

Article

A Comprehensive Analysis of Real-Time Car Safety Belt Detection Using the YOLOv7 Algorithm

Lwando Nkuzo *, Malusi Sibiya *  and Elisha Didam Markus

Department of Electrical, Electronic and Computer Engineering, Central University of Technology, Free State, Bloemfontein 9300, South Africa

* Correspondence: lwandonkuzo@gmail.com (L.N.); mrkirchoffs@gmail.com (M.S.)

Abstract: Using a safety belt is crucial for preventing severe injuries and fatalities during vehicle accidents. In this paper, we propose a real-time vehicle occupant safety belt detection system based on the YOLOv7 (You Only Look Once version seven) object detection algorithm. The proposed approach aims to automatically detect whether the occupants of a vehicle have buckled their safety belts or not as soon as they are detected within the vehicle. A dataset for this purpose was collected and annotated for validation and testing. By leveraging the efficiency and accuracy of YOLOv7, we achieve near-instantaneous analysis of video streams, making our system suitable for deployment in various surveillance and automotive safety applications. This paper outlines a comprehensive methodology for training the YOLOv7 model using the labelImg tool to annotate the dataset with images showing vehicle occupants. It also discusses the challenges of detecting seat belts and evaluates the system's performance on a real-world dataset. The evaluation focuses on distinguishing the status of a safety belt between two classes: "buckled" and "unbuckled". The results demonstrate a high level of accuracy, with a mean average precision (mAP) of 99.6% and an F1 score of 98%, indicating the system's effectiveness in identifying the safety belt status.

Keywords: vehicle occupant safety; safety belt detection status; YOLOv7; real-time detection; computer vision; object detection; deep learning; buckled; unbuckled



Citation: Nkuzo, L.; Sibiya, M.; Markus, E.D. A Comprehensive Analysis of Real-Time Car Safety Belt Detection Using the YOLOv7 Algorithm. *Algorithms* **2023**, *16*, 400. <https://doi.org/10.3390/a16090400>

Academic Editor: Arslan Munir

Received: 27 July 2023

Revised: 16 August 2023

Accepted: 21 August 2023

Published: 23 August 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The utilization of a safety belt is highly crucial in the mitigation of severe injuries and the substantial reduction in the probability of fatalities [1–4]. According to the research conducted by Khohli and Chadha, it has been estimated that a total of 3287 individuals perish on a daily basis as a result of accidents [5]. Leland et al. [6] also submit that the estimated annual number of fatalities resulting from automobile accidents is approximately 1.35 million. This phenomenon persists despite the presence of safety belt monitoring systems in modern vehicles, which serve as reminders for occupants to fasten their seat belts. In South Africa (SA), there exists a specialized branch of law enforcement that is committed to the enforcement of safety belt usage and the enhancement of overall safety measures. However, a significant portion of individuals traveling in vehicles choose not to secure their seatbelts. The subject of safety belt detection is of great significance owing to its capacity to safeguard human lives.

At present, there is a scarcity of efficient approaches to guarantee and supervise the proper fastening of safety belts, with the exception of deploying traffic officers on highways [7] and relying on safety belt warning systems that can be easily circumvented by vehicle occupants. The existing methodology is characterized by its laborious and ineffective nature, primarily attributed to the limited number of officers allocated for inspection responsibilities [7]. Additionally, a noteworthy challenge arises from the ability of vehicle occupants to manipulate the safety belt system by either fastening the belt behind their backs or utilizing a device known as a "seatbelt warning stopper" to bypass the system [2].

The existing safety belt warning systems demonstrate limited efficacy. At times, motorists observe traffic officers from a considerable distance and subsequently opt to secure their safety belts at that particular juncture. Consequently, there has been an increasing fascination with the automated identification and implementation of safety belt regulations [8]. Scholars have directed their attention towards the development of computer vision algorithms in order to meet the worldwide need of traffic safety agencies for automated safety belt detection solutions that are not solely dependent on onboard sensors.

Scholars have addressed the issue of safety belt compliance through two distinct approaches. The first approach involves the installation of cameras along highways to monitor the usage of safety belts [3–7]. However, the installed cameras may have to consider the fact that there are two modes of computer vision models (single-stage and two-stage mode), and, therefore, the Stage 1 computer vision models are the ones recommended for real-time detection [9,10]. Other authors, such as Hosameldeen, Chen et al., and Tianshu et al., also considered the installation of cameras along highways [11–13]. The second approach entails the placement of cameras within vehicles to constantly monitor the act of fastening safety belts [2,14,15]. Both proposed solutions encounter various challenges. The implementation of safety belt monitoring systems on highways effectively addresses the issue of limited law enforcement personnel, yet it fails to address the underlying concern that drivers may selectively comply with safety belt regulations in areas where they are aware of the presence of such surveillance cameras. The potential exists for the onboard monitoring system to effectively enforce the utilization of a car safety belt. Nevertheless, the process of onboard monitoring encounters the persistent challenge of fluctuating illumination, the presence of blurry images resulting from uneven road conditions, and the obstruction caused by objects within the vehicle.

As previously indicated, the existing design of the safety belt monitoring mechanism is susceptible to manipulation. The safety belt mechanism, as depicted in Figure 1, is a prevalent feature in the majority of automobiles and comprises various constituent parts, including webbing, latch plate, buckle, pillar loop, and retractor. The buckle component is outfitted with a metallic sensor designed to detect the presence of the latch plate. In addition, the seat is equipped with a pressure sensor that is responsible for measuring the weight of the occupant when they are seated. An alarm is activated in the event that either the driver, the front passenger, or both individuals are not utilizing seat belts and their weight is detected on the seat. The alarm continues until a buckle is fastened and it detects the presence of a latch plate. Nevertheless, this particular approach for identifying safety belts has its constraints, as it may not effectively distinguish between a securely fastened buckle and a latch plate that has been inserted behind the individual's back. Furthermore, there is an alternative product known as the "seatbelt warning stopper" that functions as a metallic device inserted into the buckle to mimic the presence of a latch plate [2]. The utilization of vehicle occupant weight as a measure of seat occupancy is prone to inaccuracies, as situations may arise in which a seat belt alarm is activated due to the presence of a pet or a substantial object positioned on the seat.

This paper presents a pioneering approach to address a critical safety gap in current safety belt systems through the innovative integration of computer vision technology. By leveraging advanced image analysis techniques, we have developed a novel YOLOv7 model that effectively detects the accurate status of car seat belt buckling, overcoming challenges posed by deceptive practices such as routing the belt behind the occupant's back or utilizing aftermarket devices to create false buckle signals. Unlike conventional systems susceptible to manipulation, our solution employs real-time visual data to ensure reliable and tamper-resistant seat belt status detection. The model's continuous monitoring capability, signalled by an audible alert, not only deters unsafe practices but also significantly contributes to mitigating potential fatalities by promoting genuine seat belt usage. This research marks a significant advancement in automotive safety by harnessing computer vision's precision and responsiveness to bolster seat belt enforcement, ultimately fostering a safer environment for both passengers and drivers.

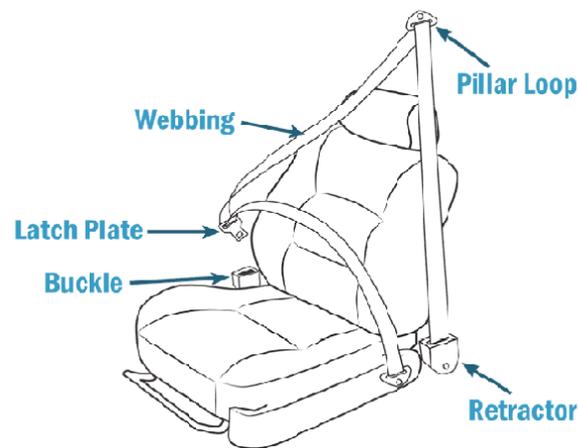


Figure 1. Example of a safety belt mechanism description.

As highlighted, this paper utilizes computer vision object detection. The domain of computer vision in object detection can be categorized into two primary methodologies: single-stage object detection and two-stage object detection [9]. Figure 2 illustrates the fundamental disparity between these methodologies, wherein the key distinction lies in the incorporation of a Region Proposal Network (RPN) within the two-stage detector. Single-stage detectors, such as YOLOv7, achieve higher computational efficiency compared to two-stage detectors by not utilizing an RPN. According to the findings of Wang et al. [10], YOLOv7 demonstrates superior performance compared to current object detection models in terms of both speed and accuracy. According to the authors, YOLOv7 demonstrates a variable speed performance, ranging from 5 to 160 frames per second (FPSs), while also achieving the highest level of accuracy with an average precision (AP) of 56.8%. This places YOLOv7 as the leading real-time object detection system in terms of accuracy.

