Grenze International Journal of Engineering and Technology, June Issue

Vulnerable Road user Detection using YOLOv8 for Effective Collision Avoidance

Manish Mohandas Narkhede[1] and Nilkanth Bhikaji Chopade[2]
[1-2]Research Centre, E&TC Department, Pimpri Chinchwad College of Engineering, Pune, India
Email: manishmn1987@gmail.com, nbchopade@gmail.com

***Abstract*— A growing number of vulnerable road users (VRUs), including cyclists, motorcyclists, and pedestrians, are involved in traffic accidents; therefore, reliable and effective detection technologies are needed to improve road safety. This paper proposes a computer vision approach using You Look Only Once version 8 (YOLOv8) for VRU detection. We adopted the method to detect VRUs by training them on a large dataset of annotated images containing road users with diverse mobility and evaluated it on a custom dataset. We achieved an average precision of 0.85 and an average recall of 0.78 over a significantly lower number of training epochs. The suggested approach can be implemented in real-time applications such as enhanced driver assistance systems, driverless cars, and surveillance systems to enhance VRU identification. This research contributes to the ongoing efforts to create safer road environments by leveraging the capabilities of modern computer vision and deep learning technologies. Integrating YOLOv8 and an onboard vision-based approach holds significant promise for reducing accidents involving vulnerable road users and fostering safer coexistence between vehicles and VRUs on today's busy roadways.**

***Index Terms*— Computer Vision, Deep Learning, VRU Detection, ADAS, Convolutional Neural Networks, Road Safety, Autonomous Vehicles.**

I. INTRODUCTION

Vulnerable road users lack external protection and face an increased risk of injury or fatality in traffic collisions [1]. The main VRU groups include pedestrians, cyclists, and motorcyclists. These groups account for around half of all global traffic fatalities despite making up a small percentage of road users [2].
Global Status Report on Road Safety 2018 by the World Health Organization highlights a stark reality: VRUs account for 54% of all global road traffic deaths [3]. Data from the European Commission reveals a worrying trend in the European Union, where pedestrians and cyclists represented 29% and 8% of all road fatalities in 2017, respectively [4]. In the United States, the National Highway Traffic Safety Administration (NHTSA) reported that urban areas witnessed 76% of pedestrian fatalities in 2018, underscoring the amplified risks in densely populated regions [5].
According to recent data, Two-Wheelers constitute India's most extensive road accident deaths, with a staggering 45.1% share. This translates to 69,385 fatalities, highlighting the acute vulnerability of two-wheeler users in traffic. Pedestrians are the next significant group affected, accounting for 18.9% of road accident deaths. This statistic is a stark reminder of pedestrians' dangers in urban and rural traffic settings. The analysis of vehicles involved in pedestrian fatalities showed that cars, taxis, vans, and light motor vehicles accounted for 19.9% of pedestrian deaths, followed by Two-Wheelers at 18.0% and Auto-Rickshaws at 16.9%. The victims on two-wheelers who

Grenze ID: 01.GIJET.10.2.110

were killed in road accidents made up around 45.2% of the total road accident deaths in 2021 [6].
The much higher degree of vulnerability for these groups stems from various factors. Since VRUs lack an external shielding structure and restraint systems, nearly all physical force in a collision is absorbed directly by the human body. This contrasts with occupants of motor vehicles that provide crash protection [7]. Further, the greater mass and velocity of motor vehicles in collisions causes more severe trauma. Difficulty seeing VRUs at night and in bad weather also heightens crash risks [8]. Improving safety outcomes for VRUs has significant implications for reducing overall traffic deaths and advancing road safety [9].
This research paper evaluates the recently introduced YOLOv8 algorithm's performance in accurately detecting and classifying vulnerable road users (VRUs) for advanced driver assistance systems (ADAS). The study addresses the challenges of suboptimal performance and real-time accuracy constraints existing systems face when detecting VRUs compared to larger vehicles.

II. Literature Review

A. Vulnerable road user detection

Accurate detection and tracking of VRUs are crucial for developing ADAS systems and autonomous vehicles [10]. Various methods have been proposed to address this issue, from traditional computer vision techniques to modern deep learning-based approaches [11].
Traditional computer vision methods for VRU detection often involve handcrafted features and classifiers. These methods rely on extracting features such as colour, texture, and shape from the input images and using classifiers like support vector machines or random forests to classify VRUs. However, these methods need help with complex scenarios, occlusions, and variations in lighting conditions [12].

