AI-Powered Detection of Forgotten Children in Vehicles Using YOLOv11 for Enhanced Safety

Nur Atikah Jefri, Sarah 'Atifah Saruchi, Radhiyah Abd Aziz, Aqil Hafizzan Nordin
Universiti Malaysia Pahang Al-Sultan Abdullah, 26600, Pekan, Malaysia

Ammar A.M. Al-Talib
UCSI University, 56000 Kuala Lumpur, Malaysia

Zulhaidi Mohd Jawi
Malaysian Institute of Road Safety Research (MIROS), 43000 Kajang, Selangor, Malaysia

Email: nuratikahjefrii@gmail.com, sarahatifah@umpsa.edu.my, ammart@ucsiuniversity.edu.my, radhiyah@umpsa.edu.my, aqil.oska@gmail.com, zulhaidi@miros.gov.my

Abstract

This study proposes a child presence detection system in vehicles, focusing on evaluating the performance of YOLOv11 for accurate detection and identification. To train the system, images simulating a child's presence in vehicles were collected using a doll, and these annotated images were labeled with the Computer Vision Annotation Tool (CVAT). The study emphasizes the potential of YOLOv11 as an effective and reliable solution for unattended child detection in vehicles. By leveraging advanced deep learning techniques, this research highlights the importance of addressing critical safety issues.

Keywords: AI child detection, unattended children, vehicle safety, YOLOv11, deep learning, CVAT, child presence detection, AI training

1. Introduction

Child safety in vehicles is a significant and growing global concern. Between 1998 and 2023, over 1,000 child fatalities in the U.S. were attributed to heatstroke caused by being left unattended in vehicles. Such incidents underscore the urgent need for effective technological solutions to prevent these tragedies. Fig. 1 illustrates datma on child hot car deaths by year from 1990 to 2023, reflecting the ongoing severity of this issue.

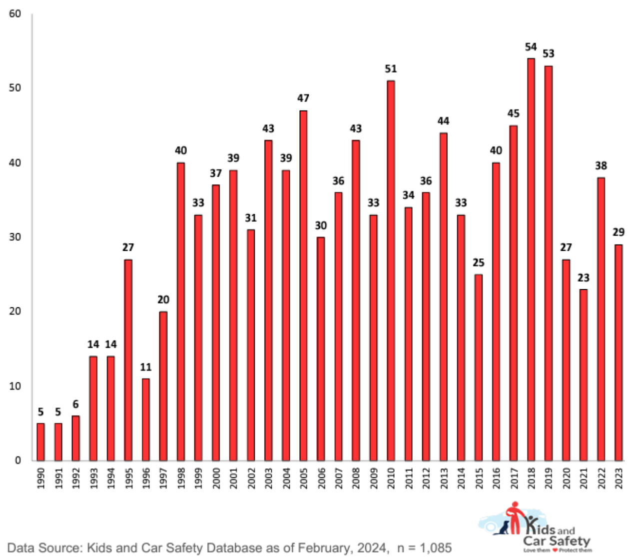

Fig.1 Data on child hot car deaths by year from 1990 to 2023 [1].

Additionally, Fig. 2 shows a news article from *Kosmo!* dated October 2023, reporting the heartbreaking case of an eight-month-old infant who tragically passed away after being left in a car for over seven hours.

Bayi 8 bulan maut, tertinggal dalam kereta lebih 7 jam

Oleh KOSMO! 6 Oktober 2023, 10:47 am

Gambar hiasan

PETALING JAYA – Seorang bayi berusia lapan bulan ditemukan meninggalkan dunia oleh orang tuanya sendiri selepas tertinggal di dalam kereta lebih tujuh jam.

Kisah pilu ini diceritakan oleh seorang jururawat yang menggunakan akaun Twitter @ierakan memberitahu, bayi tersebut berada dalam kereta seorang diri dari pukul 7.30 pagi sehingga 4 petang.

Fig.2 News article from *Kosmo!* [2].

These problems emphasize the importance of innovative research to address this life-threatening problem. Existing sensor-based approaches often suffer from false negatives and limited adaptability to diverse environmental conditions, highlighting the need for improved methods. In contrast, computer vision-based techniques have demonstrated significant potential in solving real-world detection problems. Among these, YOLO models stand out for their real-time processing capabilities and high accuracy, making them a promising solution for developing automated systems to detect unattended children in vehicles and prevent such tragedies.