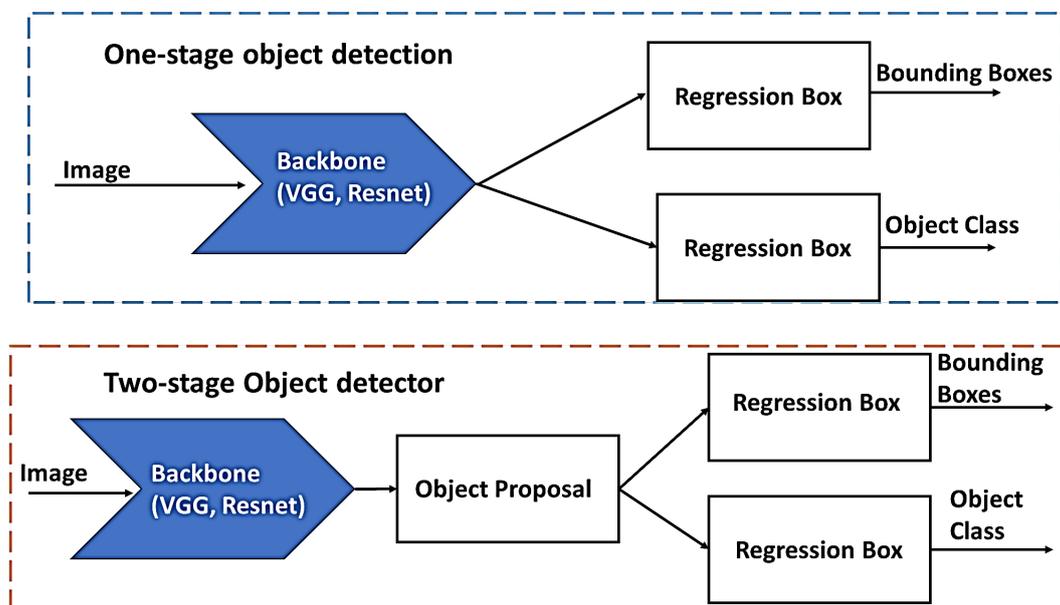


Figure 2. Example of a one-stage and two-stage detection flow diagram.

The main objectives of this research are as follows:

- To develop an in-cabin real-time YOLOv7 model for detecting the safety belt status of vehicle occupants and drivers.
- To train the model using a Graphic Processing Unit (GPU).
- To deploy the trained weights on an Nvidia Jetson Nano development board.

- To detect the safety belt buckle status even in partially occluded conditions, under different lighting conditions, in blurry images caused by vibration, and in clothing with similar colours as the safety belt.

The subsequent sections of this paper are organized in the following manner: Section 2 of this paper examines the relevant literature and provides a comprehensive explanation of the proposed YOLOv7 algorithm, along with a justification for its utilization. Section 3 provides a comprehensive description of the implementation methodology employed in our study. The experimental results are presented in Section 4. Section 5 provides an analysis of the outcomes obtained in the preceding section, highlighting the significant discoveries derived from this study. Additionally, it outlines potential areas for further investigation that can be pursued in the immediate future. Section 6 provides a comprehensive summary of this study's findings and offers concluding remarks.

2. Related Studies

Considerable research has been dedicated to addressing the safety belt detection challenge, with scholars exploring solutions both inside and outside the vehicle using image-based methods [2]. External detection of safety belts has been proposed by various researchers [3,7,11–13] as an alternative to relying on traffic officers to monitor safety belt usage on highways, given the limited availability of personnel for such tasks. However, this approach still lacks the capability to enforce safety belt usage, as car occupants and drivers can choose to wear seat belts only in areas where external cameras are installed.

Most studies focus on detecting the windshield and then identifying the safety belt [11,12]. Moreover, the safety belt status detection in many of these studies is limited to detecting one person's safety belt, either that of a front passenger or the driver. For instance, Elihos et al. [7] proposed a system to detect the front seat passenger's safety belt, while Tianshu et al. [13], Guo et al. [3], and Hossein and Fathi [8] developed systems to detect the driver's safety belt. Unfortunately, these studies overlook the safety of other vehicle occupants and concentrate solely on saving one person.

The onboard detection of safety belts is seen as a promising solution to the aforementioned issues. Jaworek-Korjakowska, Kostuch, and Skruch view in-cabin detection of safety belts as a novel approach that can enhance the functionality of new-generation cars [14]. Nevertheless, this method comes with several challenges, such as dealing with constant illumination changes while the car is in motion, potential similarities between the safety belt and the passengers' clothing colour, blurriness in images due to vibration and bumpy roads, and occlusion caused by objects like hands, hair, and clothing [2,14]. Additionally, the lack of availability of a public dataset poses another challenge.

Due to the complexity of these challenges, the topic of safety belt detection remains unsolved, both for internal and external monitoring [14]. Despite this, numerous solutions have been proposed to address these issues. As stated by Kashevinik et al. [15] and Guo et al. [3], many proposed solutions for safety belt detection rely on edge detection algorithms, which can be unstable in varying environments. Chen et al. [12] also concur with this observation and note that some studies use Hugh transform (HT) algorithms in addition to edge detection. These methods fall under the category of traditional algorithms, where the features are handcrafted. Traditional object detection algorithms face challenges related to speed, accuracy, and environmental variations, as pointed out by Zou et al. [16]. Additionally, when processing large datasets, traditional algorithms tend to be inadequate, as highlighted by Luo, Lu, and Yue [17]. In contrast, deep-learning-based techniques offer a different approach, as depicted in Figure 3. These methods automate feature extraction, providing a more efficient and accurate solution for safety belt detection compared to traditional methods.

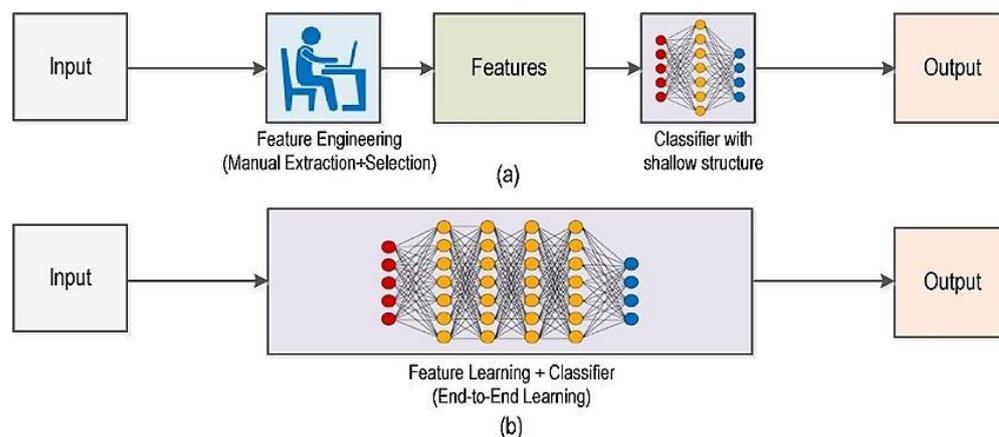


Figure 3. (a) Traditional algorithms and (b) deep learning algorithms [9].

Table 1 presents the results of various studies, and it is evident that deep-learning-based algorithms outperform traditional methods in terms of accuracy, supporting Zou et al.’s [16] submissions. Traditional algorithms fall short in accuracy compared to deep learning techniques. However, it is worth mentioning that Hu [2] claims to have achieved an accuracy of over 95% with his shape modelling algorithm, which is a notable achievement. Among the deep-learning-based algorithms, YOLOv5, as presented by Hosein and Fathi [8] in 2023, achieved the highest accuracy.

Table 1. Review summary of the literature for safety belt detection.

Author(s)	Method	Results
Bhupal Naik et al. [1]	Using SVM and CNN	CNN accuracy@91.4% SVM accuracy@87.17%
Hu [2]	Shape modelling	Accuracy@97.5% (seatbelt ON) Accuracy@100% (seatbelt OFF) Accuracy@98.11% (seatbelt under the arm)
Guo et al. [3]	Edge detection algorithms with HSV colour	Accuracy@81%
Elihos et al. [7]	SSD, CNN, and Finite Volume (FV) algorithms	Accuracy@91.9% Precision@94.5%
Hosseini and Fathi [8]	Using YOLOv5	Accuracy@99.7%. True positive rate@92% True negative rate@90% True rate@91%
Hosameldeen [11]	CNN with Alex Net approach	Average detection rate: 92.1% Accuracy@79.87% Recall@95.92% Specificity@90.32% Sensitivity@63.51%
Chen et al. [12]	CNN and SVM	Precision@63.58% F1 score@63.55%.
Jaworek-Korjakowska, Kostuch, and Skruch [14]	Adjusted CNN architecture and model interpretability	Segmentation F1 score@51% Synthetic F1 score@67%
Chun et al. [18]	Using CNN	Accuracy@89.1%
Jha et al. [19]	Segmentation and Synthetic images	
Kapdi et al. [20]	Using deep learning	

Our study utilizes the YOLOv7 algorithm for safety belt detection due to its numerous benefits. YOLO is an open-source single-stage object detector that is continuously being researched and improved. While YOLOv8 and YOLO-NAS are new variants, most researchers have opted for YOLOv7, leading us to select YOLOv7 for our study. According to Wang et al. [10], YOLOv7 is currently the most accurate real-time object detection algorithm, outperforming all known object detection methods. However, it is important to note that these claims were made prior to the release of YOLOv8 and YOLO-NAS.

3. Materials and Methods

The design method of the proposed study was structured into three main subsets: data collection, data preparation and training, and the model prediction stage, which is discussed in depth in Section 4. However, this section also discusses the history and workings of the YOLOv7 algorithm, which was selected as our model of choice. The last part of the section discusses the deployment process of the proposed study. Figure 4 depicts the flow diagram of the proposed study.

