Journal of Korea Society of Industrial Information Systems Vol.29 No.1, Feb. 2024 :91-105
http://dx.doi.org/10.9723/jksiis.2024.29.1.091 ISSN:1229-3741

YOLO v8을 활용한 컴퓨터 비전 기반 교통사고 탐지+

(Computer Vision-Based Car Accident Detection using YOLOv8)

마르와 차차 안드레아[1)], 이 충 권[2)*], 김 양 석[3)], 노 미 진[3)], 문 상 일[4)], 신 재 호[4)]
(Marwa Chacha Andrea, Choong Kwon Lee, Yang Sok Kim, Mi Jin Noh, Sang Il Moon, and Jae Ho Shin)

요 약 자동차 사고는 차량 간의 충돌로 인해 발생되며, 이로 인해 차량의 손상과 함께 인적, 물적 피해가 유발된다. 본 연구는 CCTV에 의해 촬영되어 YouTube에 업로드된 차량사고 동영상으로부터 추출된 2,550개의 이미지 프레임을 기반으로 차량사고 탐지모델을 개발하였다. 전처리를 위해 roboflow.com을 사용하여 바운딩 박스를 표시하고 이미지를 다양한 각도로 뒤집어 데이터 세트를 증강하였다. 훈련에서는 You Only Look Once 버전 8 (YOLOv8) 모델을 사용하였고, 사고 탐지에 있어서 평균 0.954의 정확도를 달성하였다. 제안된 모델은 비상시에 경보 전송을 용이하게 하는 실용적 의의를 가지고 있다. 또한, 효과적이고 효율적인 차량사고 탐지 메커니즘 개발에 대한 연구에 기여하고 스마트폰과 같은 기기에서 활용될 수 있다. 향후의 연구에서는 소리와 같은 추가 데이터의 통합을 포함하여 탐지기능을 정교화하고자 한다.

핵심주제어: 자동차 사고 탐지, 컴퓨터 비전, YOLOv8, CCTV

Abstract Car accidents occur as a result of collisions between vehicles, leading to both vehicle damage and personal and material losses. This study developed a vehicle accident detection model based on 2,550 image frames extracted from car accident videos uploaded to YouTube, captured by CCTV. To preprocess the data, bounding boxes were annotated using roboflow.com, and the dataset was augmented by flipping images at various angles. The You Only Look Once version 8 (YOLOv8) model was employed for training, achieving an average accuracy of 0.954 in accident detection. The proposed model holds practical significance by facilitating prompt alarm transmission in emergency situations. Furthermore, it contributes to the research on developing an effective and efficient mechanism for vehicle accident detection, which can be utilized on devices like smartphones. Future research aims to refine the detection capabilities by integrating additional data including sound.

Keywords: Car accidents detection, Computer vision, YOLOv8, CCTV

* Corresponding Author: cklee@kmu.ac.kr
+ This study was supported by a grant from Korea Tech. & Info. Promotion Agency (No. S3302201).
Manuscript received October 27, 2023 / revised November 21, 2023 / accepted November 25, 2023

1) 계명대학교 경영정보학과, 제1저자
2) 계명대학교 경영정보학과, 교신저자
3) 계명대학교 경영정보학과, 공동저자
4) ㈜문창, 공동저자

1. Introduction

Traffic accidents have tragically emerged as one of the foremost contributors to fatalities and lifelong disabilities on a global scale. As reported by the American Centers for Disease Control and Prevention(CDC) in 2023, a staggering 1.35 million individuals lose their lives annually as a result of road accidents encompassing a wide spectrum of vehicles, ranging from cars to buses and trucks. This alarming statistic underscores the profound toll exacted by these accidents on a worldwide scale.

The causes of such accidents are multifaceted and encompass a range of factors, including suboptimal road conditions and reckless driving behaviors, among others. These factors, often compounded by human error, significantly elevate the risk of accidents occurring. The dire consequences of traffic accidents extend far beyond mere statistics, as they result in the loss of precious human lives and inflict permanent disabilities on many. Therefore, in order to mitigate the occurrence of road accidents, various studies have been conducted, such as the detection of road cracks (Lee et al., 2019) and the implementation of a lane departure warning system through deep learning (Choi et al., 2019).

One critical factor in reducing fatalities and mitigating the severity of injuries is the timely relay of accurate information to first responders. When emergency services lack prompt access to vital accident details, there is a heightened risk of individuals suffering permanent disabilities as a result of delayed medical interventions(Nasr et al., 2016). Time plays a crucial role in saving human lives during accidents, and every second lost can have a profound impact on the outcomes for those involved(Ki et al., 2007).

In contemporary society, despite significant strides in communication technology, there persist regions and circumstances where the traditional reliance on word of mouth and manual reporting remains the primary, and sometimes sole, means of relaying crucial information, particularly in the aftermath of accidents. This reliance, however, carries inherent limitations and risks, especially in remote or underserved areas, blindspots where human presence is sparse, or during the silent hours of the night when accidents may occur unnoticed. Moreover, issues such as individuals lacking the means to report accidents due to the absence of a mobile phone, low battery, loss, or shock-induced inaction, further compound the challenges of timely emergency response.

Furthermore, a concerning aspect is the potential for false reports, either out of malicious intent or misunderstanding, which not only poses a significant burden on emergency services but also has the potential to divert their precious resources away from genuine emergencies. These circumstances underscore the need for a transformative approach to accident reporting and response.