B. Object Detection Algorithms

Object detection algorithms have significantly progressed with the advent of deep learning. These algorithms aim to detect and localize objects within images or video frames. Two popular object detection algorithms are Faster R-CNN [13] and SSD (Single Shot MultiBox Detector) [14]. Faster R-CNN uses a two-stage approach, proposing regions of interest and then classifying them [15]. On the other hand, SSD directly predicts the bounding boxes and class probabilities. These algorithms have shown promising results in detecting various objects, but their effectiveness in VRU detection needs further investigation.

C. You Look Only Once

The YOLO algorithm has gained significant attention in computer vision due to its real-time and accurate object detection capabilities. YOLO is a one-stage object detection algorithm that directly predicts bounding boxes and class probabilities from full images in one evaluation [17]. It has been widely applied in various domains, such as real-time vehicle detection [18], pedestrian location estimation from UAV imagery [19], and obstacle avoidance in micro aerial vehicle navigation [20].
YOLO Nano was proposed as a highly compact CNN for object detection, achieving a mean average precision (mAP) of approximately 69.1% on the VOC 2007 dataset while significantly smaller and requiring fewer operations than previous YOLO versions [21]. Similarly, YOLO-LITE was optimized for non-GPU computers, making real-time object detection feasible on such systems [22].
The YOLO algorithm has been enhanced for specific tasks, such as ship target detection [23], robust vehicle detection [24], subway passenger flow detection [25], marine life observation [26], and traffic flow monitoring [27]. Researchers have also proposed modifications to YOLO, such as incorporating a new residual structure, to improve its performance and address challenges in detecting objects in UAV aerial images [28]. Another study introduced an asymmetric centre point bounding box regression strategy and angle loss-based YOLO for object detection, aiming to improve the accuracy and robustness of the YOLO method [29].

D. Related Works

Some of the recent studies have focused on VRU detection using different methodologies. Teixeira (2023) proposed a multi-sensing and communication approach, leveraging smart city sensors and onboard units to predict and notify potential collisions [30]. Frampton (2022) highlighted the limitations of current sensor systems in heavy goods vehicles, suggesting the need for standardization and additional front sensors [31].
Ghorai (2020) introduced a longitudinal control algorithm for cooperative autonomous vehicles, demonstrating its effectiveness in avoiding collisions with vulnerable road users [32]. Jahangiri (2020) developed a computer vision-

based decision support system for intersection safety monitoring, providing a proactive approach to assessing safety [33].
Integrating technologies such as machine learning, mobile devices for V2X connections [34], and C-V2X communication [35] showcases the diverse approaches explored to improve VRU safety. Additionally, research on the impact of LTE-V2X connectivity on global occupancy maps [36] and the use of smartphones for path prediction and early warning systems [37] further illustrate the evolving landscape of VRU safety measures. Initiatives like probabilistic prediction of collisions between cyclists and vehicles [38] and the development of emergency braking systems for trucks [39] represent the multifaceted strategies employed to mitigate VRU collisions.

III. METHODOLOGY

A. Data collection and preprocessing

A robust dataset of 4490 images was meticulously curated to enhance the effectiveness of a road user detection system. The dataset encompasses images capturing various traffic situations sourced from vehicle cameras to offer diverse perspectives and lighting conditions. A series of preprocessing steps were undertaken to refine the dataset. This included standardizing image resolutions, eliminating extraneous data, and optimizing contrast and brightness levels to enhance clarity.
Manual annotation was conducted by delineating bounding boxes around vulnerable road users like pedestrians, cyclists, and motorcyclists. The dataset's strength lies in its inclusivity of different road user categories, environmental variables, weather conditions, and geographical contexts, ensuring a comprehensive training and validation process for the model. Subsequently, the dataset was divided into training and validation sets in an 80:20 ratio using a stratified sampling method to maintain a balanced representation of road user types across both sets.

B. YOLOv8 architecture

The architecture of YOLOv8 integrates several key features and enhancements in loss functions and network structures, contributing to an overall improvement in performance. Also, it benefits from improved feature fusion modes, which play a crucial role in enhancing detection capabilities, particularly for small objects. It builds upon the YOLOv7 architecture [40]. The YOLOv8 architecture comprises a backbone network, a feature pyramid network (FPN), and a detection head [41], as shown in Fig. 1.