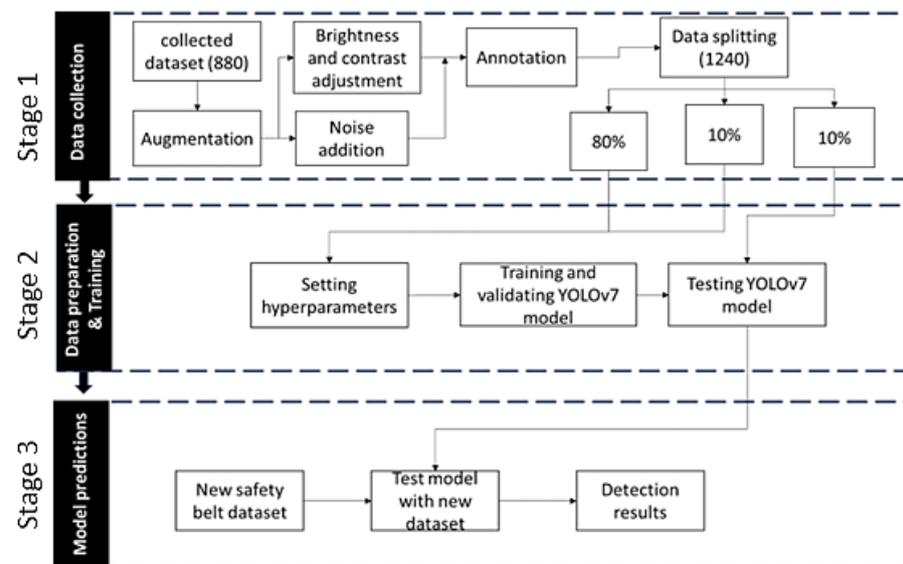


Figure 4. The flow diagram of the proposed study.

The sub-sections below elaborate on the steps presented within the stages of the flow diagram depicted in Figure 4.

3.1. Stage 1: Data Collection

The primary challenge encountered in this research is the lack of availability of a suitable dataset, which is also acknowledged by Chun et al. [18]. To address this issue, we decided to capture live video footage from within a moving car. The decision to record data in a moving vehicle was to obtain real-world data that encompass varying light intensities and blurred images caused by car vibrations. The data collection process was carried out in November and December 2022 (summer period) in Cape Town, South Africa. During this period, a diverse range of clothing types to simulate real-life scenarios was obtained. Some datasets were acquired through crowdsourcing, where we solicited video contributions from social media platforms. Consent was given by all individuals who donated their videos. Additionally, we obtained other datasets by downloading random photos from the internet to augment the diversity of the collected data. The video dataset was then converted to images using the VLC media player program. Table 2 gives figures of the collected dataset, and Figure 5 shows a representative sample of the dataset collected from video images. The sample dataset shows a few scenarios that this study was investigating. Amongst those scenarios, there are blurry images, clear images, high-light-intensity images, images with clothing similar to the safety belt, and images where the pillar loop was occluded. In all instances, the proposed study was expected to discriminate between the buckled and unbuckled status of a safety belt. The collected dataset was augmented using the noise addition and brightness contrast approaches. The benefits of these augmentation approaches were an increase in the number of datasets and the robustness of the model in changing environments inside the vehicle.

Table 2. Number of the collected datasets and their respective sources.

Dataset Source	Number of Datasets
Brightness and contrast adjustment augmentation	240
Noise augmentation	120
Original dataset	880
TOTAL	1240

**Figure 5.** Data collection sample.

The annotation process was carried out using the labelling tool, resulting in the generation of a label file in the “.txt” format. The dataset underwent annotation using the YOLO format, which entailed the application of rectangular bounding boxes. The annotation procedure encompassed the identification of the Region of Interest (ROI) as the pillar loop and the safety belt webbing. In instances where the pillar loop was not perceptible, the corresponding images were designated as unlatched. On the other hand, in cases where the pillar loop was observable, the images were designated as exhibiting buckling. The bounding rectangles were visually represented as annotations and identified in written form as follows: The first number in the line represented the category of the object, represented by an integer, while the subsequent four numbers indicated the coordinates of the bounding boxes. The initial two numerical values denoted the central point of the enclosing rectangle in relation to the X and Y axes, respectively. These values were then divided by the height and width of the image. The final two numerical values denoted the dimensions of the bounding box, with one representing the width and the other representing the height. These dimensions were then divided by the corresponding width and height of the image. The data table is organized such that each row represents the number of annotated objects.

During data splitting, the dataset was divided into three subsets: test, train, and validation datasets. The reason for the validation and testing of the dataset was to avoid

the risk of overfitting. The dataset was split with a ratio of 80% for training and 10% each for validation and testing. Table 2 below depicts the total number of datasets.

3.2. Stage 2: Preparation and Training

As stated by Poonkuntran et al. [21], data preprocessing involves removing HTML tags and URL tags, renaming the data, and resizing the images. Since a portion of the data was collected online, it contained HTML tags and URLs that needed to be eliminated before training the model. The training and performance evaluation platform for the network model was executed on a personal computer with the following configuration: The CPU of the system was an Intel (R) i5-10300H, capable of running at a maximum frequency of 2.5 GHz. The system also included an NVIDIA GeForce GTX 1650 GPU and 16G of RAM. The operating system running on the system was Windows 11 with a 64-bit architecture. For deep learning tasks, the system utilized Pytorch version 1.21 with CUDA version 12.2 for GPU acceleration. The programming language used for development was Python 3.9.16. The first step in the training process entailed the replication of the YOLOv7 library from the GitHub repository [10]. The hyperparameter configuration of the model was altered, as specified in Table 3.

Table 3. Hyperparameter configuration.

Parameter	Configuration
Image size	640 × 640
Batch size	4
Epochs	Initial setting was 300, and final setting, due to manual early stopping technique, was 30
Learning rate	0.01
Weight decay	0.0005
Optimizer	Adam

3.3. Stage 3: Predictions

Upon the completion of Stage 1 and Stage 2, an evaluation was conducted to assess the performance of the model in terms of its robustness and accuracy in predicting the status of the buckle on a safety belt. The graphical representation and comprehensive analysis of the model's outcomes and performance can be found in Section 4.1 of this scholarly article.

3.4. Model Selection

The YOLO algorithm, which was introduced in 2015, utilizes a methodology that treats object detection as a regression problem. This involves the division of bounding boxes and the corresponding class probabilities into separate spatial entities [22]. In essence, the processes of classification and bounding box location calculation occur concurrently. The algorithm initiates by dividing the image into a grid of dimensions $N \times N$. Subsequently, for each cell within the grid, it makes predictions regarding the locations of the bounding boxes, the confidence level of an object's presence, and the probability of its class. Consequently, the algorithm accomplishes the entire process of object detection and classification in a single traversal of the image. The YOLO project is currently undergoing active development, resulting in the release of newer and more advanced versions over the past few years. These versions have demonstrated improvements in the speed and performance of the algorithm [23]. These enhancements have been achieved through modifications to the algorithm's architecture or the incorporation of additional capabilities. Figure 6 illustrates the chronological development of the YOLO algorithm since its inception in 2015. Each subsequent iteration is assigned a distinct version identifier, and at present, the prevailing edition of YOLO is YOLO-NAS, constituting the ninth official release. The case study described in this research paper utilizes the YOLOv7 algorithm [10]. The rationale for this is that, although newer iterations of the YOLO algorithm have been devised, YOLOv7 has emerged as the most recent version favoured by researchers across a range of applications,

as evidenced in the existing body of literature. The YOLO algorithm demonstrates significant efficacy in the domain of automobile-related image tasks due to its exceptional inference speeds, which enable near-real-time image classification, and its ability to achieve high classification performance metrics. Indeed, it would not be an exaggeration to assert that, at present, this algorithm stands as the sole solution that provides a harmonious balance between speed and performance for tasks involving simultaneous object detection. Figure 6 depicts the historical evolution of the YOLO algorithm.

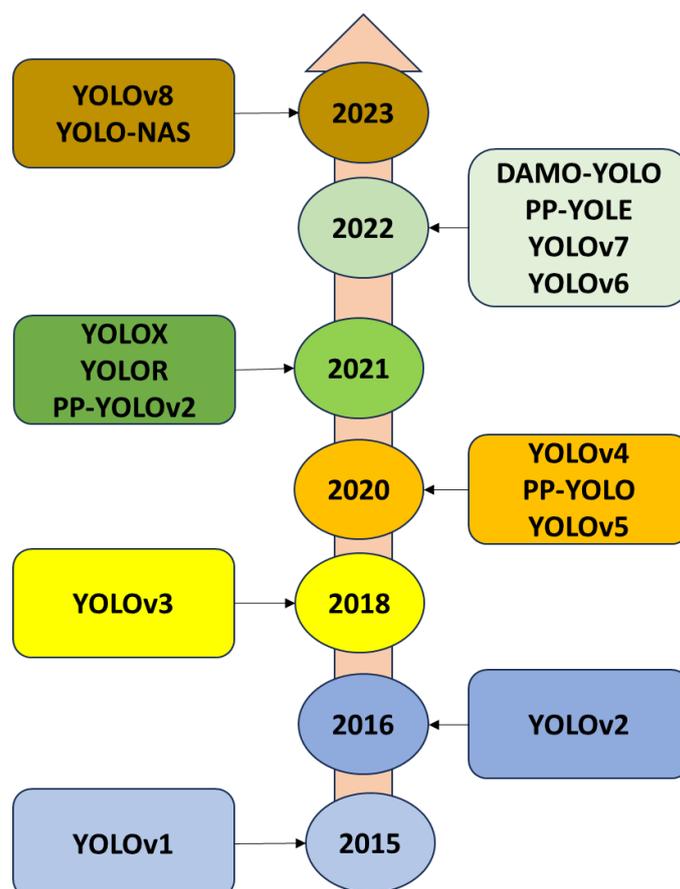


Figure 6. The evolution of the YOLO algorithm.