It is within this context that the exploration and development of reliable automated accident detection systems gain paramount importance. Such systems have the potential to mitigate the aforementioned challenges and enhance the efficiency and effectiveness of emergency response efforts. By automating the detection and reporting of accidents, these systems can drastically reduce response times, thereby mitigating the severity of injuries and loss of life. Furthermore, they offer the capability to provide coverage in areas and situations where human reporting is unreliable or impossible, thereby ensuring a comprehensive safety net for all members of society. As a result, these systems not only improve the overall efficacy of emergency response services but also contribute to the advancement of public safety and welfare.

Numerous studies have been conducted to address the challenge of accident detection and response. Many of these studies have explored the utilization of computer sensors installed in vehicles, drivers' mobile phones, or even smartwatches. However, these methods have often proven to be less accurate, providing false detections or failing to deliver information in a timely manner(Abou et al., 2022). Thus, there is a pressing need for innovative approaches that can reliably and efficiently detect accidents while overcoming the limitations of existing systems.

Computer vision, particularly in object detection models, has surged in importance with the rapid advancement of Artificial Intelligence(AI) research. However, the demand for high-performance computers with potent GPUs to support these models has led to notable challenges and consequences. One key challenge is the escalating cost of acquiring high-end computers. As AI research and applications continue to expand, the demand for powerful hardware has surged, causing prices to soar. This can pose financial barriers for researchers, institutions, and organizations. Moreover, high-end computers tend to consume substantial energy during operation. As Gupta et al.(2021) have highlighted, this not only leads to higher electricity bills but also has environmental implications. The increased energy consumption strains power plants, leading to higher energy production and a larger carbon footprint. This exacerbates global warming, underscoring the environmental costs. This situation emphasizes the pressing need for energy-efficient computing solutions. Models that require less energy not only reduce costs but also promote sustainability by decreasing the environmental impact. Opting for such models is both economically prudent and environmentally responsible, offering a path towards a more sustainable and equitable future in AI and computer vision research.

In this study, we propose a model that leverages CCTV cameras mounted on poles to detect accidents. Our aim is to develop a fast and highly accurate accident detection system that can operate effectively irrespective of weather conditions. By employing the state-of-the-art You Only Look Once version 8(YOLOv8) model, we seek to capitalize on its superior performance in object detection tasks. YOLOv8 offers rapid processing speed and lightweight design, making it more accurate than alternative detection models while requiring fewer computational resources (Terven et al., 2023). Consequently, our proposed model demonstrates cost efficiency and lower energy consumption, aligning with the contemporary global efforts to reduce carbon emissions and environmental impact.

As demonstrated in Figure 1, the proposed model operates independently in a majority of scenarios. It leverages video feeds captured from CCTV cameras and seamlessly converts these continuous video streams into individual image frames. Subsequently, these image frames undergo a meticulous analysis by the model's algorithms. If the model detects any sign of an accident, it triggers an immediate alert, signaling the need for attention.

However, recognizing the possibility of false positive detections, the system incorporates a crucial verification step. Upon receiving an alert, the system promptly forwards this information to a dedicated traffic control center. Here, human verification comes into play, with trained personnel carefully assessing the provided data. This human verification step is vital in ensuring the accuracy and reliability of the accident detection process.

Once the control center verifies the occurrence of an accident, they are empowered to take appropriate action. This action may include promptly relaying critical information to the relevant emergency services, ensuring swift and effective response to the incident. However, in cases where the verification process reveals a false positive detection, the control center retains the discretion to dismiss the alert, preventing unnecessary and potentially costly responses.

This integrated system, showcases the synergy between automated accident detection and human oversight, striking a balance between efficient response and the need to minimize the risk of erroneous alerts. It exemplifies a comprehensive approach to accident detection and response that harnesses the power of technology while retaining the judgment and expertise of human operators.

The remainder of this paper is organized as follows: Section 2 provides a review of the related work undertaken in the field of accident detection. In Section 3, we outline the methodology adopted in this study. Section 4 presents the results obtained through our experimental evaluation, and in Section 5, we draw conclusions based on our findings.

2. Related work

2.1 Accident detection

Accident detection models have been the subject of numerous studies, aiming to enhance emergency response systems and improve the outcomes of traffic accidents. Some researchers have explored the integration of GPS and GSM technologies to develop car accident detection models(Topinkatti et al., 2015; Sharma et al., 2019), and employed sensors installed in vehicles that interact with mobile devices to send distress signals to the nearest emergency services when an accident occurs. However, these models have limitations, particularly in areas with no network coverage or when the victim's mobile device has a depleted battery,