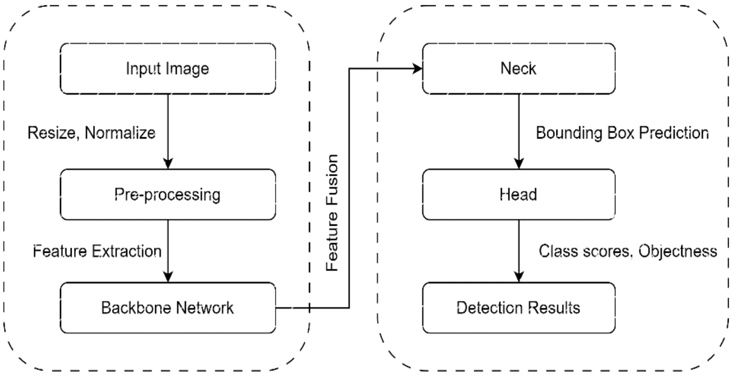

Figure 1. High-level YOLOv8 architecture

The backbone network serves as the feature extractor and captures high-level representations of the input image. We utilize a pre-trained backbone network, Darknet-53, to leverage its learned features. The FPN enhances the feature maps by combining multi-scale information, enabling the model to detect objects of different sizes. The detection head predicts each detected object’s bounding boxes and class probabilities.

C. Training Process

The YOLOv8 model was initialized with weights pre-trained on the large-scale COCO dataset to leverage existing learned feature representations pertinent to general object detection. Fine-tuning was conducted on a dataset focused on vulnerable road users to adapt the model to the target detection task. Stochastic gradient descent (SGD) optimization was employed during fine-tuning with a learning rate of 0.001 to allow for gradual parameter updates that facilitate convergence. A batch size of 32 images and 30 training epochs were used to strike an appropriate balance between update frequency and training time.

Additional optimizations included a momentum of 0.9 to accumulate gradient effects over batches for accelerated training and weight decay of 0.0005 as a regularisation method to prevent overfitting. The training process involved multiple iterations, each consisting of a forward pass, loss computation, and gradient-based update. The loss function combined localization and classification losses, primarily the mean square error (MSE) loss. L_{loc} and the cross-entropy loss L_{cls}. This provides a combined optimization target encompassing both accurate localization and classification.

$$L_{Total} = L_{loc} + L_{cls} \quad (1)$$

$$L_{loc} = \frac{1}{N}\sum(y_{true} - y_{pred})^2 \quad (2)$$

Where N = number of samples, y_{true} = True label, y_{pred} = Predicted label. This loss function L_{loc} measures the deviation between the predicted bounding box and ground truth box coordinates. Minimizing this helps get accurate box predictions.

$$L_{cls} = -\sum y_{true} * \log(y_{pred}) \quad (3)$$

This loss function L_{cls} calculates the error between predicted class probabilities and true labels. Lower values mean better classification accuracy.

Extensive data augmentation through techniques like random image rotation, scaling, and translation was also introduced to enhance the model's generalization capability by exposing it to expanded variations during training. These comprehensive strategies enabled effective learning on the vulnerable road user dataset to produce a performant detection model specialized for the safety-critical application.

D. Evaluation Metrics

A thorough evaluation assessed the algorithm's performance on the collected dataset. Multiple evaluation metrics were used to analyze the efficacy of the YOLOv8 algorithm comprehensively. These metrics include precision, recall, average precision (AP), F1-score and intersection over union (IoU).

Precision quantifies the proportion of correctly predicted vulnerable road users among all the predictions made by the algorithm. It measures the algorithm's ability to avoid false positives. Recall, however, captures the proportion of correctly predicted vulnerable road users out of all the ground truth instances. It assesses the algorithm's ability to identify all the instances accurately. F1-Score combines precision and recall to provide an overall performance assessment.

$$P = \frac{TP}{TP+FP} \quad (4)$$

$$R = \frac{TP}{TP+FN} \quad (5)$$

$$F1 = 2 * \frac{(P*R)}{P+R} \quad (6)$$

Average precision (AP) is a widely used aggregate metric that combines precision and recall at multiple thresholds. It summarizes the precision-recall curve and indicates overall algorithm performance.

$$AP = \sum_n (R_n - R_{n-1}) * P_n \quad (7)$$

Where P_n and R_n are the precision and recall at the nth threshold.