The selection of the YOLOv7 algorithm was based on its proven ability to achieve a balance between speed and accuracy. This aligns with our main goal of accurately detecting the real-time status of the safety belt. The YOLOv7 model is a highly adaptable and multi-functional model utilized in the field of computer vision. It exhibits the ability to execute a wide range of tasks, including object detection, classification, instance segmentation, and pose estimation [10]. YOLOv7 occupies a fundamental position within the YOLO series, as referenced in the scholarly literature [24]. The methodology integrates various techniques, such as the extended efficient long-range attention network (E-ELAN), model scaling through concatenation-based models [25], and convolution reparameterization [10]. These techniques enhance the model's resilience and effectiveness in addressing the safety belt detection task. The integration of various strategies in YOLOv7 achieves a favourable equilibrium between the efficiency of object detection and the accuracy of the results. The detection concept of YOLOv7 exhibits resemblances with YOLOv4 and YOLOv5, as observed in other network models belonging to the YOLO series. The architectural depiction of YOLOv7 can be located in the citation provided as Reference [10]. The YOLOv7 network is composed of four distinct modules, namely the input module, the backbone module, the head module, and the prediction module, as shown in Figure 7.

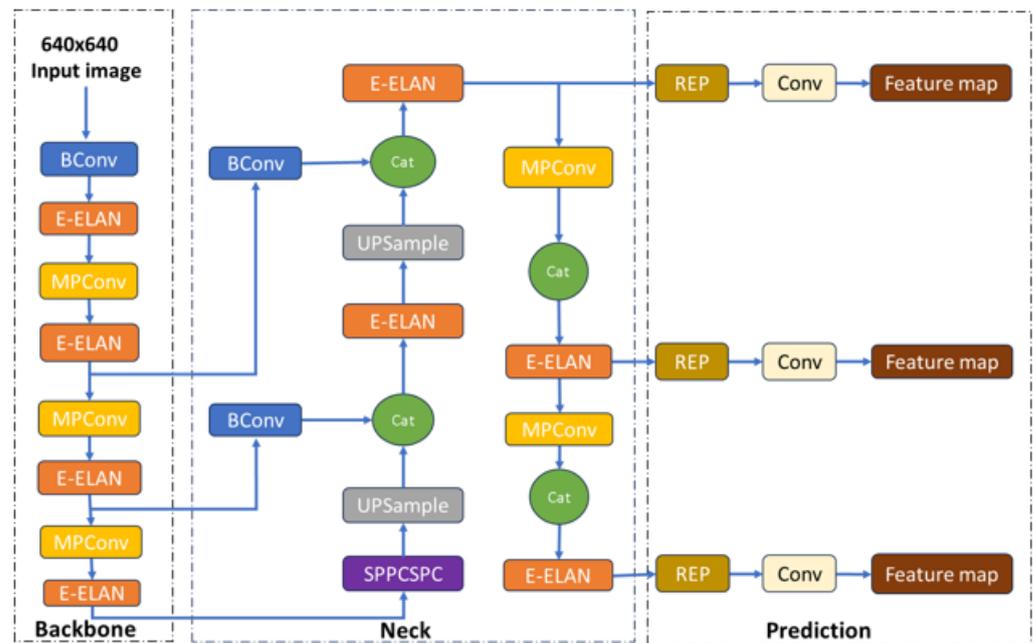


Figure 7. The original YOLOv7 architecture as presented by [10].

The backbone module comprises a series of Bconv [26] convolution layers, E-ELAN convolution layers, and max-pooling convolution (MPCConv) [26] layers. The input module of the YOLOv7 architecture is tasked with the responsibility of resizing the input image to a uniform pixel dimension in order to adhere to the input size requirements of the backbone network [27]. The backbone module comprises a series of Bconv [26] convolution layers, E-ELAN convolution layers, and max-pooling convolution (MPCConv) [26] layers. The BConv approach integrates multiple components, including a convolution layer, a batch normalization (BN) layer, and a LeakyReLU activation function [26], in order to effectively capture image features across different scales. This process enables the transmission of fundamental knowledge from lower tiers to higher tiers, thereby ensuring the efficient incorporation of characteristics across various tiers. The image channel numbers for the P3, P4, and P5 features from different scales’ outputs are adjusted by the prediction module using the rep-vgg block (REP) structure [27]. Following this, a 1×1 convolution is utilized to make predictions regarding confidence, class, and anchor frame.

3.5. Deployment

The purpose of the deployment was to enhance the implementation of inference during the safety belt detection experiments carried out within the vehicle. To achieve this objective, it was imperative to transform the weights of the trained YOLOv7 model into a more lightweight format that could be effectively processed on the Jetson Nano development board. The process consisted of three distinct steps, as illustrated in Figure 8.

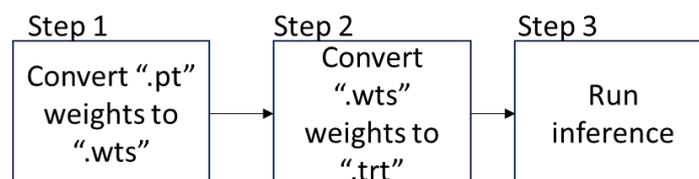


Figure 8. The steps used to deploy the YOLOv7 model to a Jetson Nano development board.

The YOLOv7 model was trained using the PyTorch platform, as stated in Stage 2 of the flow diagram. The weights for each training process were generated in the PyTorch format (.pt). The weights with the “.pt” extension are typically unsuitable for deployment on devices with limited memory capacity, such as the Jetson Nano. To facilitate compatibility

with the Jetson Nano platform, we performed a conversion process wherein the weights in the “.pt” format were transformed into TensorRT weights in the “.trt” format. This conversion was carried out with the objective of achieving lighter weights that could be effectively managed by the Jetson Nano platform. In order to convert weights to the Tensor RT format, it is necessary to first convert the “.pt” format to the “.wts” format. Subsequently, the “.wts” format needs to be converted to the “.trt” format.

4. Performance and Experimental Results

This section provides a comprehensive discussion of the performance metrics employed during the training of the model. The presentation of these metrics is supported by the inclusion of illustrations and tables. Following this, a series of experiments are conducted in order to substantiate our various hypothetical theories, which are then accompanied by visual representations.

4.1. Model Performance Metrics

The metrics frequently employed in the field of object detection include precision recall, intersection over union (IOU), average precision (AP), and mean average precision (mAP) values. The IOU metric is employed to quantify the extent of intersection between the bounding box predicted by the system and the ground truth bounding box in the original image. The calculation involves determining the intersection and concatenation ratios between the Detection Result and the Ground Truth.

$$\text{IOU} = \frac{\text{area of overlap } (B_p \cap B_{gt})}{\text{area of union } (B_p \cup B_{gt})} \quad (1)$$

The experiment involved the establishment of a threshold value of 0.5 for IOU. The Detection Result was classified as a True Positive (TP) when the IOU value, calculated between the Detection Result and the Ground Truth, exceeded the specified threshold value. This indicates the successful and accurate identification of targets. On the other hand, the Detection Result was classified as a False Positive (FP) if the IOU value was below the specified threshold, indicating an incorrect identification. The quantity of unobserved targets, known as False Negatives (FN), was computed as the count of Ground Truth occurrences that did not have a corresponding Detection Result. Precision is a metric that quantifies the ratio of correctly identified positive instances (True Positives) to the total number of instances identified as positive (True Positives + False Positives), represented as a percentage.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (2)$$

The recall metric represents the ratio of correctly identified target samples to all positive samples in the test set.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3)$$

The Precision–Recall (PR) curve is a graphical representation that showcases precision on the vertical axis and recall on the horizontal axis. This visualization effectively demonstrates the relationship between the classifier’s accuracy in correctly identifying positive instances and its capability to capture all positive instances. The average precision (AP) serves as a scalar measure that quantifies the area enclosed by the Precision–Recall (PR) curve. A higher AP value signifies the classifier’s enhanced performance.

$$\text{AP} = \int_0^1 \text{Precision}(\text{Recall}) d\text{Recall} \quad (4)$$

The mAP is a quantitative metric utilized to evaluate the accuracy of object detection models across all categories in a specific dataset. The mathematical representation of this phenomenon can be expressed as:

$$\text{mAP} = \frac{1}{N} \sum_{i=1}^N \text{AP}_i \quad (5)$$

where AP_i is the number of APs and N is the sum of the predicted classes. The F1 score is a metric that combines precision and recall in an optimal manner. It is the harmonic mean of precision and recall [28]. Mathematically, the F1 score can be expressed as:

$$\text{F1 - score} = 2 \times \left(\frac{\text{Precision} \times \text{Recall}}{\text{Recall} + \text{Precision}} \right) \quad (6)$$

Throughout the training and testing phases, we evaluated the efficacy of our model by employing a range of metrics, encompassing precision, recall, confusion matrix, F1 score, recall, precision, and mAP. The aforementioned measurements were utilized as quantitative indicators for the purpose of assessing the efficacy of the model. The confusion matrix is a widely employed tabular representation that effectively summarizes the predictions generated by a trained model. The classification framework comprises four key components: TP, FP, TN, and FN.

Experimental Results and Analysis of the Collected Dataset

The detection performance of the proposed YOLOv7 object detection was assessed through the experimental evaluation of the collected dataset. The results depicted in Figures 9 and 10 indicate a significant level of efficiency, particularly for the “buckled” class, which exhibits an AP of 99.7%. The “unbuckled” class also exhibited a high level of efficiency, as evidenced by a performance rate of 99.5%. The mAP of the model was determined to be 99.6%. Figures 9 and 10 present the graphical outcomes pertaining to precision, recall, mAP, AP, and F1 score for both categories.