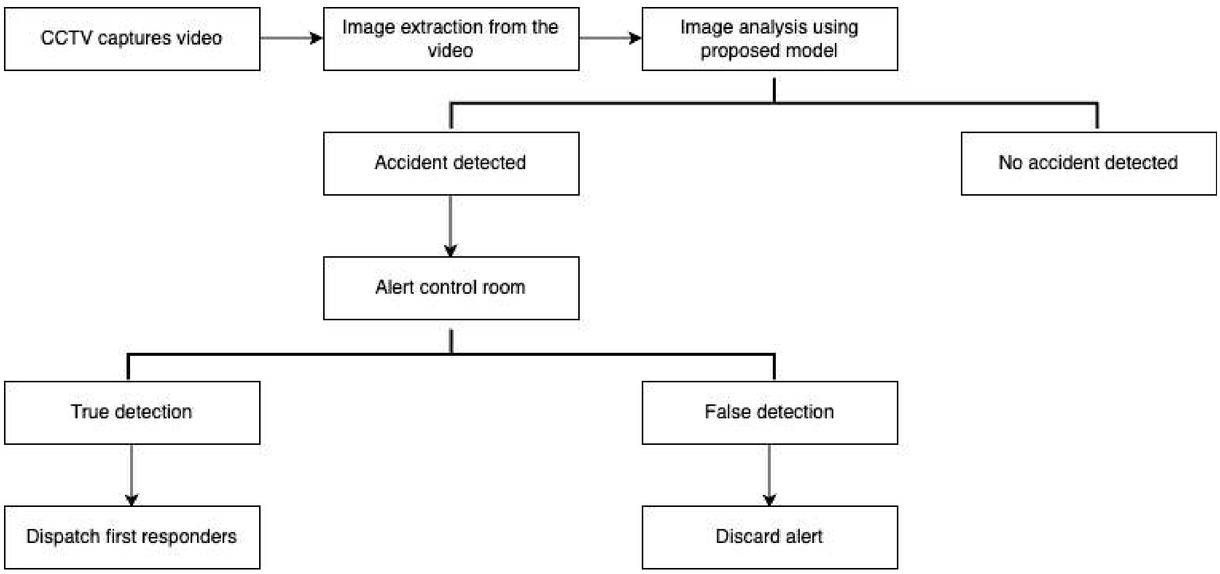

Fig. 1 Proposed model

compromising the effectiveness of the system.

In the realm of accident detection using computer vision techniques, researchers have employed older versions of the YOLO algorithm, such as YOLOv3(Gour et al., 2019; Desai et al., 2021; Tian et al., 2019). Furthermore, Choi et al. (2022) employed YOLOv5 to introduce a pedestrian detection model in crosswalks, utilizing multi-spectrum techniques with the aim of minimizing accidents at intersections. These models demonstrated promising results; however, they required substantial infrastructure investments due to their reliance on high-end GPUs for efficient performance. The resource-intensive nature of these models limits their scalability and practicality for widespread implementation.

Despite the advancements made in accident detection using computer vision, there is a need for more efficient and lightweight models that can operate effectively in diverse environments.

2.2 Comparison of detection algorithm

Table 1 Summary of the existing models

Model	Model summary
Fast R-CNN	Introduced by Girshick(2015), the Fast Region-based Convolutional Network(Fast R-CNN) method was developed with the primary aim of enhancing both the speed and accuracy of object detection. However, in a study conducted by Tian et al.(2019), Fast R-CNN exhibited inferior performance compared to other object detection models when it comes to accident detection. This performance gap can be attributed to certain limitations of the Fast R-CNN model. Notably, the model may struggle with occluded objects, distorted shapes, variations in illumination, or complex backgrounds. Moreover, it is inherently slower and computationally expensive, demanding high-end hardware for efficient operation as acknowledged by Ren et al.(2015) and Girshick (2015).
Faster R-CNN	Faster R-CNN, builds upon the improvements introduced by Fast R-CNN and further enhances the efficiency of the process by introducing a region proposal network. This innovation was first presented by Ren et al.(2016) and has since gained recognition. Despite advancements in speed, increased accuracy, and improved efficiency in object detection methods (Ren et al., 2015; Gandhi, 2018), similar to fast R-CNN, YOLOv8 still exhibits a relatively slower pace and computational expense. Its optimal operation necessitates high-end hardware configurations to ensure efficiency.
YOLOv3	YOLOv3, is a real-time object detection algorithm renowned for its ability to swiftly identify specific objects within videos, live camera feeds, or images. Operating on the backbone of deep convolutional neural networks(CNNs), YOLOv3 capitalizes on learned features to execute its object detection tasks efficiently. One notable application domain for this algorithm is in accident detection systems, as exemplified by Desai et al.(2021), where YOLOv3's capabilities were harnessed to enhance real-world safety applications. Furthermore, a comprehensive study by Tian et al.(2019) compared YOLOv3 with other object detection models and demonstrated its impressive performance, outpacing some competitors by threefold

Model	Model summary
	in terms of speed while maintaining superior accuracy in accident detection scenarios. However, YOLOv3 is not without its limitations. Research by Redmon and Farhadi(2018) and Liu et al.(2023) has pointed out its struggles with small objects and objects featuring substantial overlap with their surroundings.
YOLO-CA	Tian et al.(2019) introduced YOLO-CA, an innovative object detection model designed to address the challenges associated with detecting small objects, particularly in the context of car accidents. This model incorporates Multi-Scale Feature Fusion(MSFF) and a dynamic-weight loss function to enhance its small object detection capabilities. The experimental evaluation of YOLO-CA in the context of car accident detection yielded highly promising results. The proposed method achieved an impressive average precision rate of 90.02%. When compared to other object detection models, YOLO-CA demonstrated significant improvements in both accuracy and real-time performance.

2.3 YOLOv8

The YOLO model has gained significant popularity as an object detection and segmentation model since its initial release in 2015. Over the years, the YOLO model has undergone continuous improvements, incorporating new features and advancements. It has found application in various domains, including security and surveillance, autonomous vehicles, and medical imaging(Jiang et al., 2022; Redmon et al., 2016).

Among the developments of the YOLO model, YOLOv8, developed by Ultralytics, has emerged as a state-of-the-art solution in object detection tasks. YOLOv8 introduces several enhancements, including a more sophisticated backbone network, an anchor-free detection head, and a new loss function. These innovations have significantly improved image detection and segmentation capabilities (Ultralytics, 2023).

A notable improvement in YOLOv8 is its transition to an anchor-free model. Previous versions of YOLO relied on anchor boxes, which represented a predefined distribution of target objects. In contrast, YOLOv8 directly predicts the center of an object instead of relying on offsets from anchor boxes(Jacob, 2023). This shift allows for greater flexibility and adaptability to different datasets.

During training with YOLOv8, the model encounters slightly different variations of the input images in each epoch. This variation contributes to the increased accuracy of the model(Terven et al., 2023; Jacob, 2023). By exposing the model to diverse training samples, YOLOv8 becomes more robust and better equipped to handle various real-world scenarios.