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i \quad (8)$$

Where N is the number of classes and AP_i is the average precision for class i.

IoU measures the extent of match between the predicted bounding box and the ground truth bounding box. A higher IoU value indicates a more accurate localization of vulnerable road users.

$$IoU = \frac{Area\ of\ (B_{pred} \cap B_{gt})}{Area\ of\ (B_{pred} \cup B_{gt})} \quad (9)$$

Where B_{pred} is the area of the predicted bounding box, and B_{gt} is the area of the ground truth bounding box.

E. Real-world Testing

For real-world testing, we used a monocular vision camera mounted on the vehicle's front windshield to conduct our experiments. These experiments help verify the system's effectiveness in detecting vulnerable road users in previously unseen and unconstrained settings. The camera captures the real-time video feed of the surrounding environment, which is then processed by the YOLOv8 algorithm for vulnerable road user detection.

IV. Experimental results

A. Training the deep learning model

Our model was trained on an Ubuntu 22.04 workstation equipped with an Nvidia RTX 3090 GPU, an Intel i9-12900K CPU, and 64GB of RAM. We chose the PyTorch deep learning framework for implementation due to its robust optimization capabilities and efficient GPU utilization.

The training involved processing a comprehensive training set to ensure robust learning. We then evaluated the model's performance using a separate validation set. Analysis of evaluation results, depicted in Fig. 2, provides insights into the algorithm's effectiveness in detecting vulnerable road users.

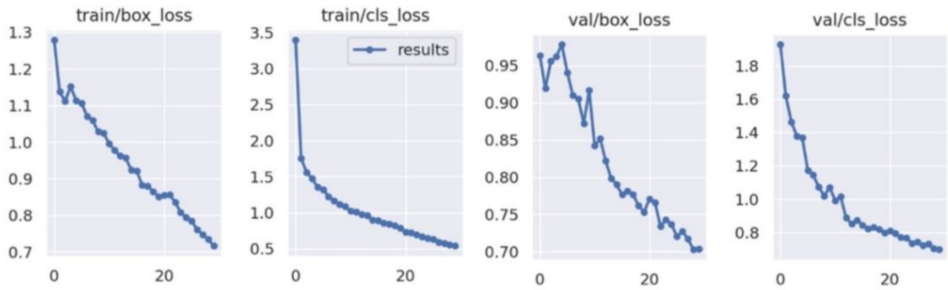

Figure 2. Graph of training and validation loss

The consistent decrease in training losses indicates effective learning, while stable and lower validation losses suggest good generalization capabilities. However, variability in classification loss warrants closer scrutiny of the model's performance in classifying different categories, possibly benefiting from further fine-tuning or data augmentation strategies.

B. Performance evaluation

Our system findings are shown in Table I. The efficacy in identifying vulnerable road users is demonstrated by the F1-score of 0.81, which indicates a solid balance between precision and recall. Using a single GPU, the program analyzed images at an average rate of 45 frames per second (FPS).

TABLE I. RESULTS OF VRU DETECTION USING YOLOV8

Metric	Value
Precision	0.85
Recall	0.78
F1-score	0.81
mAP	0.76
FPS	45

To further analyze the algorithm’s performance, we conducted experiments to measure its detection accuracy for different types of vulnerable road users. Table II presents the precision, recall, and F1-score for each class of vulnerable road users.

TABLE II. PERFORMANCE FOR DIFFERENT CLASSES

Class	Precision	Recall	F1-score
Pedestrian	0.87	0.84	0.85
Cyclist	0.82	0.78	0.80
Motorcyclist	0.88	0.85	0.86

The results demonstrate that the YOLOv8 algorithm performs well for all vulnerable road users. As shown in Fig. 3, our proposed YOLOv8 approach achieves a mAP of 0.76. The precision of our proposed approach is measured to be 0.85, indicating that the algorithm can correctly identify vulnerable road users among other objects. The recall value of 0.78 implies that our approach accurately detects a significant proportion of vulnerable road users in the dataset.