The results of the proposed study were analysed using a confusion matrix. In the matrix, the predicted class is represented by each column, while the true proportions of each class in the dataset are represented by each row, as shown in Figure 11. The analysis of Figure 11’s findings illustrates that the model exhibits the utmost level of accuracy in predicting both the “buckled” and “unbuckled” classes, with a 100% prediction rate.

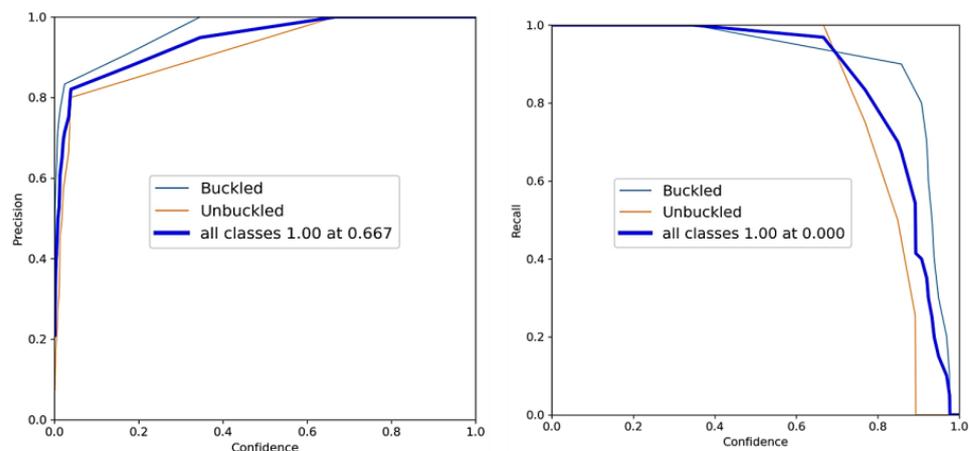


Figure 9. Depiction of the precision curve and the recall curve.

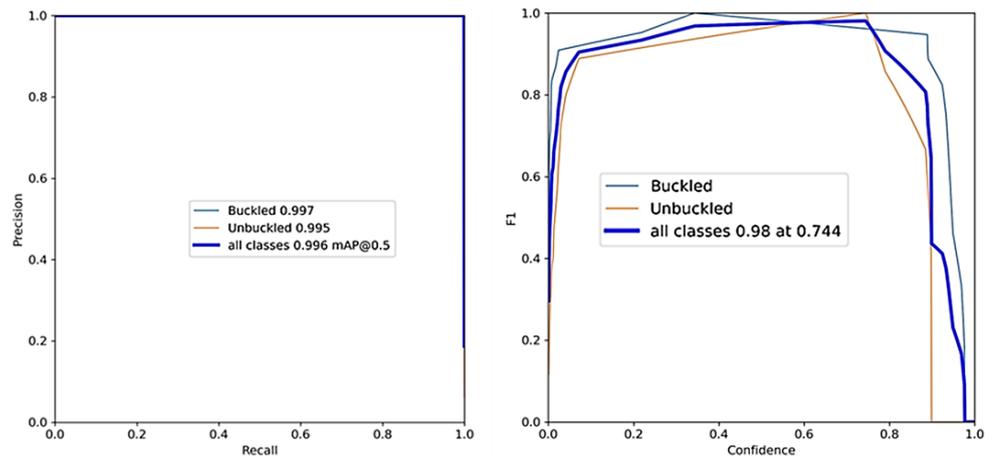


Figure 10. Precision–recall curve (PR) and F1 score graph.

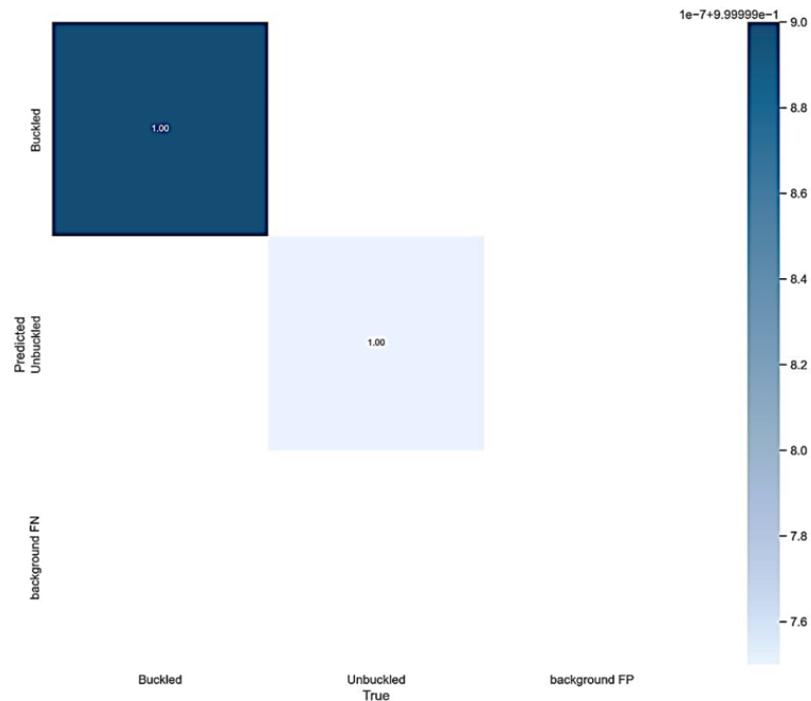


Figure 11. The confusion matrix of the proposed study.

Moreover, this study provides a visual depiction of the fluctuations in the loss values, specifically the Box loss, Objectness loss, and Classification loss. The YOLOv7 model incorporates the Generalized Intersection over Union (GIOU) loss as its chosen loss function for bounding boxes. The Box loss is calculated as the average of the GIOU loss function [29]. A lower value of the loss function signifies a higher level of accuracy. The Objectness loss refers to the mean value of the target detection loss, where a lower value indicates a higher level of accuracy [29]. The Classification loss refers to the average value of the Classification loss metric, where a lower value signifies a higher level of accuracy, as shown in Figure 12. As illustrated in Figure 12, the loss values exhibited a consistent decline, ultimately reaching a stable state after 30 epochs. To prevent overfitting, a manual early stopping technique was randomly employed at the 30th epoch with satisfactory model performance metrics.

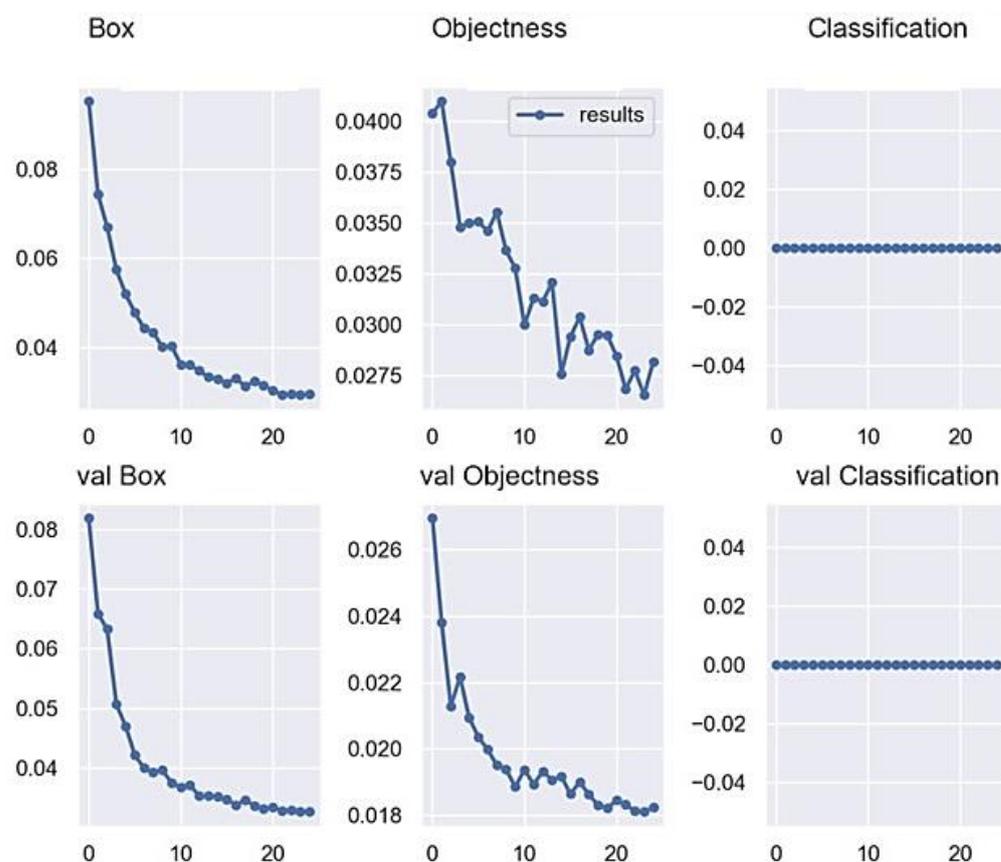


Figure 12. The measured losses of the proposed study.

To further demonstrate the superiority of the proposed study, a comparison of results was carried out between the reviewed related studies, as shown in Table 4. As demonstrated by the results in Table 4, the proposed model output outperforms almost all the reviewed object detection models. However, Hossein and Fathi [8] claim to have obtained an accuracy of 99.7%, which is 0.1 higher than our 99.6% mAP rate. However, the proposed model demonstrates the practical advantages of in-cabin detection of the safety belt.

Table 4. Result comparison with related studies.

