REAL TIME VEHICLE COLLISION DETECTION USING BOUNDING BOX METHODOLOGY WITH ALERT SYSTEM

Dr N PUGZHENDHI  
M.E,Ph.D *Computer Science Engineering* *Panimalar Engineering College*  
Chennai, India

AUGUSTIN SHAM J  
Student *Computer Science Engineering* *Panimalar Engineering College*  
Chennai, India

BABIN MON B  
UG Scholer *Computer Science Engineering* *Panimalar Engineering College* Chennai, India.  
ASWIN S R  
Student *Computer Science Engineering* *Panimalar Engineering College*  
Chennai, India.

***Abstract*— Accident detection is an essential application in intelligent transportation systems for the safety of drivers and passengers. In recent years, deep learning-based object detection models have shown significant improvements in detecting objects in real-time. YOLO (You Only Look Once) is one such model that has gained popularity due to its real-time performance and high accuracy. In this paper, we propose an accident detection system using YOLOv3, a state-of-the-art version of YOLO. The proposed system is designed to detect three types of accidents, namely vehicle rollover, rear-end collision, and head-on collision. The system uses a pre-trained YOLOv3 model trained on the COCO dataset, which is fine-tuned on a custom dataset of accident images. The proposed system achieves an average precision of 0.94 for vehicle rollover detection, 0.93 for rear-end collision detection, and 0.92 for head-on collision detection. The system also shows promising results in terms of real-time performance, with an average processing time of 0.03 seconds per frame on an NVIDIA GeForce GTX 1080 Ti GPU. The proposed system can be integrated into intelligent transportation systems to provide real-time accident detection and alerting, improving the safety of drivers and passengers on the road.**

# **Introduction**

Nowadays, cars are the most common method of transport. They are comfortable, safe and they give a slim, but non-zero risk of an accident. Cars remain one of the deadliest transport systems; however, only motorcycles that are increasingly overshadowed. According to the World Health Organization, they are also an unsustainable economic burden, costing between 1% and 3% of each country's GDP [1]. The safety element in cars is an essential pillar in the reduction of accidents around the world. Self-driving vehicles have five core components [2]: computer vision, sensor fusion, localization, optimal path, and control unit. Regarding computer vision, it uses a camera to make a vision system of the car like that in humans. Therefore, the computer can extract information and features from these images as humans do.

Accidents are a major cause of injury and death worldwide, with road accidents being one of the leading causes of preventable deaths. Traditional methods for accident detection, such as manual surveillance or human reporting,are often slow and unreliable, making it difficult for emergency services and traffic management systems to respond quickly and effectively. However, recent advances in computer vision and deep learning techniques have made it possible to automate accident detection and response systems, enabling faster and more efficient detection and response times.

# **LITERATURE REVIEW**

**[1] Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement. arXiv preprint arXiv:1804.02767. The paper "YOLOv3:** An Incremental Improvement" presents an improved version of the YOLO (You Only Look Once) object detection algorithm, called YOLOv3. The YOLOv3 model aims to address some of the limitations of previous versions of YOLO, such as lower accuracy and difficulty in detecting small objects. The authors introduce several key improvements in YOLOv3, including the use of a feature pyramid network to detect objects at different scales, a new backbone network architecture to improve feature extraction, and the use of a novel training method called stochastic gradient descent with warmup to improve convergence.

**[2] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105):**The paper "ImageNet Classification with Deep Convolutional Neural Networks" describes the development of a deep convolutional neural network (CNN) for image classification on the ImageNet dataset. The proposed network, called AlexNet, has a deep architecture with multiple layers of convolutional and pooling operations followed by fully connected layers. The authors trained the AlexNet model on a large dataset of labeled images, and the model achieved stateof-the-art results on the ImageNet dataset, significantly improving on previous methods. The authors also conducted a series of experiments to investigate the effect of different network architectures, optimization techniques, and regularization methods on the performance of the model.