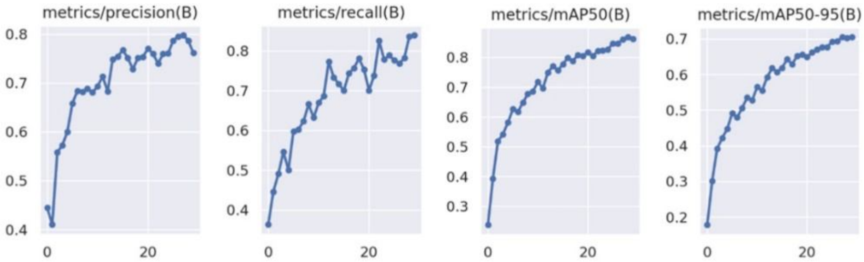

Figure 3. Graph of performance metrics

We analyzed the model's performance under different scenarios, such as varying lighting conditions, weather conditions, and occlusion levels. This analysis helps us understand the model's strengths and limitations in real-world scenarios, as shown in Fig. 4.

Figure 4. VRU detection on real-world images

C. Comparison with Other Models

We selected widely used object detection models for comparison, including Faster R-CNN, SSD, and previous versions of YOLO. These models were chosen due to their prevalence in VRU detection research and their varied approaches to object detection. The comparison of their performance on our VRU dataset is presented in Table III and Table IV.

TABLE III. COMPARATIVE EVALUATION OF PERFORMANCE METRICS

Model	Precision	Recall	F1-Score	mAP	Speed (FPS)
Faster R-CNN	0.80	0.72	0.76	0.71	16
SSD	0.78	0.70	0.74	0.69	20
YOLOv5	0.76	0.68	0.72	0.67	30
YOLOv7	0.83	0.75	0.79	0.74	35
YOLOv8	0.85	0.78	0.81	0.76	45

TABLE IV. FALSE POSITIVES AND NEGATIVES COMPARISON

Model	False Positives	False Negatives
Faster R-CNN	203	178
SSD	182	157
YOLOv5	211	194
YOLOv7	172	136
YOLOv8	153	124

The assessment of model resilience in difficult real-world conditions indicates that YOLOv8 experienced the slightest performance degradation across metrics. Table V highlights the robustness of each object detection model against real-world perceptual and environmental challenges pertinent to VRU detection applications under diverse weather conditions.

TABLE V. MODEL ROBUSTNESS IN VARIOUS CONDITIONS

Model	Low Light	Occlusion	Weather Variability
Faster R-CNN	Accuracy suffers significantly	Noticeable degradation	Dense rain/snow problematic
SSD	Major decline in accuracy	Small objects missed	Rain/fog causes issues
YOLOv5	High miss rate	Many missed detections	Large variability with conditions
YOLOv7	Moderate drop in accuracy	Match YOLOv8	Nearly as good
YOLOv8	Minimal effect on accuracy	Small impact	Handles well

D. Ablation Study

We conducted an ablation study using different variables to further analyze our system's performance. Table VI presents the ablation study's results, showcasing each factor's impact on the system's performance. The factors considered in the study include using different input image resolutions, the effects of data augmentation techniques, and the impact of using different confidence thresholds for detection.

TABLE VI. ABLATION STUDY RESULTS OF PROPOSED MODEL

Factor	Precision	Recall	F1-score	mAP
Baseline (Original Configuration)	0.85	0.78	0.81	0.76
Increased Image Resolution	0.88	0.81	0.84	0.79
Data Augmentation	0.86	0.79	0.82	0.77
Lower Confidence Threshold	0.83	0.76	0.79	0.74

Increasing the image resolution improved performance with increased precision, recall, F1-score, and mAP. This indicates that higher-resolution images provide more detailed information, leading to better detection results. Using data augmentation techniques also resulted in a slight performance improvement, demonstrating the effectiveness of such methods in enhancing the system's ability to generalize to different scenarios.

On the other hand, decreasing the confidence threshold for detection decreased precision, recall, F1-score, and mAP. This suggests that lowering the threshold leads to more false positives and missed detections, reducing the system's overall performance. Therefore, it is crucial to carefully select an appropriate confidence threshold to balance between false positives and missed detections.

V. DISCUSSION

Our vulnerable road user detection system demonstrates computer vision-based systems' potential to detect vulnerable road users in real-time scenarios effectively. The ablation study provided insights into the impact of different factors on the system's performance, highlighting the importance of factors such as image resolution and confidence threshold selection.