Author(s)	Method	Results
Guo et al. [3]	Edge detection and HSV	Accuracy@81%
Elihos et al. [7]	SSD, CNN, and FV algorithms	Accuracy@91.9% Precision@94.5%
Hossein and Fathi [8]	YOLOv5	Accuracy@99.7%
Hosameldeen [11]	CNN with Alex Net	True positive rate@92% True Negative rate@90% True rate@91%
Chen et al. [12]	CNN and SVM	Average Precision@92.1%
Chun et al. [18]	CNN	Precision@63.58% Sensitivity@ 63.51% F1 score@63.55%
Ours	YOLOv7	mAP @99.6% F1 score@ 98%

4.2. Analysis of the Trained YOLO Model Based on Real-Time Experiments

This section encompasses a series of test experiments aimed at substantiating diverse hypothetical scenarios. Each of these experiments is accompanied by a comprehensive explanation of the test procedures employed, followed by a detailed analysis of the test results. The setup of all the experiments that were conducted onboard is shown in Figure 13.

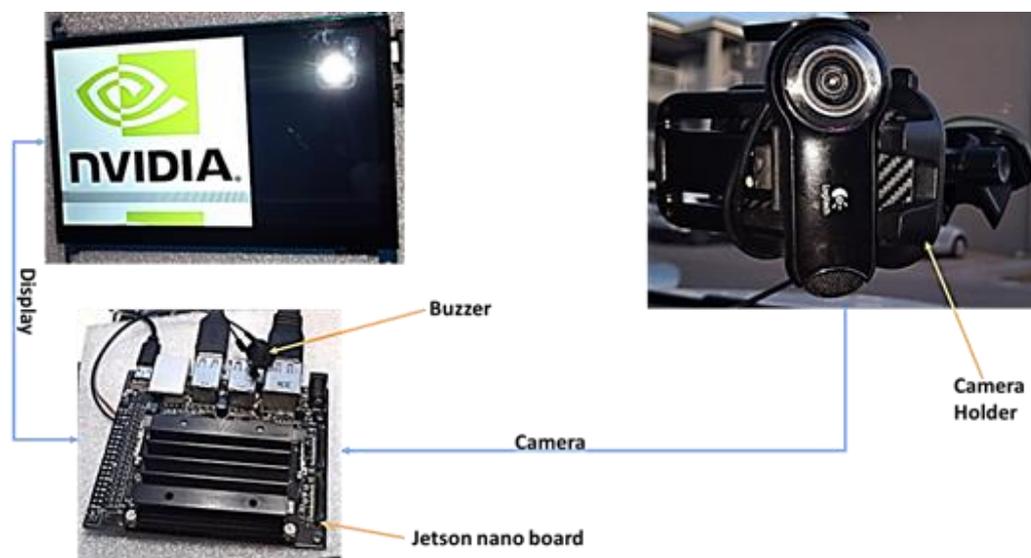


Figure 13. The setup of all the experiments that were conducted inside the vehicle.

The setup consisted of an Nvidia Jetson Nano board, a camera, a display, a buzzer, and a camera holder. The camera was utilized as an acquisition device for inferring, and the Jetson Nano was employed for processing the incoming frames to detect the status of a safety belt. The display was utilized to verify the status of the detection and the configuration of the camera. The buzzer was connected to the Jetson Nano and would only be triggered when the status of the safety belt was detected as unbuckled. Lastly, the camera holder was mounted onto the windshield and utilized to securely hold the camera in place, preventing any potential falls while the car was in motion.

4.2.1. Experiment 1: Concept Testing

The underlying premise of our study necessitates the installation of a camera on the windshield with appropriate adjustments to ensure the visibility of the safety belt pillar loop. The detection system is designed to categorize the safety belt status as “unbuckled” if the pillar loop remains invisible, and conversely as “buckled” if the person is securely fastened and the pillar loop is visible. Additionally, in cases where an individual secures the safety belt behind their back but the pillar loop is visible, the system must correctly identify the status as “unbuckled”. Consequently, a buzzer is activated whenever the safety belt status is detected as unbuckled. The choice of using the pillar loop is motivated by unforeseen scenarios where occupants could be wearing clothing with stripes that are similar to a safety belt or when a person is carrying a bag that has webbing that looks like a safety belt. In both of these cases, if a pillar loop is not used as the determining factor, the model would detect the status of a safety belt as buckled even though people might not necessarily be buckled.

The pillar loop approach was validated in four cases, as shown in Figure 14:

1. When a vehicle occupant buckled behind their back and the pillar loop of the safety belt was visible.
2. When two vehicle occupants wore their seat belts and their pillar loops were visible.
3. When two vehicle occupants were unbuckled and their pillar loops were visible.

4. A detection status when one vehicle occupant was buckled but had an invisible pillar, and a detection status when the other occupant was buckled with a visible pillar loop.

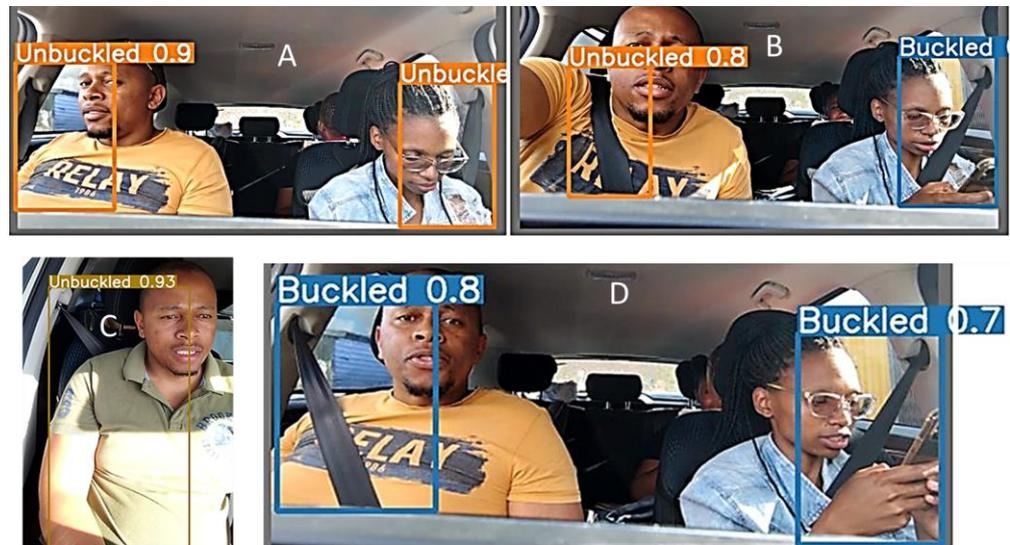


Figure 14. (A–D) depict the safety belt detection status when the model was tested in conditions where the safety pillar loop was visible and where it was invisible.

4.2.2. Experiment 2: Occlusion

This experiment established the ability of our model to detect a safety belt buckling status when the pillar loop was partially occluded. The camera was adjusted as explained in experiment 1. The experiment was conducted in a stationary car. The experiment tested the ability of the model to detect the safety belt status in occlusion conditions. Figure 15 depicts various scenes (A, B, and C) where the model showed superiority in detecting the safety belt status in different occlusion conditions.

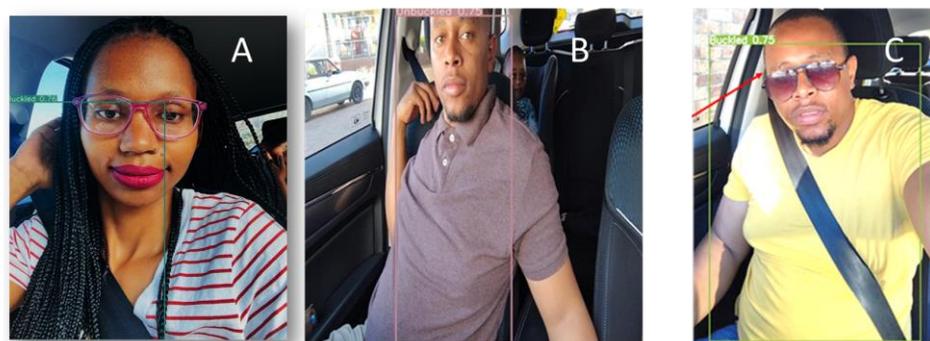


Figure 15. Detecting a partially occluded car safety belt. (A) shows the detection status of a safety belt when the hand and hair partially obscured the webbing and the pillar loop was only 50% visible. (B) shows the safety belt detection status with the hand blocking the webbing while buckled. (C) shows safety belt detection with the pillar loop partially visible.

4.2.3. Experiment 3: Clothing

Experiment 3 aimed to ascertain the detection status of safety belts when the individuals occupying the vehicle are attired in garments that possess a colour resembling that of the safety belt webbing. The camera underwent adjustment procedures as outlined in Experiment 1. The study was conducted within the confines of a vehicle in motion. The objective of the experiment was to evaluate the efficacy of the model in detecting the status of the safety belt under specified conditions. Figure 16 illustrates the observations of the experiment.

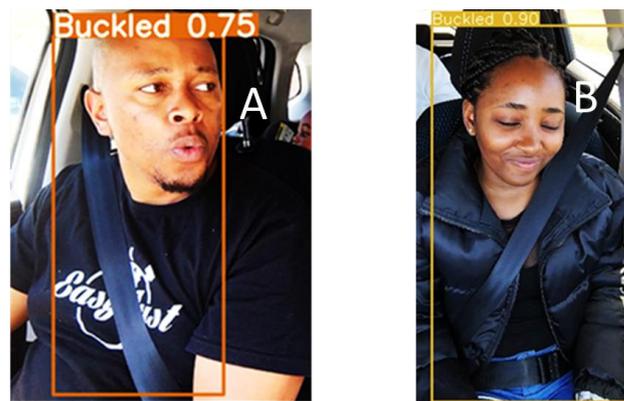


Figure 16. The safety belt detection status of (A,B) was represented when occupants were wearing clothing that was of a similar colour to the webbing.