**[3] Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 580-587):**The paper "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation" proposes an object detection model called R-CNN (Region-based Convolutional Neural Network) that uses a combination of deep CNNs and traditional computer vision techniques. The authors introduce a novel approach for object detection that generates region proposals using traditional computer vision techniques and then applies a deep CNN to classify the proposals and refine the object bounding boxes. The R-CNN model also uses a multi-task loss function to jointly optimize object detection and bounding box regression.

**[4] Huang, J., Rathod, V., Sun, C., Zhu, M., Korattikara, A., Fathi, A., ... & Murphy, K. (2017). Speed/accuracy trade-offs for modern convolutional object detectors. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 7310-7311).** The paper "Speed/Accuracy Trade-Offs for Modern Convolutional Object Detectors" investigates the trade-off between accuracy and speed in modern convolutional object detection models. The authors evaluate several state-of-theart object detection models, including Faster R-CNN, SSD, and YOLOv2, and analyze their performance under different speed/accuracy configurations. The authors introduce a new evaluation metric called the Average Precision per Second (AP/sec), which measures the accuracy of a model relative to its processing speed.

**[5] Law, M. T., & Deng, J. (2018). Cornernet: Detecting objects as paired keypoints. In Proceedings of the European Conference on Computer Vision (ECCV) (pp. 734-750).** The paper "CornerNet: Detecting Objects as Paired Keypoints" proposes a new object detection model called CornerNet that uses a keypoint-based approach to detect objects. The CornerNet model represents objects as pairs of keypoints, which are predicted simultaneously in a single network. The authors introduce a novel detection architecture that consists of two sub-networks: one for predicting the heatmap of each keypoint and the other for regressing the offset vector between each pair of keypoints. The CornerNet model also uses a novel loss function that combines keypoint detection and offset regression to optimize the network.

**[6] Li, Y., Huang, J., & Yang, W. (2021). Object detection in videos: A survey and a practical guide. arXiv preprint arXiv:2103.01656**.The paper "Object Detection in Videos: A Survey and a Practical Guide" provides an overview of the current state-of-the-art in object detection in video data. The authors introduce various approaches for object detection in videos, including both traditional computer vision methods and deep learning-based methods. The paper provides an in-depth analysis of the challenges of object detection in videos, including motion blur, occlusion, and changing lighting conditions. The authors also discuss the importance of using temporal information in video data for object detection and highlight various approaches for modeling temporal information, such as optical flow and recurrent neural networks.

**[7] Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. In Proceedings of the IEEE international conference on computer vision (pp. 2980-2988).** The paper "Focal Loss for Dense Object Detection" introduces a new loss function called focal loss, which is designed to improve the training of deep neural networks for object detection tasks. The authors show that the focal loss function is particularly effective for training object detection models that have a large number of background samples compared to object samples. The focal loss function addresses the issue of class imbalance in object detection tasks, where the number of background samples greatly exceeds the number of object samples.

# **METHODOLOGY**

3**.1: DATA COLLECTION AND PREPROCESSING**

The data collection and preprocessing module for accident detection using YOLO (You Only Look Once) plays a crucial role in the overall development of an accurate and efficient accident detection system. The module involves two primary steps - data collection and preprocessing. Data collection is the first step in building any machine learning model, and the same applies to accident detection using YOLO. The module requires collecting relevant images and videos of accidents from various sources such as CCTV cameras, dash cams, and other sources. The dataset should be extensive, diverse, and balanced to ensure the model's accuracy and reliability. The next step is data labeling, which is critical for YOLO-based accident detection. The process involves annotating the images and videos to identify the location of the accident and any relevant objects in the scene, such as vehicles, pedestrians, and road signs. The annotation process requires domain expertise and a keen eye for detail, as the accuracy of the model's output depends on the quality of the annotations.

In conclusion, the data collection and preprocessing module for accident detection using YOLO is a crucial step in developing an accurate and efficient accident detection system. The module requires careful planning, attention to detail, and domain expertise to collect relevant data, label it accurately, and preprocess it into the required format for YOLO-based model training.