The YOLOv8 architecture involves passing an input image through a backbone network, which is a convolutional neural network that extracts relevant image features. These features are then fed into a neck, a series of layers that combine and merge different features obtained from the backbone network. Finally, the merged features are processed by the head, where object detection predictions are made.

A distinguishing characteristic of YOLOv8

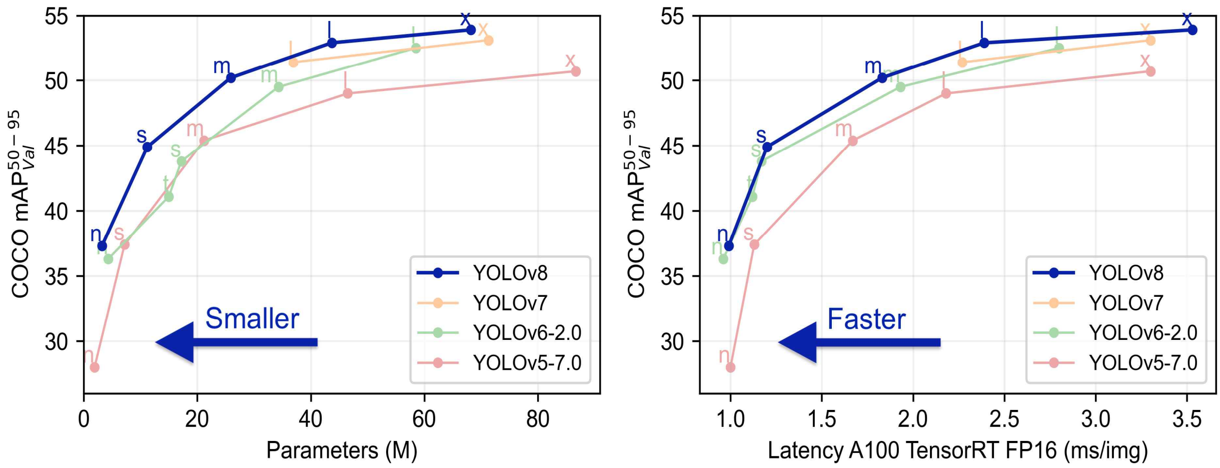

Fig. 2 Comparison between YOLOv8 and othe previous versions of YOLO source https://github.com/ultralytics/ultralytics

is its single-stage detection approach, unlike other models such as Region-Based Convolutional Neural Networks(R-CNN). This approach enables YOLOv8 to perform highly accurate real-time detections, making it a lightweight solution in terms of infrastructure requirements (Terven et al., 2023; Jacob, 2023).

The efficiency and accuracy of YOLOv8 make it an ideal choice for our proposed accident detection system, leveraging smart pole-mounted cameras for timely and reliable accident detection regardless of environmental conditions.

3. Methodology

3.1 Dataset

In this study, we employed a dataset consisting of 2,550 car accident image frames. These frames were sourced from YouTube accident compilation videos, offering a diverse and representative compilation of accident scenarios. To facilitate the annotation and organization of data for computer vision object detection models, we leveraged Roboflow.com, an invaluable tool.

Following the extraction of images from YouTube videos using a Python script, we utilized Roboflow.com to preprocess the images. This involved drawing bounding boxes on areas depicting occurrences of accidents. Such annotation is crucial for training computer vision models to accurately recognize and locate specific objects. Additionally, Roboflow provided a practical solution for partitioning the labeled data into training, validation, and test datasets. The images were thus distributed into three subsets: a training set comprising 1,785 images, a validation set of 510 images, and a testing set with 255 images.

One notable feature of Roboflow is its capacity to augment data by rotating images into various formats, such as 180 degrees or 90 degrees. This data augmentation technique generated additional images with variations in orientation, thereby enriching the training dataset. By exposing the model to diverse perspectives of the same objects, the algorithm could effectively learn to recognize objects under different conditions, ultimately