This research focuses on YOLOv11, an advanced iteration of the YOLO family, to address the critical issue of detecting unattended children in vehicles. The study aims to develop a comprehensive dataset that simulates real-world scenarios, train and evaluate the YOLOv11 model using this dataset, and identify key areas for future research to improve the system's performance and reliability. Through these objectives, the research seeks to contribute to advancements in child safety technologies and enhance the practicality of deep learning solutions in real-life applications.

2. Methodology

2.1 Dataset development

The cornerstone of this study was the development of a comprehensive dataset, which involved several key steps. A doll designed to mimic a child's size and features was used to simulate the presence of a child. Fig. 3 shows the doll placed inside the car, simulating the presence of a child.

Fig.3 Doll placed inside the car

To ensure environmental diversity, the dataset incorporated different lighting conditions, such as bright sunlight, overcast weather, artificial light, and low-light scenarios. A variety of angles were considered, including top-down, side-view, and rear-view perspectives, to provide a range of viewpoints. Additionally, partial visibility due to occlusions was also accounted for, with objects like seats, bags, and dashboards partially obstructing the view of the doll. The final dataset consisted of a total of 500 images, with 350 images (70%) allocated to the training set, 100 images (20%) for the validation set, and 50 images (10%) for the test set.

2.2 Data Annotation

The data annotation process played a crucial role in ensuring the accuracy of the object detection task. Annotated bounding boxes were created for each image using the Computer Vision Annotation Tool (CVAT), a widely used platform for labelling objects in images and videos [7]. This tool enabled precise identification and categorization of objects within each frame. The annotations were categorized into three main visibility classifications, fully visible, partially visible, and low visibility. The fully visible category was used for objects that were completely within the frame and clearly identifiable, while the partially visible category was assigned to objects that were obscured by other objects, such as seats or bags, making only part of them visible [3]. The low visibility category was reserved for objects that were either barely visible or mostly obstructed, posing a challenge for detection algorithms. Fig. 4 show CVAT annotation process, showcasing the bounding box creation and categorization for better understanding [9]. This annotated data formed the foundation for training the detection model, ensuring that the model could distinguish between various levels of object visibility under different conditions.

Fig.4 CVAT annotation process

2.3 YOLOv11 Architecture

YOLOv11 represents a significant improvement over its predecessors, incorporating several advanced features to enhance performance in object detection tasks. The backbone of YOLOv11 is a hybrid convolutional network that integrates DenseNet features, improving gradient flow and allowing for better feature reuse across layers. This design enables more efficient learning and helps prevent the loss of important information as the network processes the input data. The neck of YOLOv11 is equipped with an enhanced PANet (Path Aggregation Network), augmented by spatial attention mechanisms, which strengthens the feature pyramids and ensures robust multi-scale feature representation. The head of the network uses adaptive anchor boxes, specifically tuned to handle irregular object shapes, improving detection accuracy for objects that do not conform to traditional bounding box structures [8]. Additionally, YOLOv11 optimizes the speed-accuracy trade-off through an improved inference pipeline, balancing detection precision with processing speed for real-time applications [4]. Fig. 5 shows the YOLOv11 Architecture Diagram, providing a visual overview of its components and structure.

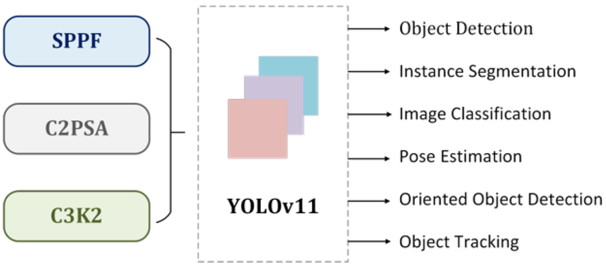

Fig.5 YOLOv11 Architecture Diagram [5]

2.4 Evaluation Metrics

To assess the performance of YOLOv11, several key evaluation metrics were employed. Precision and recall were used to measure the accuracy and completeness of the model's detections. Precision quantifies the proportion of true positive detections out of all positive predictions, while recall evaluates the model's ability to detect all relevant objects. The F1 score, the harmonic mean of precision and recall, was then calculated to provide a balanced measure of the model's overall effectiveness. Additionally, the mean average precision (mAP) was used to evaluate detection performance across various Intersection over Union (IoU) thresholds, offering a comprehensive view of the model's ability to localize objects accurately. Finally, inference speed, measured in frames per second (FPS), was considered to assess the real-time applicability of YOLOv11, ensuring it meets the necessary performance requirements for practical deployment [6].