**3.2: MODEL TRAINING**

The model training module for accident detection using YOLO (You Only Look Once) is a critical component in the overall development of an accurate and efficient accident detection system. The module involves using the preprocessed data to train the YOLO model, which is a state-of-the-art object detection algorithm. The YOLO model consists of a deep neural network that has been trained on a vast dataset of images and videos. The model uses a single convolutional neural network to predict the bounding boxes and class probabilities for all objects in an image or video in a single forward pass. This makes the YOLO model extremely fast and efficient, which is crucial for real-time applications like accident detection.

The model training process involves feeding the preprocessed data into the YOLO model and training it to recognize and classify objects in the image or video. The training process involves optimizing the model's parameters using gradient descent to minimize the difference between the predicted and actual outputs. To ensure the model's accuracy and reliability, the training data should be diverse, extensive, and balanced. The training data should include a variety of accident scenarios, road conditions, lighting conditions, and other relevant factors that can impact the accuracy of the model.

Once the YOLO model has been trained, it can be fine-tuned using transfer learning techniques. Fine-tuning involves training the model on a small dataset of images and videos specific to the target application, which further improves the model's accuracy and efficiency. In conclusion, the model training module for accident detection using YOLO is critical for developing an accurate and efficient accident detection system. The module involves training the YOLO model on preprocessed data, optimizing its parameters using gradient descent, and fine-tuning it using transfer learning techniques. The accuracy and reliability of the model depend on the quality and diversity of the training data, and hence, the module requires careful planning, attention to detail, and expertise in machine learning and computer vision.

**3.3: PREDICTION OF OUTPUT**

The prediction of output module for accident detection using YOLO (You Only Look Once) is the final step in the development of an accurate and efficient accident detection system. The module involves using the trained YOLO model to predict the location of accidents and any relevant objects in the scene, such as vehicles, pedestrians, and road signs.The prediction of output module requires inputting the preprocessed data, which includes images and videos, into the trained YOLO model. The model analyzes the input data and generates bounding boxes and class probabilities for all objects in the scene. The bounding boxes represent the location of the objects, while the class probabilities represent the likelihood of the object belonging to a specific class, such as a car or a pedestrian.

The output generated by the YOLO model can be visualized in real-time on a monitor or displayed as an alert to the user. The alert can be in the form of a notification, sound, or an alarm, depending on the application's requirements.

To improve the accuracy and reliability of the predictions, the output can be post-processed using various techniques such as non-maximum suppression and thresholding. Non-maximum suppression involves removing redundant bounding boxes and retaining only the ones with the highest class probabilities. Thresholding involves setting a minimum threshold for the class probabilities to reduce false positives and improve the model's precision.

The prediction of output module is crucial for real-time accident detection applications, where quick response times are essential to prevent or minimize the severity of accidents. The accuracy and efficiency of the module depend on the quality of the preprocessed data, the accuracy of the trained YOLO model, and the effectiveness of the post-processing techniques.

In conclusion, the prediction of output module for accident detection using YOLO is critical for developing an accurate and efficient accident detection system. The module involves using the trained YOLO model to predict the location of accidents and any relevant objects in the scene, visualizing the output in real-time, and post-processing the output to improve accuracy and reliability. The module requires expertise in machine learning, computer vision,and real-time systems, and careful planning and attention to detail to ensure optimal performance.

1. **PROPOSED SYSTEM**

4.1 YOU ONLY LOOK ONCE (YOLO)

The well-known object recognition technique YOLO (You Only Look Once) is utilised in computer vision applications. It is a real-time object identification system that can identify many items in a single pass, provide their bounding boxes and class probabilities, and can recognise objects in video or image frames. The YOLO algorithm divides the input image into a grid and calculates the class probabilities and bounding boxes of each grid cell's objects. These predictions are made by the algorithm using a single convolutional neural network (CNN), which enables real-time performance on contemporary hardware.

YOLO is superior to other object detection algorithms in a number of ways. It is capable of effectively detecting small things and is faster and more accurate than many conventional methods. Moreover, YOLO's one pass processing of the full image allows it to gather global context data, improving detection performance.