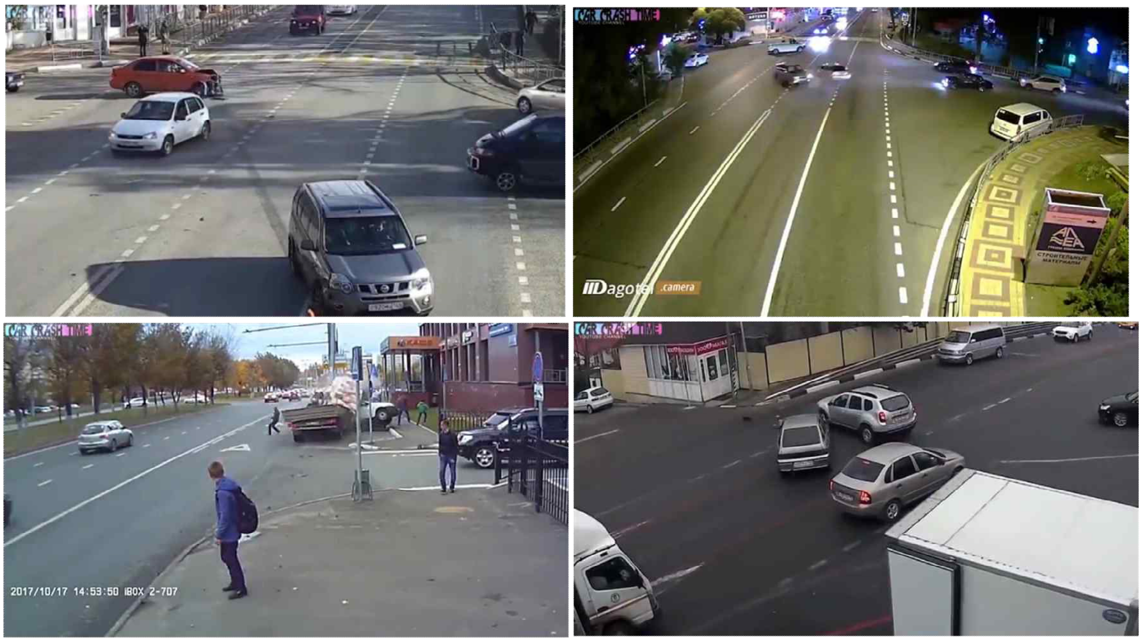

Fig. 3 Some accident images for training the model

enhancing its accuracy and robustness. Following the aforementioned data preparation steps, the preprocessed images were subsequently input into the YOLOv8 model for training.

3.2 Model training

In the pursuit of identifying the optimal accident detection model for this research study, pre-trained weights from YOLOv8 small and medium architectures were leveraged. The training process across all these models was standardized to consist of 50 epochs. During training, the input images were resized to a dimension of 800 pixels to ensure consistent and uniform processing. To facilitate these training and testing operations, Google Colab GPU was utilized, offering ample computational power for model training tasks. The Tesla T4 GPU provided 16GB of memory and CUDA version 12.0, enabling accelerated computation and faster training times. Due to hardware constraints, the model was tested in the same environment, implying that its performance can be expected to be favorable under similar hardware specifications. Furthermore, there is an anticipation that the model may exhibit even better performance when deployed on more advanced GPU configurations.

3.3 Performance metrics

In the performance analysis of the accident detection model, Mean Average Precision (MAP) is employed as a key evaluation metric. MAP is widely used in computer vision models and provides a measure between 0 and 1, with higher values indicating better performance. The calculation of MAP involves several components, including recall, Intersection over Union(IoU), precision, precision-recall curve, and average precision (AP)(Jacob, 2023).

Precision measures the accuracy of the model and is calculated using a specific formula (Padila et al., 2020). It quantifies the

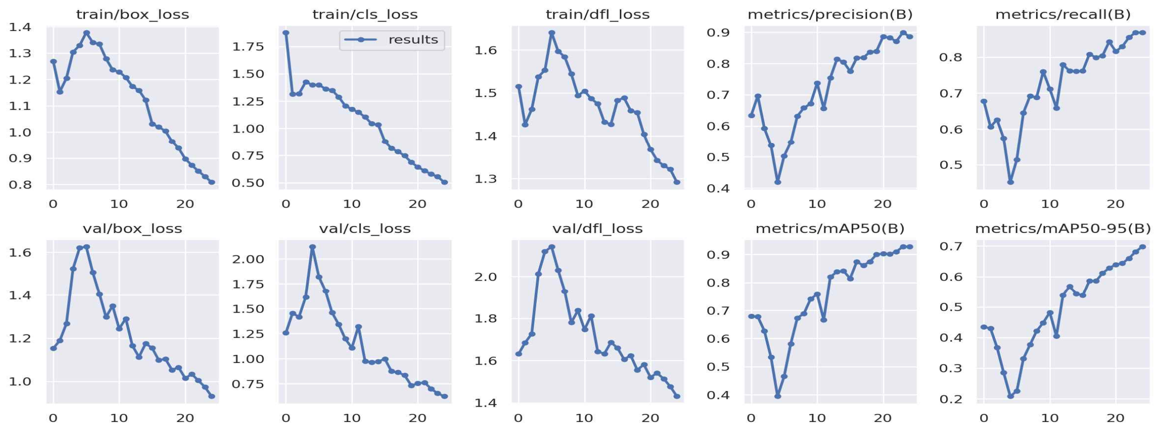

Fig. 4 Results attained by the trained model

proportion of correctly predicted positive instances out of all predicted positive instances. the formular for calculating precision is

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} = \frac{\text{TP}}{\text{all detections}}$$

Recall, on the other hand, assesses the model's ability to identify true positive instances. the formular is

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} = \frac{\text{TP}}{\text{all ground truths}}$$

Intersection over Union (IoU) is a metric that defines the overlap between predicted bounding boxes and ground truth bounding boxes. It is calculated by dividing the intersection area of the two boxes by the union area.

The precision-recall curve illustrates the relationship between precision and recall for the computer vision model. A larger area under the curve indicates a model with strong precision and recall, while a smaller area suggests weaker precision and recall. A model that maintains high precision as recall increases is considered to perform better.

Average Precision(AP) is calculated as the weighted mean of the precisions achieved at each threshold, with the weight derived from the increase in recall from the previous threshold. AP represents the area under the precision-recall curve(AUC). By comparing AP values, we can assess the performance of different accident detection models. the formula for calculating AP is

$$\text{AP} = \sum_n (R_n - R_{n-1}) P_n$$

Where Pn and Rn are the precision and recall at the nth threshold.

4. Results

4.1 MAP

In this study, we propose the use of the pre-trained YOLOv8 medium version as the accident detection model. The trained model achieved a mean average precision (MAP50) of 0.954, equivalent to 95.4%, and a box precision of 0.919, equivalent to 91.9%. The results, as illustrated in table 2, indicate that the proposed model outperformed other versions evaluated in this study and demonstrated better performance compared to previous results reported by other researchers. Due to hardware restrictions, we were not able to perform training on large version of YOLOv8 model.