4.2.4. Experiment 4: Lighting

The objective of this experiment was to assess the resilience of the proposed model under varying lighting conditions. The model underwent testing under three different lighting conditions: clear lighting (A), dark lighting (B), and high-light-intensity conditions. In order to simulate different lighting conditions, the vehicle was operated in shaded areas to represent low-lighting conditions, in areas with direct sunlight to simulate high light intensity, and in areas with moderate sunlight to represent good lighting conditions for approaching vehicle occupants. No detections were observed under conditions of high light intensity, which can be attributed to the distortion of safety belt features caused by excessive sunlight exposure on the camera. Figure 17 illustrates the detection of safety belts under different lighting conditions, specifically in clear lighting (A) and dark lighting (B).

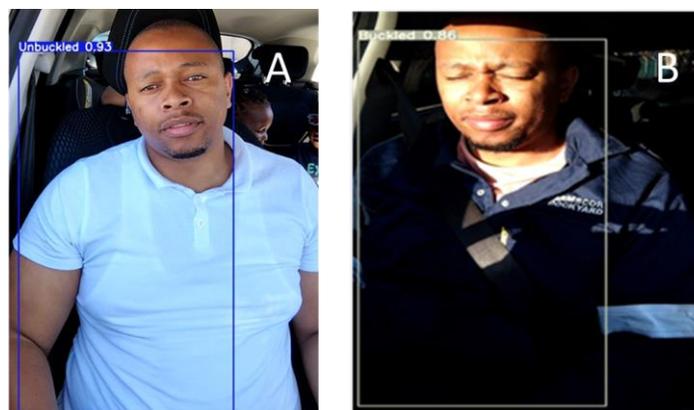


Figure 17. This illustrates the detection status of safety belts in both well-lit conditions (A) and low-light conditions (B).

4.2.5. Experiment 5: Blurry Scenes

The present study successfully demonstrated the ability to detect the presence of a safety belt in images with low clarity. This phenomenon was observed in situations where the camera experienced rapid movement due to factors such as a rough road surface or vibrations. The camera was affixed in accordance with the methodology outlined in Experiment 1, and its stability was ensured through the meticulous adjustment of the holder screws. Following that, we proceeded to traverse a speed bump at a velocity of 40 km per hour, resulting in an observation of significantly distorted visual perception, as illustrated in Figure 18.



Figure 18. The visual distortion observed in the scene is a result of the vehicular vibrations induced by traversing a speed hump at a velocity of 40 km per hour.

4.3. Analysing Statistical Insights and Robustness of Safety Belt Buckle Detection across Diverse Experimental Scenarios

The results we present here are based on a total of 15 tests carried out for each of the five distinct experimental scenarios. These scenarios were designed to cover a wide range of situations that could affect the detection performance of the safety belt buckle. The outcomes of these tests are depicted in Figure 19, which illustrates the object detection model's accuracy and performance across different conditions. This comprehensive approach ensures that our findings are grounded in a robust and representative dataset, allowing us to draw meaningful insights about the model's effectiveness in real-world scenarios.



Figure 19. Comparative Analysis of Safety Belt Buckle Detection Performance.

The line plot visualization of experimental outcomes shown in Figure 19 provides a comprehensive overview of the YOLOv7-based object detection model's accuracy in identifying safety belt buckles. The analysis highlights the model's proficiency in distinguishing between buckled and unbuckled belts across diverse scenarios while also pinpointing opportunities for refinement. These insights contribute to the advancement of intelligent safety systems, encouraging further exploration and enhancements in safety belt detection technology for safety applications.

5. Discussion

The primary objective of this study was to assess the efficacy of the YOLOv7 algorithm in detecting the presence of a safety belt. According to Reference [5], the failure to utilize safety belts leads to a significant number of fatalities, with a daily death toll of 3285 and an

annual toll of 1.35 million. The investigation of safety belt buckling is a significant area of research due to the potential for grave injuries and fatalities resulting from the act of unbuckling [1–4]. The majority of research studies conducted thus far have primarily focused on conventional algorithms, which have been found to exhibit instability and inaccuracy when confronted with dynamic environments [15]. Previous approaches have involved the utilization of cameras installed on highways [3,7,11–13] as a means of identifying the status of safety belts. This approach aimed to address the challenge of the limited number of traffic officers available for conducting manual inspection tasks. Nevertheless, this approach anticipates the potential for vehicle occupants to easily fasten their safety belts in regions where the cameras are installed. Previous studies [3,8] have primarily concentrated on the identification of the driver's utilization of a safety belt, neglecting the safety of other individuals present in the vehicle. Figures 15 and 16 illustrate the efficacy of our model in detecting multiple vehicle occupants, thereby enhancing the safety of a larger number of individuals within the vehicle.

Based on the findings of previous studies [3,12,15], the majority of research in this field has focused on exploring seat belt buckling solutions through the utilization of conventional algorithms. According to Zou et al. [16], conventional object detection algorithms face difficulties in terms of their speed, accuracy, and adaptability to different environmental conditions. The model under consideration was derived from YOLOv7, which belongs to the same category as other deep-learning-based approaches. Consequently, the study conducted demonstrated a notable level of precision, as evidenced by a mean average precision (mAP) of 99.6%, an F1 score of 98%, and a confidence level of 100% for both the buckled and unbuckled categories. The model exhibited satisfactory performance during the evaluations conducted inside the vehicle on occluded pillar loop scenes, the instances where occupants were attired in clothing resembling the webbing of the safety belt, the scenes with reduced clarity, and the scenarios involving diverse lighting conditions. Experiments 2 through 5 illustrate the experimental outcomes obtained.

Nevertheless, the study that was carried out had certain limitations. During the conducted experiments, it was observed that an abundance of sunlight results in the distortion of specific safety belt features, consequently leading to the unsuccessful detection of the safety belt status. Given that the proposed studies rely on the safety belt pillar loop as the primary determinant of the safety belt status, the utilization of a single camera positioned at a fixed angle may result in inaccurate detections. Consequently, the model may erroneously identify occupants as unbuckled, despite the possibility that they are indeed properly secured. The model's training did not include cases involving incorrect buckling of occupants except for when people were buckled behind their backs, leading to a noticeable decrease in confidence when detecting the safety belt status of individuals who are improperly buckled. The model was not evaluated on nocturnal scenes as a result of the absence of a camera capable of supporting night vision capabilities.

Safety belt detection using computer vision has witnessed significant advancements in recent years; however, several gaps and potential avenues for future research remain to be addressed. One of the primary challenges pertains to the robustness of detection algorithms across diverse environmental conditions. While existing models excel in controlled settings, the ability to accurately identify seat belts in varying lighting conditions, camera angles, seat designs, and occupant body sizes warrants further investigation. Real-time performance is another critical consideration, particularly in dynamic scenarios. The imperative to balance accuracy and real-time processing necessitates the development of algorithms optimized for swift and accurate detection. The quality and diversity of training data emerge as a central concern; comprehensive datasets that encompass a wide array of real-world scenarios are indispensable for enhancing model generalization. The intricate task of detecting safety belts amidst multiple occupants or occlusions demands innovative solutions. Future research directions encompass strategies to mitigate privacy concerns inherent in capturing passenger images while ensuring effective seat belt detection. Furthermore, the adaptation of detection algorithms to unconventional seating arrangements, as observed

in certain vehicle types, requires specialized attention. Integrating safety belt detection seamlessly into user interfaces, such as infotainment systems or driver assistance displays, emerges as a viable strategy to enhance passenger awareness and compliance. Deep learning, although efficacious, presents challenges in terms of interpretability, particularly in safety-critical applications. Exploring techniques to render deep learning models more transparent and interpretable is thus a pertinent research trajectory. Moreover, optimizing detection algorithms for resource-constrained hardware and edge computing installations is crucial for widespread deployment within vehicles. The longevity of system performance and accuracy necessitates continuous monitoring, calibration, and maintenance strategies. Lastly, the alignment of safety belt detection technology with industry standards and regulatory frameworks underscores the importance of collaborative efforts between researchers, industry stakeholders, and regulatory bodies to establish benchmarks and guidelines for effective implementation. As the field of computer vision evolves, ongoing research endeavours are poised to bridge these gaps and refine the efficacy of safety belt detection systems.

6. Conclusions Remarks

The present study investigated the identification of the safety belt status of a vehicle using the YOLOv7 algorithm. The model was constructed utilizing a dataset of 1240 images sourced from Cape Town as well as crowdsourcing and random online image acquisition methods. The dataset was divided into training, testing, and validation sets with a ratio of 80:10:10 using a Python script to automate the partitioning procedure. During the process of training the model, the findings of the present study exhibited a higher level of effectiveness in comparison to other relevant studies that were reviewed. The study under consideration successfully attained a mean average precision (mAP) of 99.6% at a threshold of 0.50. Additionally, it achieved an F1 score of 98% and precision rates of 99.7% and 99.5% for the buckled and unbuckled classes, respectively, within the model. Following a successful evaluation, the model was effectively deployed onto the Jetson Nano development board for the purpose of conducting onboard experiments. The experiments conducted revealed that the model operated in accordance with the proposed design concept, which involved utilizing a pillar loop as the key factor for distinguishing between the buckled and unbuckled states of a safety belt. Additional observations were made regarding the model's ability to detect various challenging scenarios. These included the detection of blurry scenes, partially occluded environments, instances where occupants were seated with their seatbelts fastened behind their backs, and scenes with low-lighting conditions. Nevertheless, it was observed that the model exhibited limitations in detecting instances of missing safety belt features under conditions of high light intensity.