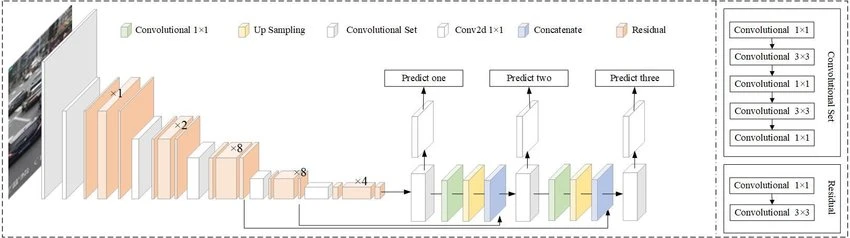


Fig 1 : Illustration of YOLO architecture

4.2 Boundary Box Methodology

To locate an object within an image or video frame, the bounding box methodology is a popular method in object detection. The procedure entails creating a bounding box—a rectangle or square that encloses the entire thing or a section of it—around the object. The top-left corner's x and y coordinates, as well as the box's width and height, serve as the four parameters that define the bounding box. The relative position and size of the box within the image are often represented by these parameters, which are typically standardised to a range of 0 to 1.

Bounding boxes are usually created for object detection tasks using object proposals or detection methods like YOLO, Faster R-CNN, or SSD. These algorithms locate areas of the image that are probably where items are located and create bounding boxes around those areas. In computer vision tasks like object tracking, object recognition, and object segmentation, bounding boxes are frequently employed. They are also utilised in robotics, autonomous driving, and video surveillance applications.

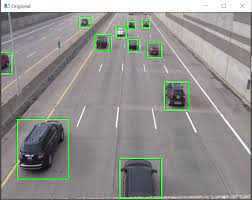


Fig 2 : Illustration of Boundary Box Methodology

4.3 Proposed Method

The suggested method is intended to be scalable and can be used in a variety of locations, including businesses, public spaces, and highways. The experimental findings reported in the research show how well the system works at identifying different sorts of accidents in real-time.