Table 2 Results comparison between different accident detection models in (%)

Previous Models	Medium	Small
Fast R-CNN	80.3	49.2
Faster R-CNN	86.7	56.7
YOLOv3	91.1	67.7
YOLO-CA	91.5	76.5
YOLOv8	95.4	94.7

4.2 Precision

The model's performance in the validation set is noteworthy, with a box precision score of 0.918. This metric reflects the accuracy and precision with which the model identifies and predicts object bounding boxes, indicating a high level of precision in the model's object localization capabilities. Additionally, when examining the precision-confidence curve in figure 5, a remarkable achievement is observed. The model attains a precision score of 1 at a confidence score threshold of 0.9.

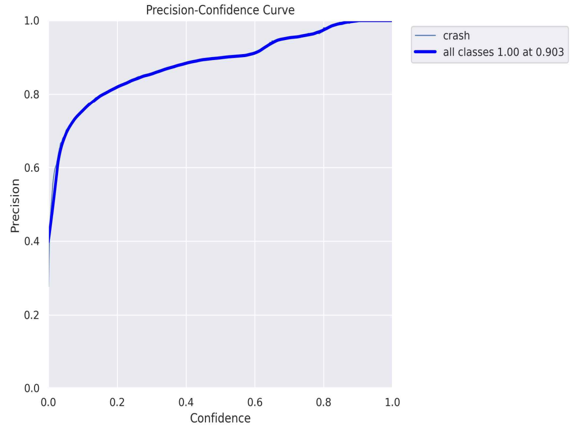

Fig. 5 Precision – Confidence curve

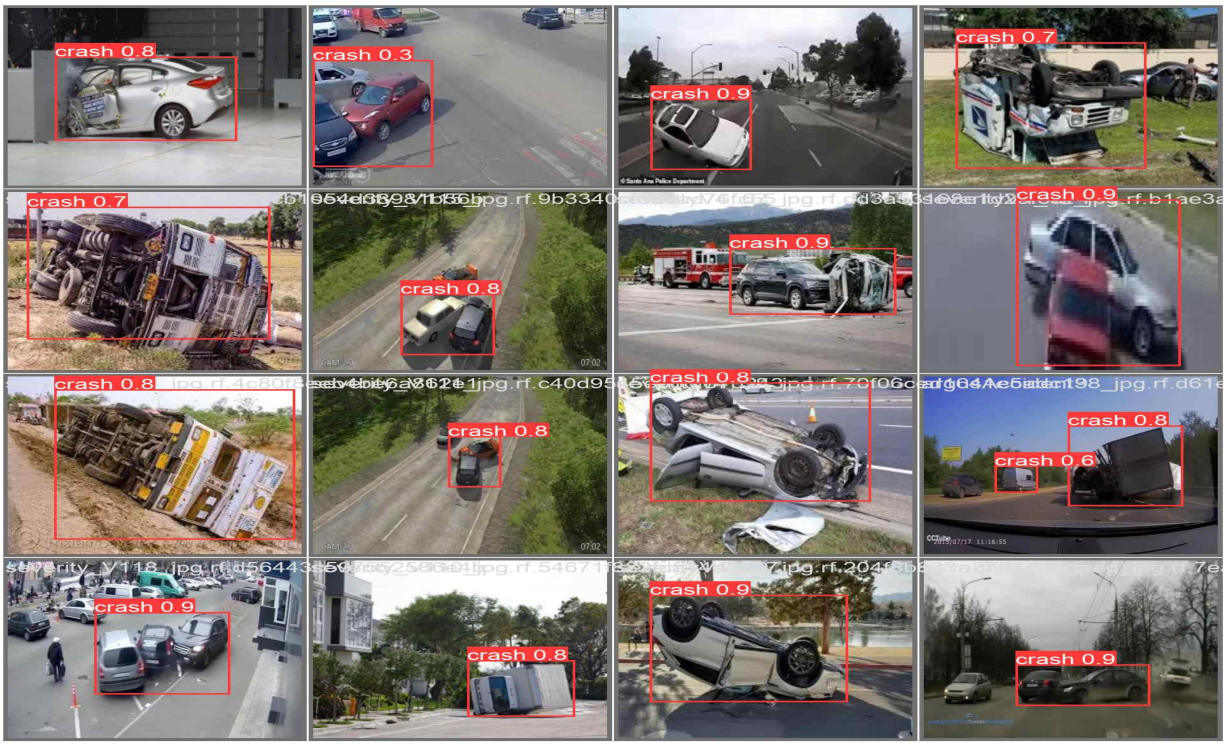

Fig. 6 Accident detection from validation set

This signifies that when the model makes predictions with a confidence score of 0.9 or higher. Such a high precision score at a relatively high confidence threshold underscores the model's robustness and reliability in its predictions, particularly in scenarios where precision is of utmost importance, even at the cost of some recall.

This level of precision is critical in applications such as object detection, where minimizing false positives and ensuring accurate localization are paramount.

Table 3 box precision and recall

Model	Box(P)	Recall	Speed
YOLOv8m	0.91	0.89	39.8 ms
YOLOv8s	0.94	0.85	20.4 ms

Table 3 provides a comprehensive comparison of the two models in terms of speed, recall, and box precision. The YOLOv8m model achieved a box precision of 0.91 and a recall score of 0.89. In contrast, the YOLOv8s model attained a higher box precision of 0.94 but a slightly lower recall score of 0.85. Additionally, in terms of detection speed, YOLOv8s outperformed YOLOv8m by processing, inferring, and post-processing each image 19.4 milliseconds faster. Upon analyzing the comparison, it is evident that while the smaller model exhibited faster processing with a commendable box precision score, it lagged behind in identifying true positives (recall score) when compared to the YOLOv8m model.