This finding has significant implications for road safety and VRU protection. We can significantly minimize the danger of accidents involving pedestrians, cyclists, and other non-motorized road users by improving the accuracy and responsiveness of VRU identification and collision avoidance systems.

An alert or system response can be deployed upon detecting vulnerable road users to notify the driver and prevent potential collisions. Responses may include visual alerts, auditory warnings, or haptic feedback. Visual alerts could manifest as warning symbols or vehicle dashboard heads-up displays. Auditory warnings might involve loud beeps or voice commands, while haptic feedback could vibrate the driver's seat or steering wheel. Careful consideration of the timing and intensity of these alerts is crucial to ensure they are informative and manageable for the driver.

While the study results are promising, certain limitations need to be acknowledged. The dataset may only cover specific scenarios, potentially limiting the findings' generalizability. Also, the performance evaluation focused solely on vulnerable road user detection, necessitating further analysis to assess the algorithm's performance in complex traffic situations.

VI. CONCLUSION

This paper presented a real-time vulnerable road user detection system using the YOLOv8 deep learning architecture. The results demonstrate that YOLOv8 can detect VRUs with high accuracy, achieving an average precision of 0.85 and recall of 0.78 over a range of conditions. The high processing speed of 45 FPS makes this system well-suited for safety-critical applications, including collision avoidance and driver assistance.

This research contributes towards the broader goals of reducing traffic injuries and fatalities by enhancing the ability of vehicles to detect vulnerable road users rapidly and reliably. As autonomous driving matures, precise and low-latency VRU detection will become essential for ensuring safe coexistence between cars and pedestrians/cyclists.

Future research can focus on deploying and evaluating the proposed approach in real-world driving scenarios. Additional work on optimizing system performance and expanding detectable VRU categories is also warranted.

ACKNOWLEDGEMENT

We express our gratitude to the E&TC Research Centre at Pimpri Chinchwad College of Engineering, Pune, as well as our colleagues for their invaluable help and support during the execution of this research.

REFERENCES

[1] "ITS & Vulnerable Road Users." Accessed: Aug. 15, 2023. [Online]. Available: https://transport.ec.europa.eu/transport-themes/intelligent-transport-systems/road/action-plan-and-directive/its-vulnerable-road-users_en

[2] J. Scholliers et al., "Impact Assessment of Its Applications for Vulnerable Road Users," Transportation Research Procedia, vol. 14, pp. 4515–4524, 2016, doi: 10.1016/j.trpro.2016.05.374.

[3] World Health Organization, "GLOBAL STATUS REPORT ON ROAD SAFETY," 2018.

[4] ERTRAC, "ADAS European Road Safety 2016," European COMMISSION, 2016.

[5] N. Highway Traffic Safety Administration and U. Department of Transportation, "Comparing Demographic Trends in Vulnerable Road User Fatalities and the U.S. Population, 1980–2019," 1980. [Online]. Available: https://cloud.safe.nsc.org/

[6] "ROAD ACCIDENTS IN INDIA ROAD ACCIDENTS IN INDIA ROAD ACCIDENTS IN INDIA," 2021, Accessed: Dec. 02, 2023. [Online]. Available: www.morth.nic.in

[7] D. Shinar, "Safety and mobility of vulnerable road users: Pedestrians, bicyclists, and motorcyclists," Accid Anal Prev, vol. 44, no. 1, pp. 1–2, 2012, doi: 10.1016/j.aap.2010.12.031.

[8] P. Sewalkar and J. Seitz, "Vehicle-to-pedestrian communication for vulnerable road users: Survey, design considerations, and challenges," Sensors (Switzerland), vol. 19, no. 2, 2019, doi: 10.3390/s19020358.

[9] P. Dollár, C. Wojek, B. Schiele, and P. Perona, "Pedestrian detection: An evaluation of the state of the art," IEEE Trans Pattern Anal Mach Intell, vol. 34, no. 4, pp. 743–761, 2012, doi: 10.1109/TPAMI.2011.155.

[10] M. Dimitrievski, D. Van Hamme, P. Veelaert, and W. Philips, "Cooperative Multi-Sensor Tracking of Vulnerable Road Users in the Presence of Missing Detections," Sensors, vol. 20, no. 17, pp. 1–40, Sep. 2020, doi: 10.3390/S20174817.