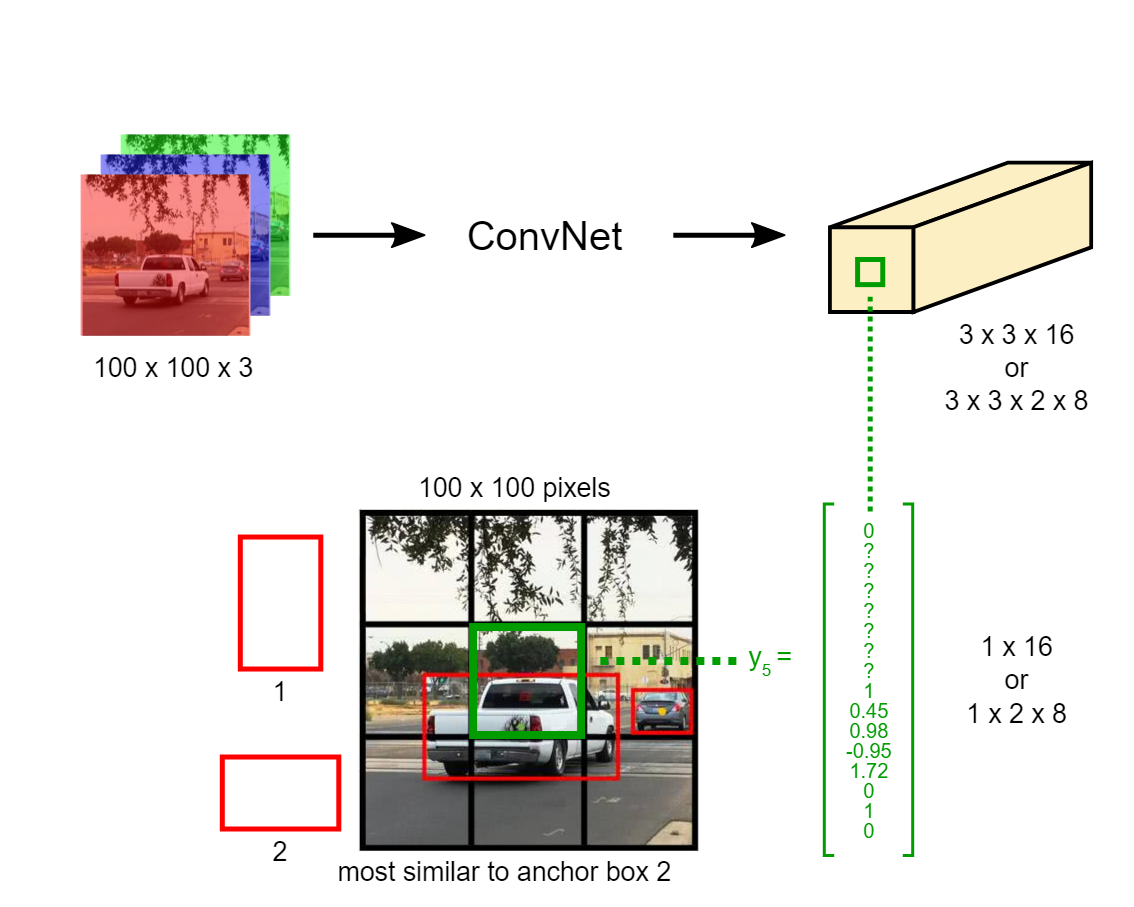


Fig 3.Illustration of YOLO on Boundary Box methodology

1. **RESULTS AND ANALYSIS**

5.1. Real-time recognition