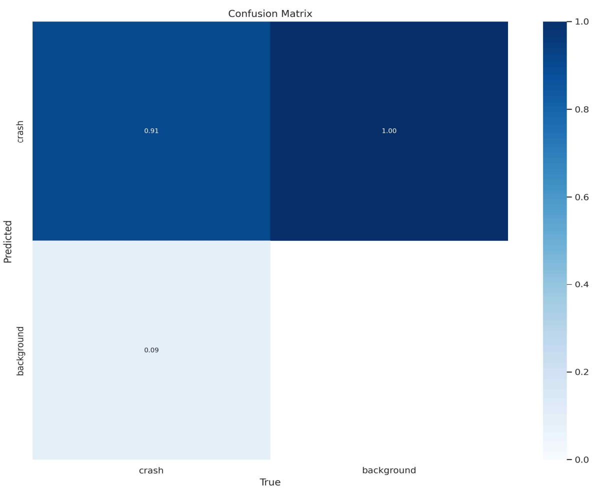

Fig. 7 Confusion Matrix

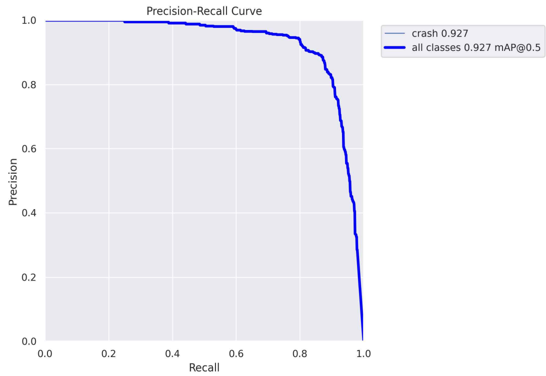

Fig. 8 Precision-Recall curve

4.3 Confusion Matrix

Figure 7 presents the results of the confusion matrix, offering valuable insights into the model's performance. Notably, the model achieved an accuracy rate of 0.91 in predicting accidents. This accuracy rate indicates that the model performed well in correctly identifying accident-related instances within the dataset. A high accuracy rate in predicting accidents signifies that the model's predictions align closely with the actual occurrences of accidents. The results underscore the model's effectiveness and its ability to make precise and dependable accident predictions, affirming its potential utility in real-world applications where the timely detection of accidents is essential for prompt response and intervention.

Fig. 9 Accident detection from test set

4.4 AUC

In Figure 8, the precision-recall curve showcases a consistently substantial area under the curve (AUC), reflecting the strong precision and recall characteristics of the model. This outcome carries significant implications, as it signifies the model's ability to maintain a high level of precision as recall increases. In essence, the model consistently achieves a high level of accuracy in detecting accidents while also ensuring that the instances it identifies as accidents are indeed genuine.

The sustained high precision across various recall levels is a noteworthy finding, indicating the model's effectiveness and reliability. It exemplifies the model's capability to consistently and accurately identify accidents, even as it becomes more inclusive in its recognition, without compromising precision.

4.5 Prediction on test set

Figure 9 provides a visual representation of the detected accidents from the test images, offering valuable insights into the model's performance this set of data remained entirely unseen by the model during both the training and validation phases. Notably, a striking majority of the detections depicted in the figure showcase confidence scores surpassing the 0.85 threshold. This observation holds significant significance, as it underscores the model's exceptional level of certainty in its predictions.

A confidence score exceeding 0.85 suggests that the model is highly confident in its ability to accurately identify accidents within the test images. Such a high degree of certainty speaks to the reliability and accuracy of the model's predictions. It indicates that the detected accidents can be considered highly credible and trustworthy, providing strong evidence of the model's competence in the task of accident detection.

5. Conclusion

In conclusion, this study represents a noteworthy advancement in the domain of accident detection through the application of a lightweight model, specifically the pre-trained YOLOv8 medium version. The model's performance stands out significantly, with a mean average precision(MAP50) reaching an impressive score of 0.954 and a box precision of 0.919. These results not only demonstrate the model's effectiveness but also surpass the performance of alternative versions and previous research findings, including those presented by Tian et al. in 2019.

The proposed system, distinguished by its utilization of a lightweight model, offers a compelling solution for rapid and precise accident detection, all while reducing the necessity for high-intensity computational infrastructures. This characteristic renders it suitable for real-world applications, where agility and efficiency are paramount. Notably, this attribute holds significant environmental implications, as the model's efficiency helps reduce the carbon footprint typically associated with resource-intensive computations. Moreover, the model's effectiveness during nighttime or in areas with limited human presence, such as blind spots, contributes to its potential impact in accident detection and response, aiding first responders in critical situations.

For academics and researchers, this study lays a solid foundation for further advancements in the realm of automated accident detection systems. It serves as a springboard for the exploration of innovative approaches that aim to reduce or eliminate the need for human intervention in accident detection such improving the model to combine audio and motion. The research findings point towards the potential for more accurate and comprehensive accident detection systems that integrate a range of features and data sources, moving beyond sole reliance on images.

However, it is imperative to acknowledge that, despite its proficiency in detecting true positives, this model necessitates human intervention. The human verification step remains integral to ensure that alerts are validated as genuine accidents and not false positives. This underscores the importance of maintaining a balance between automation and human oversight in critical applications like accident detection.

Furthermore, it is worth noting that this study relied on a relatively limited dataset of car accident image frames extracted from YouTube videos. Limitation to this is some of the videos had lower resolutions, while others were captured by car dash cameras.

Owing to GPU limitations, we encountered constraints in training the model with substantial weights derived from the YOLOv8 model. Consequently, our results were confined to the utilization of medium and small weights.