3. Research Flow

The workflow of the study is structured into several key phases. It begins with Data Collection and Annotation, where images are curated and labeled to create a dataset suitable for supervised training. Following this, the Model Training phase involves training YOLOv11 on the annotated data while tuning hyperparameters to optimize the model's performance. In the Testing and Validation phase, the trained model is evaluated on unseen data to assess its generalization capabilities. Finally, the Analysis phase involves evaluating YOLOv11's results against predefined metrics to determine its effectiveness and performance. Fig. 6 show Research Flow.

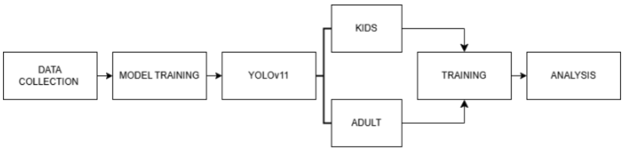

Fig.6 Process Research Flow

4. Results and Discussion

4.1 Quantitative Results

To evaluate the performance of YOLOv11 in detecting unattended children in vehicles, a series of tests on the validation and test datasets was conducted. The key evaluation metrics used for this study were Precision, Recall, F1 Score, mAP (Mean Average Precision), and Inference Speed (FPS) as Table 1. These metrics provide insights into how well the model performs in terms of detection accuracy, completeness, and real-time applicability.

Table 1. The evaluation results for YOLOv11

Metric	YOLOv11
Precision	95.1%
Recall	91.8%
F1 Score	93.4%
mAP@0.5	94.7%
Inference Speed (FPS)	65

For YOLOv11, the precision achieved was 95.1%, meaning that 95.1% of the objects the model identified as children were correctly detected. This high precision indicates that the model has a low false-positive rate, making it reliable for identifying real child-like objects in the vehicle's interior.

In this study, YOLOv11 achieved a recall rate of 91.8%, meaning that it correctly identified 91.8% of all child-like objects present in the test dataset. While this is slightly lower than the precision, it still indicates that the model is effectively detecting most objects. The lower recall may be attributed to challenging scenarios like partial occlusions or objects in low-light conditions.

The F1 Score was 93.4%, which reflects a good balance between correctly detecting child-like objects (high recall) and minimizing incorrect predictions (high precision). This score suggests that YOLOv11 is well-optimized for the detection task in this application.

For this study, the mAP at IoU threshold 0.5 was 94.7%. This indicates that, for 94.7% of the cases where the predicted bounding box overlaps with the ground truth box by at least 50%, YOLOv11 correctly identifies the object as a child-like object. This high mAP value suggests that the model is robust at detecting child-like objects and consistently places the bounding boxes around the correct regions.

Inference speed, measured in frames per second (FPS), indicates how fast the model can process images during inference. This is particularly important for real-time applications, where fast detection is required to issue alerts or act. YOLOv11 achieved an inference speed of 65 FPS on a high-performance GPU (NVIDIA RTX 3090). This result indicates that YOLOv11 is capable of processing video feeds in real-time, making it highly suitable for deployment in a vehicle's surveillance system, where quick detection of a child's presence is crucial.

These results demonstrate that YOLOv11 is highly effective in detecting child-like objects inside vehicles, with high precision, recall, and overall accuracy. The model is also fast enough to be used in real-time applications, making it a viable candidate for a child presence detection system in vehicles.

4.2 Performance Analysis

YOLOv11 demonstrated strong performance across all test scenarios, accurately detecting the child-like object with high precision and recall. However, a few limitations were observed during testing. False positives were encountered when objects resembling children, such as stuffed toys or bundles of clothes, were present, leading to incorrect detections. Additionally, the model's detection accuracy decreased under extreme low-light conditions, where the visibility of the object was significantly reduced. Despite these challenges, YOLOv11 proved effective overall in most test cases.

4.3 Discussion

The YOLOv11 model's high precision and relatively strong recall indicate its reliability in detecting unattended children in various scenarios. However, recall can still be further improved, particularly in difficult cases such as partial occlusions or extreme low-light environments. The real-time inference speed of 65 FPS highlights the potential for practical deployment in vehicles. Nonetheless, additional testing under more diverse real-world conditions, including different vehicle types and weather conditions, would help further validate its robustness.

