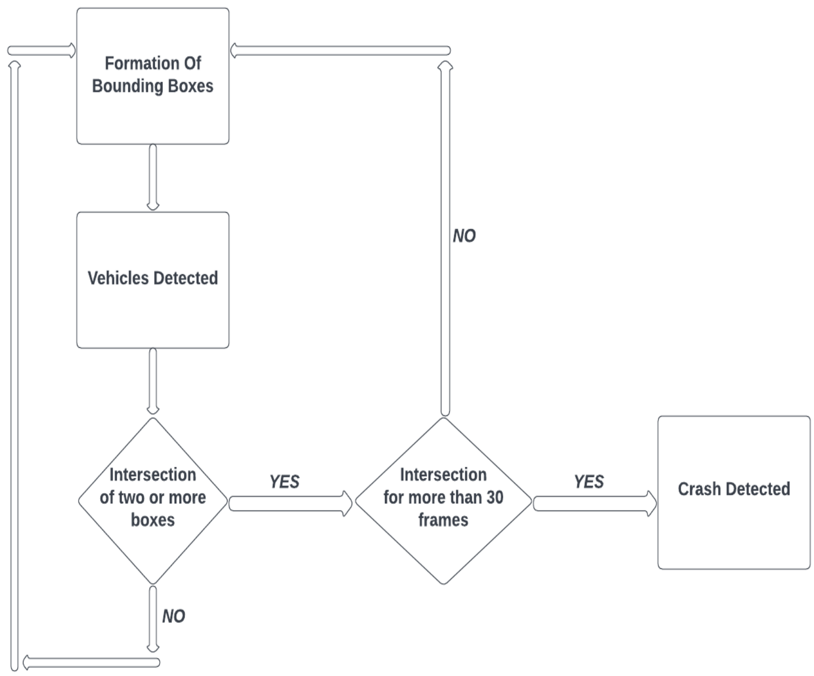

FIGURE 3: OBJECT CRASH DETECTION

III.METHOD BASED VIDEO FEATURES

With the development of machine vision and artificial neural network technology, more and more applications based on video processing have been applied in transportation and vehicle fields. Under this background, some researchers utilized video features of the car accident to detect it. Reference presented a Dynamic-Spatial-Attention Recurrent Neural Network (RNN) for anticipating accidents in dashcam videos, which can predict accidents about 2 seconds before they occur with 80% recall and 56.14% precision. Reference [26] proposed a car accident detection system based on first-person videos, which detected anomalies by predicting the future locations of car participants and then monitoring the prediction accuracy and consistency metrics. These methods also have some limitations because of low penetration of vehicular intelligent devices and shielding effects between vehicles. There are also some other methods which use roadside devices instead of vehicular equipments to obtain and process video. Reference proposed a novel accident detection system at intersection, which composed background images from image sequence and detected accidents by using Hidden Markov Model. Reference outlined a novel method for modelling of interaction among multiple moving objects, and used the Motion Interaction Field to detect and localize car accidents. Reference proposed a novel approach for automatic road accident detection, which was based on detecting damaged vehicles from footage received from surveillance cameras installed in roads. In this method, Histogram of gradients (HOG) and Gray level co-occurrence matrix features were used to train support vector machines. Reference presented a novel dataset for car accidents analysis based on traffic Closed-Circuit Television (CCTV) footage, and combined Faster Regions-Convolutional Neural Network (R-CNN) and Context Mining to detect and predict car accidents. The method in achieved 1.68 seconds in terms of Time-To-Accident measure with an Average Precision of 47.25%. Reference proposed a novel framework for automatic car accident detection, which learned feature representation from the spatio-temporal volumes of raw pixel intensity instead of traditional hand-crafted features. The experiments of method in demonstrated it can detect on average 77.5% accidents correctly with 22.5% false alarms.