Future research could greatly benefit from a high end GPU and more extensive and diverse dataset captured from selected high resolution CCTV to comprehensively evaluate the model's performance and generalizability across various scenarios. Moreover, to diminish the need for human intervention and achieve a fully autonomous system, the model can be synergistically integrated with other models, such as a sound detection model. This additional model has the capability to discern the occurrence of an accident based on audio signals captured by a microphone, thereby enhancing overall accuracy.

This study lays a foundation for the development of efficient accident detection systems. It highlights the importance role of combining accuracy, speed, and environmental consciousness to enhance emergency response and augment road safety.

References

Abou, L., Fliflet, A., Hawari, L., Presti, P., Sosnoff, J. J., Mahajan, H. P., and Rice, L. A. (2022). Sensitivity of Apple Watch fall detection feature among wheelchair users, *Assistive technology*, 34(5), 619-625.

CDC. (2019). *Road Traffic Injuries and Deaths: A Global Problem tical Programming Tools*, https://www.cdc.gov/injury/features/global-road-safety/index.html

Choi, D. H., Lee, J., & Lee, D. (2022). A study on the detection of pedestrians in crosswalks using multi-spectrum. Journal of the Korea Industrial Information Systems Research, 27(1), 11-18.

Choi, S., Lee, K., Kim, K., & Kwak, S. (2019). Lane departure warning system using deep learning. Journal of the Korea Industrial Information Systems Research, 24(2), 25-31.

Desai, R., Jadhav, A., Sawant, S., and Thakur, N. (2021). Accident Detection Using ML and AI Techniques, *Engpaper Journal*.

Gandhi R. (2018). R-CNN, Fast R-CNN, Faster R-CNN, YOLO — Object Detection Algorithms,https://towardsdatascience.com/r-cnn-fast-r-cnn-faster-r-cnn-yolo-object-detection-algorithms-36d53571365e

Girshick, R. (2015). Fast r-cnn. In proceedings of the IEEE international conference on computer vision (pp. 1440-1448).

Gour, D., and Kanskar, A. (2019). Optimised YOLO: algorithm for CPU to detect road traffic accident and alert system, *International. J. Eng. Res. Technol*, 8. 160-163.

Gupta, U., Kim, Y. G., Lee, S., Tse, J., Lee, H. H. S., Wei, G. Y., and Wu, C. J. (2021, February). Chasing carbon: The elusive environmental footprint of computing. In 2021 IEEE International Symposium on High-Performance Computer Architecture (HPCA) (pp. 854-867). IEEE.

Jacob S, Francesco. (2020). *What is Mean Average Precision (mAP) in Object Detection?*, https://blog.roboflow.com/mean-average-precision/

Jiang, Peiyuan, Daji Ergu, Fangyao Liu, Ying Cai, and Bo Ma. (2022). A Review of Yolo algorithm developments, *Procedia Computer Science*, 199. 1066-1073.

Ki, Y. K., and Lee, D. Y. (2007). A traffic accident recording and reporting model at intersections, *IEEE Transactions on Intelligent Transportation Systems*, 8(2), 188-194.

Lee, M. H., Nam, K. W., & Lee, C. W. (2019). Crack Detection on the Road in Aerial Image using Mask R-CNN. Journal of the Korea Industrial Information Systems Research, 24(3), 23-29.

Liu, B., He, F., Du, S., Li, J., & Liu, W. (2022). An advanced YOLOv3 method for small object detection. arXiv preprint arXiv:2212.02809..

Nasr, E., Kfoury, E., and Khoury, D. (2016). An IoT approach to vehicle accident detection, reporting, and navigation, *In 2016 IEEE international multidisciplinary conference on engineering technology*, 231-236.

Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767.

Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). You only look once:

Unified, real-time object detection, *In Proceedings of the IEEE conference on computer vision and pattern recognition*, 779-788.

Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural information processing systems, 28.

Sharma, S., and Sebastian, S. (2019). IoT based car accident detection and notification algorithm for general road accidents, *International Journal of Electrical & Computer Engineering*, 9(5). 2088-8708.

Terven, J., and Cordova-Esparza, D (2023). A comprehensive review of YOLO: From YOLOv1 to YOLOv8 and beyond, *arXiv preprint*, arXiv:2304.00501.

Tian, D., Zhang, C., Duan, X., and Wang, X. (2019). An automatic car accident detection method based on cooperative vehicle infrastructure systems, *IEEE Access*, 7. 127453-127463.

Topinkatti, A., Yadav, D., Kushwaha, V. S., and Kumari, A. (2015). Car accident detection system using GPS and GSM, *International Journal of Engineering Research and General Science*, 3(3). 1025-1033.

Ultralytics. (2023). *Ultralytics YOLOv8 documentation*, https://docs.ultralytics.com

Marwa Chacha Andrea
He completed a bachelor's degree in Information Technology from the Institute of Finance Management in 2015. In 2023 he obtained a master's degree in MIS from Keimyung University where he is currently a Ph.D. student.

Choong Kwon Lee
He is a professor in MIS at Keimyung University. His research interests are big data and IT workforce.

Yang Sok Kim
He is an Associate Professor in MIS at Keimyung University. He received his Ph.D. from University of Tasmania, Australia. His research interests are Machine Learning and Recommender Systems.

Mi Jin Noh
She is an assistant professor of Business Big Data at Keimyung University. Her research interest is data analysis in social network.

Sang Il Moon
He is studying at the Graduate School of Business Administration in Yeungnam University, South Korea. He is the manager of the Strategic Planning Department of Moon Chang Co., Ltd. His research interest is Integrated Water Management Systems.

Jae Ho Shin
He received his B.S. degree from MIS in Keimyung University, South Korea. He has been serving as the head of the IT Strategy Division of Moon Chang Co., Ltd. His research interests are AIoT and Smart City.