In contrast to the current safety belt monitoring system, which can be easily deceived by methods such as fastening the safety belt behind the back or utilizing a safety belt stopper to mimic the latch plate of a safety belt, our proposed model employs a camera as the sensing mechanism. This model is capable of accurately distinguishing between the states of being buckled and unbuckled, even in situations where vehicle occupants intentionally fasten their belts behind their backs. The study under consideration exhibited the greatest level of precision in comparison to the most recent studies put forth in the existing body of research. Nevertheless, there remains potential for further enhancement. In order to advance the field, it is recommended that future studies place greater emphasis on improving the ability of models to accurately detect the presence of safety belts in situations characterized by high levels of illumination. One possible approach to resolving the challenges related to lighting conditions is to explore recent iterations of the YOLO algorithms, which have demonstrated superior levels of precision and resilience in comparison to alternative object detection algorithms employed in comparable research endeavours. These algorithms have the potential to be integrated with advanced camera technologies that possess the capability to adjust and adapt to different lighting conditions.

Author Contributions: Conceptualization, M.S.; methodology, L.N.; software, L.N.; validation, M.S.; formal analysis, L.N.; investigation, L.N.; resources, E.D.M.; data curation, L.N.; writing—original draft preparation, L.N.; writing—review and editing, L.N.; visualization, L.N.; supervision, M.S. and E.D.M.; project administration, E.D.M.; funding acquisition, M.S. and E.D.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding except for the publication funding provided by the Central University of Technology.

Data Availability Statement: The data are available on request when a reasonable amount of time is given.

Acknowledgments: Wholeheartedly, we would like to thank the Central University of Technology (CUT), Department of Electrical, Electronic, and Computer Engineering, for affording us the resources and the publication funding for the conducted study. We would also like to thank all the individuals who participated during the data collection stage of the project.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Naik, D.B.; Lakshmi, G.S.; Sajja, V.R.; Venkatesulu, D.; Rao, J.N. Driver's seat belt detection using CNN. *Turk. J. Comput. Math. Educ. TURCOMAT* **2021**, *12*, 776–785.
2. Hu, F. Robust Seatbelt Detection and Usage Recognition for Driver Monitoring Systems. *arXiv* **2022**, arXiv:2203.0081.
3. Guo, H.; Lin, H.; Zhang, S.; Li, S. Image-based seat belt detection. In Proceedings of the 2011 IEEE International Conference on Vehicular Electronics and Safety (ICVES 2011), Beijing, China, 10–12 July 2011; IEEE: Piscataway, NJ, USA, 2011; pp. 161–164. [[CrossRef](#)]
4. World Health Organization. *Save Lives: A Road Safety Technical Package*; World Health Organization: Geneva, Switzerland, 2017.
5. Kohli, P.; Chadha, A. Enabling Pedestrian Safety Using Computer Vision Techniques: A Case Study of the 2018 Uber Inc. Self-driving Car Crash. In *Advances in Information and Communication*; Arai, K., Bhatia, R., Eds.; Springer International Publishing: Cham, Switzerland, 2020; pp. 261–279.
6. Leland, J.; Stanfill, E.; Cherian, J.; Hammond, T. Recognizing Seatbelt-Fastening Behavior with Wearable Technology and Machine Learning. In Proceedings of the Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems. Presented at the CHI '21: CHI Conference on Human Factors in Computing Systems, Yokohama, Japan, 8–13 May 2021; ACM: New York, NY, USA, 2021; pp. 1–6. [[CrossRef](#)]
7. Elihos, A.; Alkan, B.; Balci, B.; Artan, Y. Comparison of Image Classification and Object Detection for Passenger Seat Belt Violation Detection Using NIR & RGB Surveillance Camera Images. In Proceedings of the 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), Auckland, New Zealand, 27–30 November 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–6. [[CrossRef](#)]
8. Hosseini, S.; Fathi, A. Automatic detection of vehicle occupancy and driver's seat belt status using deep learning. *SIVIP* **2023**, *17*, 491–499. [[CrossRef](#)]
9. Daoud, E.; Khalil, N.; Gaedke, M. Implementation of a One-Stage Object Detection Solution to Detect Counterfeit Products Marked with A Quality Mark. *IADIS Int. J. Comput. Sci. Inf. Syst.* **2022**, *17*, 37–49.
10. Wang, C.-Y.; Bochkovskiy, A.; Liao, H.-Y.M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *arXiv* **2022**, arXiv:2207.02696.
11. Hosameldeen, O. Deep learning-based car seatbelt classifier resilient to weather conditions. *IJET* **2020**, *9*, 229. [[CrossRef](#)]
12. Chen, Y.; Tao, G.; Ren, H.; Lin, X.; Zhang, L. Accurate seat belt detection in road surveillance images based on CNN and SVM. *Neurocomputing* **2018**, *274*, 80–87. [[CrossRef](#)]
13. Tianshu, W.; Zhijia, Z.; Zhanning, L.; Yunpeng, L.; Shixian, W. Detection and Implementation of Driver's Seatbelt Based on FPGA. *J. Phys. Conf. Ser.* **2019**, *1229*, 012075. [[CrossRef](#)]
14. Jaworek-Korjakowska, J.; Kostuch, A.; Skruch, P. SafeSO: Interpretable and Explainable Deep Learning Approach for Seat Occupancy Classification in Vehicle Interior. In Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Nashville, TN, USA, 19–25 June 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 103–112. [[CrossRef](#)]
15. Kashevnik, A.; Ali, A.; Lashkov, I.; Shilov, N. Seat Belt Fastness Detection Based on Image Analysis from Vehicle In-Abin Camera. In Proceedings of the 2020 26th Conference of Open Innovations Association (FRUCT), Yaroslavl, Russia, 23–25 April 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 143–150. [[CrossRef](#)]
16. Zou, Z.; Chen, K.; Shi, Z.; Guo, Y.; Ye, J. Object Detection in 20 Years: A Survey. *Proc. IEEE* **2023**, *111*, 257–276. [[CrossRef](#)]
17. Luo, J.; Lu, J.; Yue, G. Seatbelt detection in road surveillance images based on improved dense residual network with two-level attention mechanism. *J. Electron. Imag.* **2021**, *30*, 1–14. [[CrossRef](#)]

18. Chun, S.; Ghalehjegh, N.H.; Choi, J.; Schwarz, C.; Gaspar, J.; McGehee, D.; Baek, S. NADS-Net: A Nimble Architecture for Driver and Seat Belt Detection via Convolutional Neural Networks. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), Seoul, Republic of Korea, 27–28 October 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 2413–2421. [[CrossRef](#)]
19. Jha, S.; Brooks, I.; Ray, S.J. Seatbelt Segmentation Using Synthetic Images. Available online: https://ecs.utdallas.edu/research/researchlabs/msp-lab/publications/Jha_2023_2.pdf (accessed on 16 March 2023).
20. Kapdi, R.A.; Khanpara, P.; Modi, R.; Gupta, M. Image-based Seat Belt Fastness Detection using Deep Learning. *SCPE* **2022**, *23*, 441–455. [[CrossRef](#)]
21. Poonkuntran, S.; Balusamy, B.; Kumar Dhanraj, R. *Object Detection with Deep Learning Models: Principles and Applications*, 1st ed.; Chapman and Hall/CRC: Boca Raton, FL, USA, 2022.
22. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788. [[CrossRef](#)]
23. Ajala, J. *Object Detection and Recognition Using YOLO: Detect and Recognize URL(s) in an Image Scene*; St. Cloud State University: St. Cloud, MN, USA, 2021.
24. Lai, Y.; Ma, R.; Chen, Y.; Wan, T.; Jiao, R.; He, H. A Pineapple Target Detection Method in a Field Environment Based on Improved YOLOv7. *Appl. Sci.* **2023**, *13*, 2691. [[CrossRef](#)]
25. Wang, C.Y.; Bochkovskiy, A.; Liao, H. Scaled-YOLOv4: Scaling cross stage partial network. In Proceedings of the IEEE/CVPR, Nashville, TN, USA, 20–25 June 2021; IEEE: Piscataway, NJ, USA, 2021. [[CrossRef](#)]
26. Jiang, T.T.; Cheng, J.Y. Target Recognition Based on CNN with LeakyReLU and PReLU Activation Functions. In Proceedings of the 2019 IEEE Conference on Sensing, Diagnostics, Prognostics, and Control (SDPC), Beijing, China, 15–17 August 2019; pp. 718–722.
27. Ding, X.H.; Zhang, X.Y.; Man, N.N.; Han, J.G.; Ding, G.G.; Sun, J. RepVGG: Making VGG-style ConvNets great again. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 13728–13737.
28. Sibiya, M.; Sumbwanyame, M. *Classification and Severity Prediction of Maize Leaf Diseases Using Deep Learning CNN Approaches*; University of South Africa: Pretoria, South Africa, 2021.
29. Liu, K.; Sun, Q.; Sun, D.; Peng, L.; Yang, M.; Wang, N. Underwater Target Detection Based on Improved YOLOv7. *JMSE* **2023**, *11*, 677. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.