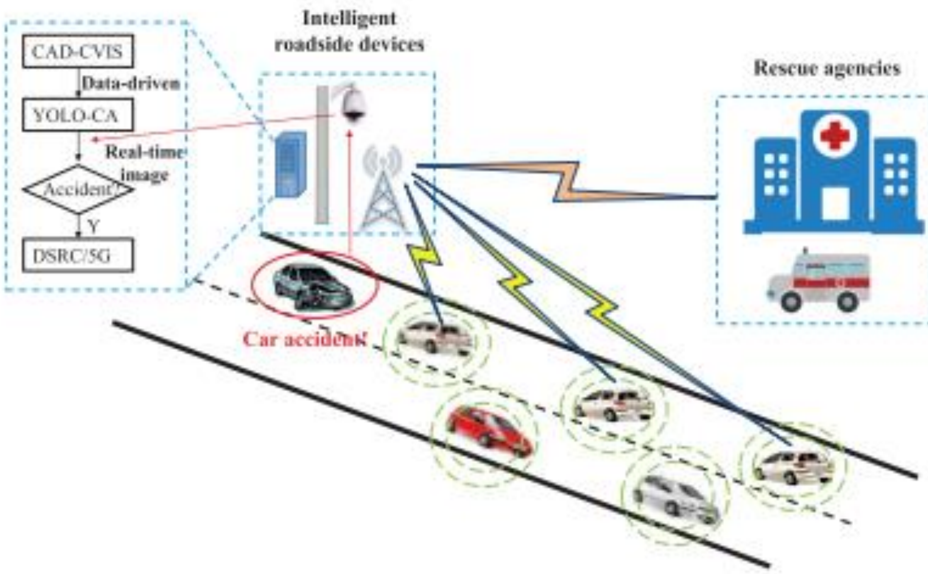

FIGURE 4: THE APPLICATION SCENARIO OF THE AUTOMATIC CAR ACCIDENT DETECTION METHOD BASED ON CVIS

Compared with the methods based on vehicle running condition, these methods improve the detection accuracy and some of them even can predict accidents about 2 seconds before they occur. To some extent, these methods are significant in decreasing the accident rate and improving traffic safety. However, the detection accuracy of these methods is low and the error rate is high, and the wrong accident information will have a great impact on the normal traffic flow. Concerning the core issue mentioned above, in order to avoid the drawbacks of vehicular cameras, our proposed method utilizes the roadside intelligent edge devices to obtain traffic video and process image. Moreover, for sake of improving the accuracy of accident detection method based on intelligent roadside devices, we establish the CAD-CVIS dataset based on video sharing websites, which is consisted of various kinds of accident types, weather conditions and accident locations. Moreover, we develop the model YOLO-CA to improve the reliability and real-time performance among different traffic conditions by combining deep learning algorithms and MSFF method.

IV. DETECTION PRINCIPAL

The detection principle of YOLO-CA, which includes extracting feature map and predicting bounding box. YOLO-CA divides the input image into 13×13 grid and 26×26 grid. The first grid is responsible for detecting the large objects, whereas the second grid makes up for the inaccuracy of small target detection in the first grid. The feature extraction networks corresponding to these two grids are different, but the detection models of the objects is similar. For ease of presentation, we regard the first grid as example to explain the training steps of YOLO-CA. The center of car accident region falls into the grid cell (7, 5), so this cell is responsible for detecting this car accident in the whole training process.