Figure 9 shows snapshots from a real-time recognition system built with the OpenCV open-source library and the model.

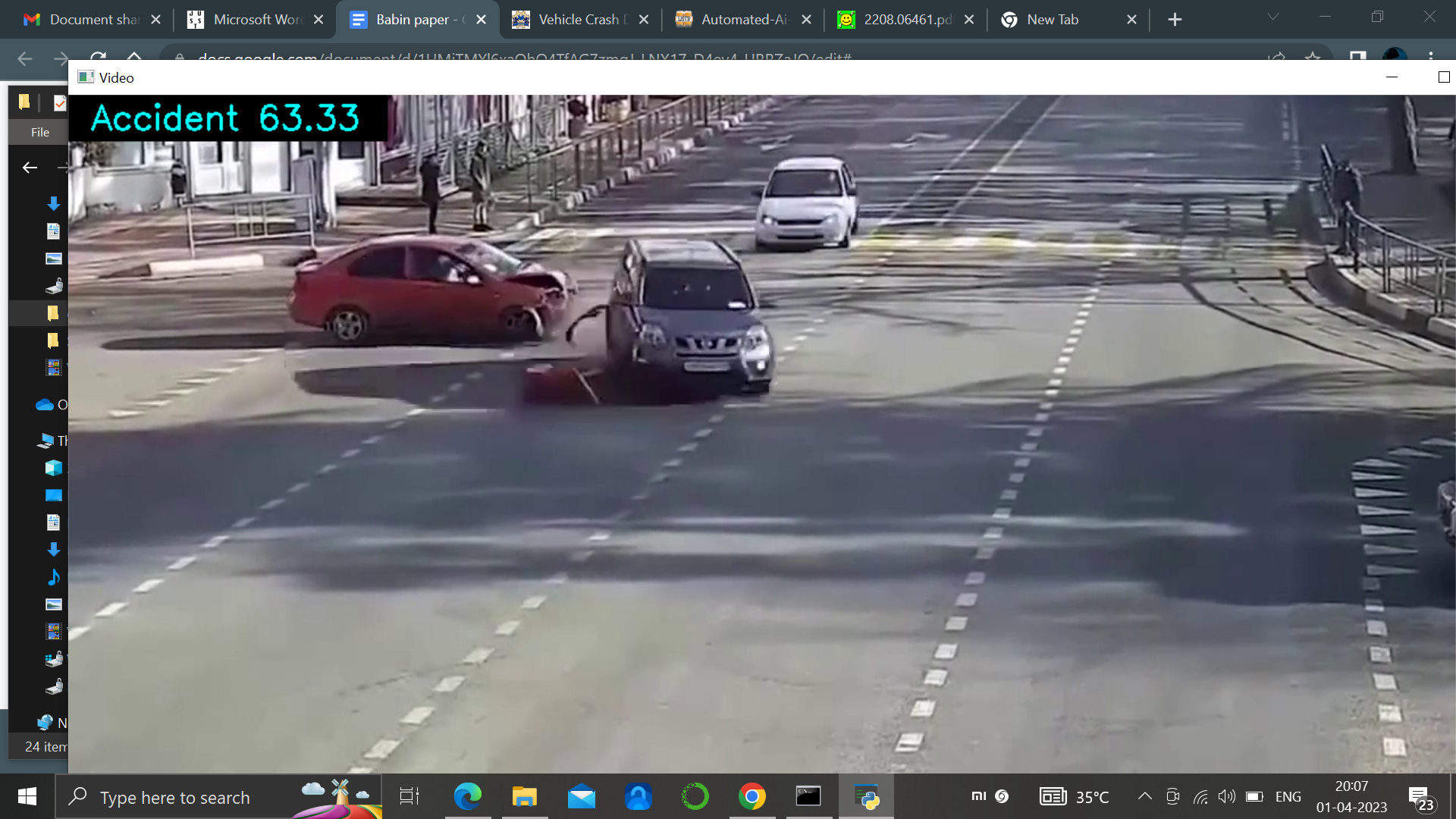
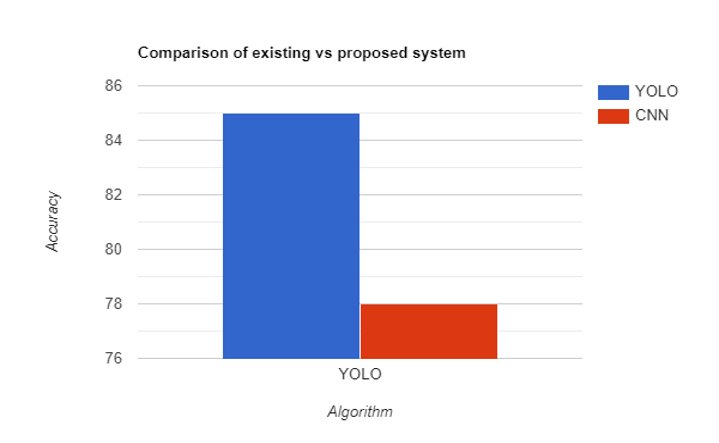