[11] A. Bighashdel and G. Dubbelman, "A Survey on Path Prediction Techniques for Vulnerable Road Users: From Traditional to Deep-Learning Approaches," 2019 IEEE Intelligent Transportation Systems Conference, ITSC 2019, no. 783190, pp. 1039–1046, 2019, doi: 10.1109/ITSC.2019.8917053.

[12] D. Charlebois et al., "The ideal vulnerable road user - a study of parameters affecting VRU detection," Traffic Inj Prev, vol. 24, no. sup1, pp. S62–S67, 2023, doi: 10.1080/15389588.2022.2159762.

[13] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," IEEE Trans Pattern Anal Mach Intell, 2017, doi: 10.1109/tpami.2016.2577031.

[14] W. Liu et al., "SSD: Single shot multibox detector," Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 9905 LNCS, pp. 21–37, 2016, doi: 10.1007/978-3-319-46448-0_2/FIGURES/5.

[15] X. Zhao, W. Li, Y. Zhang, T. A. Gulliver, S. Chang, and Z. Feng, "A faster RCNN-based pedestrian detection system," IEEE Vehicular Technology Conference, vol. 0, 2016, doi: 10.1109/VTCFall.2016.7880852.

[16] Z. Q. Zhao, P. Zheng, S. T. Xu, and X. Wu, "Object Detection with Deep Learning: A Review," IEEE Trans Neural Netw Learn Syst, vol. 30, no. 11, pp. 3212–3232, 2019, doi: 10.1109/TNNLS.2018.2876865.

[17] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection." [Online]. Available: http://pjreddie.com/yolo/

[18] H. Wang, X. Lou, Y. Cai, Y. Li, and L. Chen, "Real-Time Vehicle Detection Algorithm Based on Vision and Lidar Point Cloud Fusion," vol. 2019, 2019.

[19] O. Kainz, M. Gera, M. Michalko, and F. Jakab, "Experimental Solution for Estimating Pedestrian Locations from UAV Imagery," Applied Sciences, vol. 12, no. 19, Oct. 2022, doi: 10.3390/APP12199485.

[20] S. Karlsson, C. Kanellakis, S. S. Mansouri, and G. Nikolakopoulos, "Monocular Vision-based Obstacle Avoidance Scheme for Micro Aerial Vehicle Navigation," 2021 International Conference on Unmanned Aircraft Systems (ICUAS), pp. 1321–1327, Jun. 2021, doi: 10.1109/ICUAS51884.2021.9476793.

[21] A. Wong, M. Famuori, M. J. Shafiee, F. Li, B. Chwyl, and J. H. Chung, "YOLO Nano: a Highly Compact You Only Look Once Convolutional Neural Network for Object Detection", doi: 10.48550/ARXIV.1910.01271.

[22] J. Pedoeem and R. Huang, "YOLO-LITE: A Real-Time Object Detection Algorithm Optimized for Non-GPU Computers", doi: 10.48550/ARXIV.1811.05588.