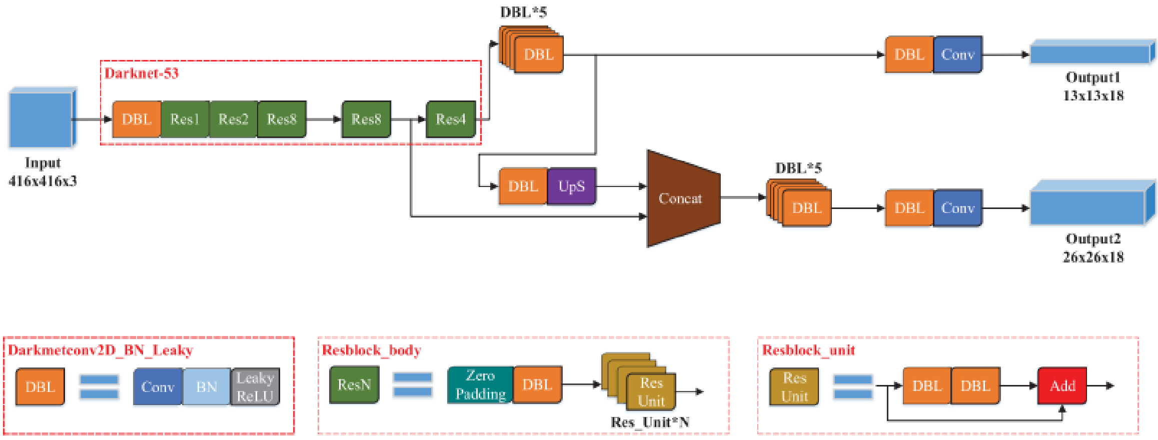

Figure 5: The network structure of YOLO-CA

Then the cell (7, 5) will predict three bounding boxes, and each boxes includes six parameters: x, y,w, h,CS, p. The (x, y) is the center point of the bounding box, and the (w, h) is the ratio of width and height of the bounding box to the whole image. The CS is confidence score of bounding box, which represents how confident the model is that the bounding box contains an object and how accurate it thinks the box is that it predicts. Lastly, each bounding box will predict class probability of car accident p. After the training of a batch of images, the loss of model will be calculated, which is utilized to adjust the weights of parameters. In the calculation of loss, let the ground truth of an object is $x*$, $y*$, $w*$, $h*$, $CS*$, $p*$. $S \times S$ is the number of cells in grid, and B is the number of predicted bounding boxes of each grid cell. For each grid cell, the Pr(Objects) equals 1 when the cell contains center of object, whereas it equals 0 when there is not center of object in the cell.

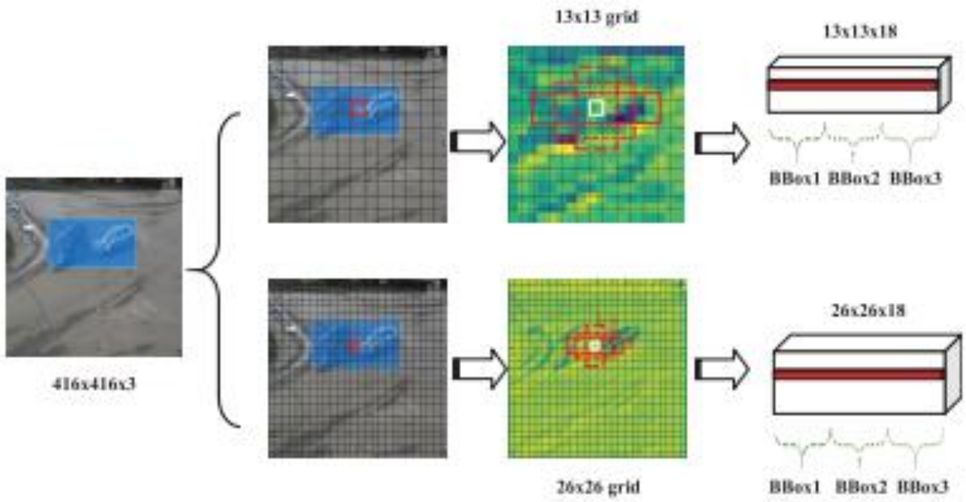

FIGURE 6: THE DETECTION PRINCIPAL OF YOLO-CA

V. RESULT ANALYSIS

1) Training results of YOLO-CA

The training results of YOLO-CA, including the changes of precision, recall, IoU and loss of each batch in iteration process. In the training process of YOLO-CA, we regard the prediction result with IoU over 0.5 and right classification as true result, and other predictions are all false results. As shown in Table. 3, the prediction results can be divided into four parts:

(1)TP: Truth Positive.

(2) FP: False Positive.

(3) FN: False Negative.