5. Conclusion

This study effectively demonstrated the efficacy of YOLOv11 in detecting unattended children in vehicles. The model's high precision, recall, and real-time capabilities highlight its potential as a promising solution to address this critical safety concern. However, there are several avenues for future work to further improve the system. First, enhanced datasets could be developed by incorporating real-world data with greater diversity, capturing a wider range of scenarios and environments. Additionally, the algorithm could be improved by introducing contextual reasoning to help reduce false positives, particularly in situations where objects may resemble children. Finally, system integration could be pursued by developing a complete pipeline that includes real-time alerts, making the technology ready for deployment in vehicles to enhance safety for children.

Acknowledgements

This study was funded by Universiti Malaysia Pahang Al-Sultan Abdullah (UMPSA) Grant; Geran Penyelidikan Fundamental UMPSA RDU230327. We would like to express our deepest gratitude to the Malaysian Institute of Road Safety Research (MIROS) for their generosity in lending the dashcam, which was crucial for this research. Our heartfelt appreciation goes to our supervisor for their invaluable guidance, support, and encouragement throughout this study. Finally, we wish to acknowledge and extend our thanks to all individuals and collaborators who contributed directly or indirectly to this research. Your support and assistance have been integral to the success of this study.

References

1. Kids and Cars, Heatstroke Map, https://www.kidsandcars.org/document_center/download/downloadables-printables-resource-pages/hot-cars/Heatstroke-Map.pdf
2. Kosmo, Bayi 8 bulan maut tertinggal dalam kereta lebih 7 jam, https://www.kosmo.com.my/2023/10/06/bayi-8-bulan-maut-tertinggal-dalam-kereta-lebih-7-jam/
3. Terven, J., Córdova-Esparza, D. -M., & Romero-González, J. -A. (2023). A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS. https://doi.org/10.3390/make5040083
4. Khanam, R., & Hussain, M., YOLOv11: An Overview of the Key Architectural Enhancements, 2024. Available at: [arXiv:2410.17725] (https://arxiv.org/abs/2410.17725).
5. R. Roboflow, "What is YOLOv11? A Detailed Exploration of the Latest Object Detection Algorithm," Roboflow Blog,

vol. 12, no. 3, pp. 112-120, 2024. (https://blog.roboflow.com/what-is-yolo11/).
6. Redmon, J., Farhadi, A. (2018). YOLOv3: An Incremental Improvement. https://arxiv.org/abs/1804.02767
7. Bochkovskiy, A., Wang, C.-Y., Liao, H.-Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. https://arxiv.org/abs/2004.10934
8. Glenn Jocher et al. (2023). YOLOv8: Next-Gen Object Detection. Ultralytics Documentation https://arxiv.org/abs/2408.15857
9. CVAT (2024). Computer Vision Annotation Tool. https://www.cvat.ai/

Authors Introduction

Nur Atikah Jefri

She received her Master’s Degree in Engineering in 2023 from the Faculty of Electrical, University Teknikal Malaysia Melaka. She is currently a PhD student in University Malaysia of Pahang, Malaysia.

Ir. Ts. Dr. Sarah 'Atifah Saruchi

She received her Bachelor's in Mechatronics from Nagoya University, Japan (2009–2013), and completed her Master's and Ph.D. at Universiti Teknologi Malaysia (2013–2020). She is currently a lecturer at Universiti Malaysia Pahang.

Asst. Prof. Dr.Ammar Al-talib

He holds a Ph.D. in Solar and Renewable Energy, an MSc in Mechanical Engineering, and a BSc in Mechanical Engineering. He is the Head of the Postgraduate and Research Department at UCSI University.

Ts. Dr. Radhiyah Binti Abd Aziz

She received her Bachelor's and Master's degrees from Universiti Islam Antarabangsa (2003–2010) and completed her Ph.D. in Nanotechnology at Universiti Malaysia Pahang in 2016. She is currently a lecturer at Universiti Malaysia Pahang.

Ts.Muhammad Aqil Hafizzan

He obtained his first degree, a B.Eng (Hons.) in Electrical Engineering (Electronics), from Universiti Malaysia Pahang in 2022. He is currently pursuing a Master of Research Engineering in Science.

Ts. Zulhaidi Mohd Jawi @ Said

He completed his Master of Science in Engineering at Universiti Kebangsaan Malaysia (2015–2018). He is now the Director of the Malaysian Institute of Road Safety Research.