[23] J. Zhou, P. Jiang, A. Zou, X. Chen, and W. Hu, "Ship Target Detection Algorithm Based on Improved YOLOv5," J Mar Sci Eng, vol. 9, no. 8, Aug. 2021, doi: 10.3390/JMSE9080908.
[24] S. Kumar, M. Jailia, S. Varshney, N. Pathak, S. Urooj, and N. A. Elmunim, "Robust Vehicle Detection Based on Improved You Look Only Once," Computers, Materials &Amp; Continua, vol. 74, no. 2, pp. 3561–3577, 2023, doi: 10.32604/CMC.2023.029999.
[25] X. Li, Y. Zhang, D. He, X. Teng, B. Liu, and Y. Chen, "Passenger Flow Detection in Subway Stations Based on Improved You Only Look Once Algorithm," Transportation Research Record: Journal of the Transportation Research Board, vol. 2677, no. 9, pp. 397–409, Sep. 2023, doi: 10.1177/03611981231159128.
[26] H. Stavelin, A. Rasheed, O. San, and A. J. Hestnes, "Marine life through You Only Look Once's perspective", doi: 10.48550/ARXIV.2003.00836.
[27] C. Y. Cao, J. C. Zheng, Y. Q. Huang, J. Liu, and C. F. Yang, "Investigation of a Promoted You Only Look Once Algorithm and Its Application in Traffic Flow Monitoring," Applied Sciences, vol. 9, no. 17, Sep. 2019, doi: 10.3390/APP9173619.
[28] P. Zhang, H. Deng, and Z. Chen, "RT-YOLO: A Residual Feature Fusion Triple Attention Network for Aerial Image Target Detection," Computers, Materials &Amp; Continua, vol. 75, no. 1, pp. 1411–1430, 2023, doi: 10.32604/CMC.2023.034876.
[29] X. Wang, L. Kong, Z. Zhang, H. Wang, and X. Lu, "ACP-YOLO: Asymmetric Center Point Bounding Box Regression Strategy and Angle Loss-based YOLO for Object Detection", doi: 10.21203/RS.3.RS-3103985/V1.
[30] P. Teixeira, S. Sargento, P. Rito, M. Luis, and F. Castro, "A Sensing, Communication and Computing Approach for Vulnerable Road Users Safety," IEEE Access, vol. 11, pp. 4914–4930, 2023, doi: 10.1109/ACCESS.2023.3235863.
[31] R. J. Frampton and J. E. Millington, "Vulnerable Road User Protection from Heavy Goods Vehicles Using Direct and Indirect Vision Aids," Sustainability 2022, Vol. 14, Page 3317, vol. 14, no. 6, p. 3317, Mar. 2022, doi: 10.3390/SU14063317.
[32] P. Ghorai and A. Eskandarian, "Longitudinal Control Algorithm for Cooperative Autonomous Vehicles to Avoid Accident with Vulnerable Road Users," 2020 IEEE 23rd International Conference on Intelligent Transportation Systems, ITSC 2020, Sep. 2020, doi: 10.1109/ITSC45102.2020.9294180.
[33] A. Jahangiri et al., "Developing a Computer Vision-Based Decision Support System for Intersection Safety Monitoring and Assessment of Vulnerable Road Users," Mineta Transportation Institute Publications, Mar. 2020, doi: 10.31979/mti.2020.1853.
[34] B. Ribeiro, M. J. Nicolau, and A. Santos, "Using Machine Learning on V2X Communications Data for VRU Collision Prediction," Sensors 2023, Vol. 23, Page 1260, vol. 23, no. 3, p. 1260, Jan. 2023, doi: 10.3390/S23031260.
[35] L. Miao, S. F. Chen, Y. L. Hsu, and K. L. Hua, "How Does C-V2X Help Autonomous Driving to Avoid Accidents?," Sensors 2022, Vol. 22, Page 686, vol. 22, no. 2, p. 686, Jan. 2022, doi: 10.3390/S22020686.
[36] N. Mouawad, V. Mannoni, B. Denis, and A. P. Da Silva, "Impact of LTE-V2X Connectivity on Global Occupancy Maps in a Cooperative Collision Avoidance (CoCA) System," IEEE Vehicular Technology Conference, vol. 2021-April, Apr. 2021, doi: 10.1109/VTC2021-SPRING51267.2021.9449034.
[37] Vourgidis, L. Maglaras, A. S. Alfakeeh, A. H. Al-Bayatti, and M. A. Ferrag, "Use Of Smartphones for Ensuring Vulnerable Road User Safety through Path Prediction and Early Warning: An In-Depth Review of Capabilities, Limitations and Their Applications in Cooperative Intelligent Transport Systems," Sensors 2020, Vol. 20, Page 997, vol. 20, no. 4, p. 997, Feb. 2020, doi: 10.3390/S20040997.
[38] D. Pan et al., "Probabilistic Prediction of Collisions between Cyclists and Vehicles Based on Uncertainty of Cyclists' Movements," https://doi.org/10.1177/03611981221121270, vol. 2677, no. 3, pp. 1151–1164, Sep. 2022, doi: 10.1177/03611981221121270.
[39] X. Wang, Y. Zhou, and S. D. Cafiso, "Automatic Emergency Braking System for Trucks for Two-Wheeler Longitudinal Collision Avoidance," https://doi.org/10.1177/03611981231166374, vol. 2677, no. 11, pp. 565–577, May 2023, doi: 10.1177/03611981231166374.
[40] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," pp. 7464–7475, Jul. 2022, doi: 10.1109/cvpr52729.2023.00721.
[41] "What is YOLOv8? The Ultimate Guide." Accessed: Dec. 02, 2023. [Online]. Available: https://blog.roboflow.com/whats-new-in-yolov8/#what-is-yolov8