(4) TN: True Negative.

The precision is defined as precision=TPTP+FP and recall is defined as recall=TPTP+FN.

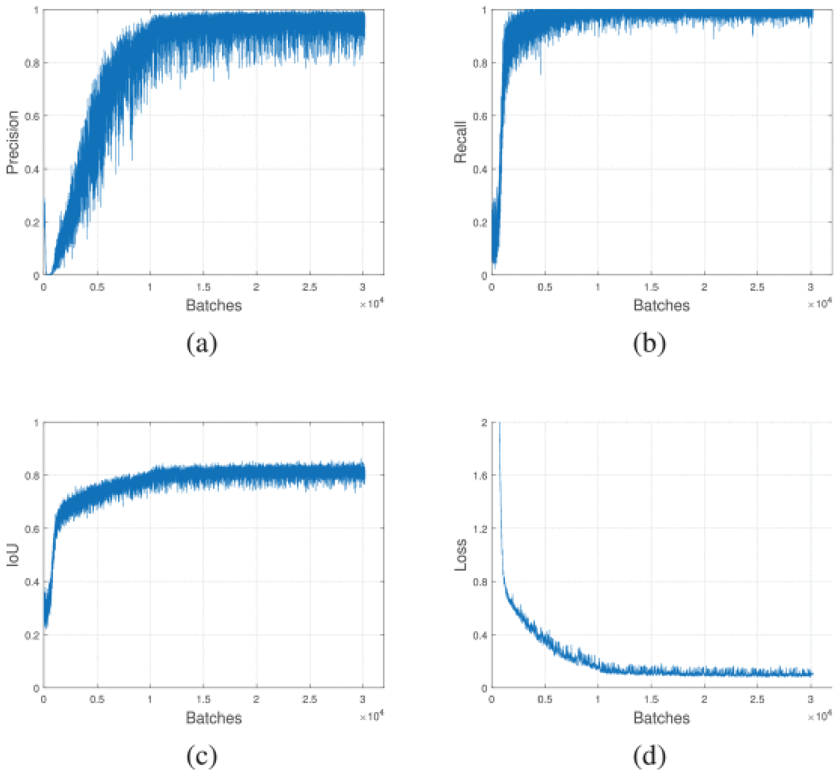

Figure 7: The training results of YOLO-CA. (a) Precision (b) Recall (c) IoU (d) Loss

		Ground truth	
		Positive	Negative
Prediction Result	Positive	TP	FP
	Negative	FN	TN

FIGURE 8: DISTRIBUTION MAP OF PREDICTION RESULTS

In the training process of YOLO-CA, we regard the prediction result with IoU over 0.5 and right classification as true result, and other predictions are all false results. As shown in Table. 3, the prediction results can be divided into four parts: (1) TP: Truth Positive. (2) FP: False Positive. (3) FN: False Negative. (4) TN: True Negative. The precision is defined as precision = TP TP+FP and recall is defined as recall = TP TP+FN. As shown in Fig. 7a, with the increasing of iterations, the precision of YOLO-CA is increasing gradually and converge over 90%. Moreover, recall eventually converges to more than 95%. In terms of locating performance of YOLO-CA in training set, IoU finally stabilizes above 0.8. The decreasing process of loss of YOLO-CA in (7), and the final convergence of loss is less than 0.2.

2) Comparative experiments and visual result:

The comparative experiments are conducted for comparing seven detection models: (1) One-stage models: SSD, our proposed YOLO-CA, traditional YOLO-v3 and YOLO-v3 without MSFF (Multi-Scale Feature Fusion). (2) Two-stage models: Fast R-CNN, Faster R-CNN and Faster R-CNN with FPN. In order to comparatively demonstrate the validation of YOLO-CA as well as confirm its strength in terms of the comprehensive performance on the accuracy and real-time, the following indexes are selected for comparison among the seven models:

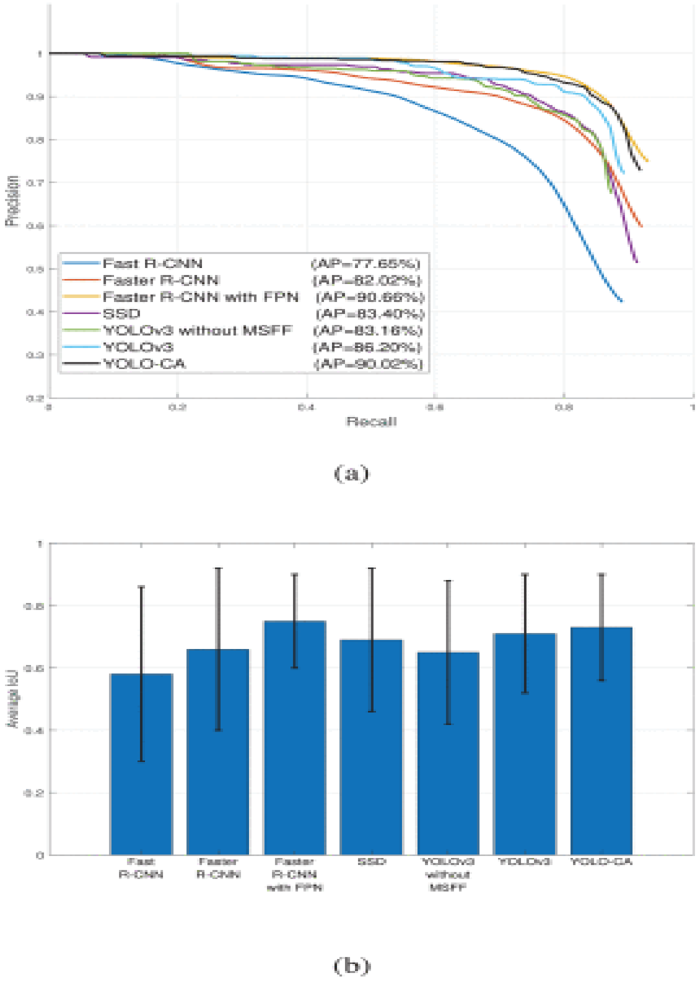

FIGURE 9: THE AP AND IOU RESULTS OF DIFFERENT MODELS. (A) PRECISION-RECALL CURVE (B) AVERAGE IOU

• Average Precision (AP) that is defined as the average value of precision under different recall, which can be changed by adjusting threshold of classification confidence. AP index evaluate the accuracy performance of detection models. The average precision can be calculated

$$AP = \sum precision(r)$$

where r is recall.

• Average Intersect over Union (Average IoU) that is utilized to evaluate the object locating performance of detection models. The Average IoU is the average value of IoUs between every prediction bounding box and corresponding ground truth.

• Frames Per Second (FPS). Inference time is defined as the average time cost of detecting a frame among test set. FPS is the reciprocal of inference time, which is defined as the average number of frames that can be detected in one second.

VI. CONCLUSIONS

In this study, the proposed accident detection system can be trained by using regression based algorithm called YOLO (you only look once) algorithm on the sample vehicle datasets and the vehicle detection process has been successfully performed by the trained model vehicle detector being tested on the test data set with the live video feeds from the webcam. The proposed system is faster than other object detection methods and predicts the object better other object detection algorithm such as Faster-CNN or Fast CNN. The input can also be optimized and give better results. Further the system alerts via a wireless communication devices to nearby emergency vehicles.

REFERENCES

[1]. WHO. Global Status Report on Road Safety 2018. Accessed: Dec. 2018. [Online]. Available: https://www.who.int/violence_injury_prevention/road_safety_status/2018/en/

[2]. H. L. Wang and M. A. Jia-Liang, ‘‘A design of smart car accident rescue system combined with WeChat platform,’’ J. Transp. Eng., vol. 17, no. 2, pp. 48–52, Apr. 2017.