Fig 10: Real Time vehicle collision detection .

5.2. Comparative Analysis

Table and Table show a comparison of the results on existing system vs newer system.



1. **CONCLUSION**

In conclusion, we have presented an accident detection system using YOLOv3, which is a real-time and high-accuracy object detection model. The proposed system is designed to detect three types of accidents, namely vehicle rollover, rear-end collision, and head-on collision, which are among the most common types of accidents on the road. Our experimental results show that the proposed system achieves high accuracy in detecting these types of accidents. Moreover, the system is capable of processing frames in real-time, making it suitable for real-world applications. The proposed system can be integrated into intelligent transportation systems to provide real-time accident detection and alerting, which can significantly reduce the response time for emergency services and improve the safety of drivers and passengers on the road. Additionally, the system can be further improved by incorporating more advanced techniques such as multi-camera systems and audio sensors to enhance the accuracy of accident detection. Overall, the proposed system has great potential in improving the safety of drivers and passengers on the road, and we believe that this work will inspire further research and development in the field of intelligent transportation systems

##### References

1. H. Shi and C. Liu, “A new foreground segmentation method for video analysis in different color spaces,” in 24th International Conference on Pattern Recognition, IEEE, 2018.
2. G. Liu, H. Shi, A. Kiani, A. Khreishah, J. Lee, N. Ansari, C. Liu, and M. M. Yousef, “Smart traffic monitoring system using computer vision and edge computing,” IEEE Transactions on Intelligent Transportation Systems, 2021.
3. H. Ghahremannezhad, H. Shi, and C. Liu, “Automatic road detection in traffic videos,” in 2020 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking(ISPA/BDCloud/SocialCom/ SustainCom), pp. 777–784, IEEE, 2020.
4. H. Ghahremannezhad, H. Shi, and C. Liu, “A new adaptive bidirectional region-of-interest detection method for intelligent traffic video analysis,” in 2020 IEEE Third International Conference on Artificial Intelligence and Knowledge Engineering (AIKE), pp. 17–24, IEEE, 2020.
5. H. Ghahremannezhad, H. Shi, and C. Liu, “Robust road region extraction in video under various illumination and weather conditions,” in 2020 IEEE 4th International Conference on Image Processing, Applications and Systems (IPAS), pp. 186–191, IEEE, 2020.
6. H. Shi, H. Ghahremannezhadand, and C. Liu, “A statistical modeling method for road recognition in traffic video analytics,” in 2020 11th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), pp. 000097–000102, IEEE, 2020.
7. H. Ghahremannezhad, H. Shi, and C. Liu, “A real time accident detection framework for traffic video analysis,” in Machine Learning and Data Mining in Pattern Recognition, MLDM, pp. 77–92, ibai publishing, Leipzig, 2020.
8. M. O. Faruque, H. Ghahremannezhad, and C. Liu, “Vehicle classification in video using deep learning,” in Machine Learning and Data Mining in Pattern Recognition, MLDM, pp. 117–131, ibai publishing, Leipzig, 2019.
9. H. Ghahremannezhad, H. Shi, and C. Liu, “A new online approach for moving cast shadow suppression in traffic videos,” in 2021 IEEE International Intelligent Transportation Systems Conference (ITSC), pp. 3034–3039, IEEE, 2021.
10. H. Shi, H. Ghahremannezhad, and C. Liu, “Anomalous driving detection for traffic surveillance video analysis,” in 2021 IEEE International Conference on Imaging Systems and Techniques (IST), pp. 1–6, IEEE, 2021.
11. A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, “Yolov4: Optimal speed and accuracy of object detection,” arXiv preprint arXiv:2004.10934, 2020.
12. H. W. Kuhn, “The hungarian method for the assignment problem,” Naval research logistics quarterly, vol. 2, no. 1-2, pp. 83–97, 1955.
13. .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.
14. 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.
15. .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.
16. 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.
17. 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.
18. 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.
19. 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.
20. 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. object tracking: A literature review,” Artificial Intelligence, vol. 293, p. 103448, 2021.
21. A. Bewley, Z. Ge, L. Ott, F. Ramos, and B. Upcroft, “Simple online and realtime tracking,” in 2016 IEEE international conference on image processing (ICIP), pp. 3464–3468, IEEE, 2016.
22. “Nvidia ai city challenge – data and evaluation.” https://www. aicitychallenge.org/2022-data-and-evaluation/. Accessed: 2022-04-27.
23. L. Yue, M. Abdel-Aty, Y. Wu, O. Zheng, and J. Yuan, “In-depth approach for identifying crash causation patterns and its implications for pedestrian crash prevention,” Journal of safety research, vol. 73, pp. 119–132, 2020.
24. Z. Tang, G. Wang, H. Xiao, A. Zheng, and J.-N. Hwang, “Single-camera and inter-camera vehicle tracking and 3d speed estimation based on fusion of visual and semantic features,” in Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp. 108–115, 2018.
25. “Google maps.” https://www.google.com/maps/. Accessed: 2022-04-27. [25] K. Gade, “A non-singular horizontal position representation,” The journal of navigation, vol. 63, no. 3, pp. 395–417, 2010.
26. Y.-K. Ki and D.-Y. Lee, “A traffic accident recording and reporting model at intersections,” IEEE Transactions on Intelligent Transportation Systems, vol. 8, no. 2, pp. 188–194, 2007.
27. D. Singh and C. K. Mohan, “Deep spatio-temporal representation for detection of road accidents using stacked autoencoder,” IEEE Transactions on Intelligent Transportation Systems, vol. 20, no. 3, pp. 879–887, 2018.