[3]. Y. K. Ki and D. Y. Lee, ‘‘A traffic accident recording and reporting model at intersections,’’ IEEE Trans. Intell. Transp. Syst., vol. 8, no. 2, pp. 188–194, Jun. 2007.

[4]. W. Hao and J. Daniel, ‘‘Motor vehicle driver injury severity study under various traffic control at highway-rail grade crossings in the United States,’’ J. Saf. Res., vol. 51, pp. 41–48, Dec. 2014.

[5]. J. White, C. Thompson, H. Turner, H. Turner, and D. C. Schmidt, ‘‘Wreckwatch: Automatic traffic accident detection and notification with smartphones,’’ Mobile Netw. Appl., vol. 16, no. 3, pp. 285–303, Jun. 2011.

[6]. S. Sadek, A. Al-Hamadi, B. Michaelis, and U. Sayed, ‘‘Real-time automatic traffic accident recognition using HFG,’’ in Proc. 20th Int. Conf. Pattern Recognit., Aug. 2010, pp. 3348–3351.

[7]. A. Shaik, N. Bowen, J. Bole, G. Kunzi, D. Bruce, A. Abdelgawad, and K. Yelamarthi, ‘‘Smart car: An IoT based accident detection system,’’ in Proc. IEEE Global Conf. Internet Things (GCIoT), Dec. 2018, pp. 1–5.

[8]. D. Singh and C. K. Mohan, ‘‘Deep spatio-temporal representation for detection of road accidents using stacked autoencoder,’’ IEEE Trans. Intell. Transp. Syst., vol. 20, no. 3, pp. 879–887, Mar. 2019.

[9]. M. Zheng, T. Li, R. Zhu, J. Chen, Z. Ma, M. Tang, Z. Cui, and Z. Wang, ‘‘Traffic accident’s severity prediction: A deep-learning approach-based CNN network,’’ IEEE Access, vol. 7, pp. 39897–39910, 2019.

[10]. L. Zheng, Z. Peng, J. Yan, and W. Han, ‘‘An online learning and unsupervised traffic anomaly detection system,’’ Adv. Sci. Lett., vol. 7, no. 1, pp. 449–455, 2012.

[11]. Y. Fangchun, W. Shangguang, L. Jinglin, L. Zhihan, and S. Qibo, ‘‘An overview of Internet of vehicles,’’ China Commun., vol. 11, no. 10, pp. 1–15, Oct. 2014.

[12]. C. Ma, W. Hao, A. Wang, and H. Zhao, ‘‘Developing a coordinated signal control system for urban ring road under the vehicle-infrastructure connected environment,’’ IEEE Access, vol. 6, pp. 52471–52478, 2018.

[13]. S. Zhang, J. Chen, F. Lyu, N. Cheng, W. Shi, and X. Shen, ‘‘Vehicular communication networks in the automated driving era,’’ IEEE Commun. Mag., vol. 56, no. 9, pp. 26–32, Sep. 2018.

[14]. Y. Wang, D. Zhang, Y. Liu, B. Dai, and L. H. Lee, ‘‘Enhancing transportation systems via deep learning: A survey,’’ Transp. Res. C, Emerg. Technol., 2018.

[15]. G. Wu, F. Chen, X. Pan, M. Xu, and X. Zhu, ‘‘Using the visual intervention influence of pavement markings for rutting mitigation—Part I: Preliminary experiments and field tests,’’ Int. J. Pavement Eng., vol. 20, no. 6, pp. 734–746, 2019.

[16]. S. Ramos, S. Gehrig, P. Pinggera, U. Franke, and C. Rother, ‘‘Detecting unexpected obstacles for self-driving cars: Fusing deep learning and geometric modeling,’’ in Proc. IEEE Intell. Vehicles Symp. (IV), Jun. 2017, pp. 1025–1032.

[17]. T. Qu, Q. Zhang, and S. Sun, ‘‘Vehicle detection from high-resolution aerial images using spatial pyramid pooling-based deep convolutional neural networks,’’ Multimedia Tools Appl., vol. 76, no. 20, pp. 21651–21663, 2017.

