Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol. 14, Issue 18s (2025)
https://www.jneonatalsurg.com

OPEN ACCESS

Helmet Detection System for Two-Wheeler Riders Using Yolo Machine Learning Algorithms

Yogeshwer Sharma[1], Dr. Prof. Pinky Rane[2]

[1]Research Scholar, Department of Computer Science and Engineering, Medi-Caps University, Indore, Madhya Pradesh, India, en23cs501038@medicaps.ac.in

[2]Assistant Professor, Department of Computer Science and Engineering, Medi-Caps University, Indore, Madhya Pradesh, India, pinky.rane@medicaps.ac.in

Cite this paper as: Dr. Geeta Gupta, Dr. Yogesh Gupta, (2025) Helmet Detection System for Two-Wheeler Riders Using Yolo Machine Learning Algorithms. *Journal of Neonatal Surgery*, 14 (18s), 323-328.

ABSTRACT

Helmets are essential safety measures for anyone using two-wheeled vehicles (motorcycles, bicycles, and e-scooters), and the absence of protective helmets may result in serious or fatal injuries. The principal technique for helmet detection is now a series of Convolutional Neural Network methods. Detection accuracy, speed prediction, and ease of deployment are essential criteria for achieving road safety. Conventional object identification methods often fail to provide consistent performance across all domains. This study presents a helmet identification application using the latest You only look once version 7 (YOLOv7) algorithm enhanced by an attention-based approach. The model's performance was assessed using a collection of helmet test photos, achieving an average accuracy (mAP@0.5) of 91.4%. The findings demonstrate great detection accuracy and minimal computing requirements, making the model appropriate for practical use. Consequently, the suggested model may aid in addressing the issue of helmet detection on two-wheeled vehicles.

Keywords: Object detection, helmet, YOLO, Machine learning, Two wheelers

1. INTRODUCTION

In addition to being popular, eco-friendly, and usually not requiring the search for a parking spot, riding two-wheeled motor vehicles such as motorcycles, bicycles, e-scooters, etc. The popularity of cycling, whether propelled or not, has skyrocketed in the last several years. According to many studies (Useche et al., 2022; Störmann et al., 2020; Useche et al., 2021). Two-wheeled vehicles pose a greater risk of injury since riders often lack adequate protection while riding them. The number of motorcycle, e-scooter and bicycle accidents in the US reached 124,002 in 2021, with 906 lives lost (NTSB, 2022). There were 473 motorbike-licensed deaths, 372 bicycle-related deaths (131 of which included e-bikes), and 56 e-scooter-related deaths. Riders on motorbikes, bicycles, or electric scooters accounted for 38% of all traffic injuries and 35% of all road deaths in 2022 (NTSB, 2022). Despite this, a lot of people still don't think it's a good idea to wear a helmet. The city's transport infrastructure has become less resilient due to the increased number of accidents caused by the increased traffic in recent years (Lacinák, 2021). While constructing a resilient city, it is important to keep economic factors in mind (Ristvej et al., 2018). In order to keep track of traffic, many public and private buildings now have cameras installed. The data collected by these cameras is quickly finding new applications thanks to the development of computerised image processing. Persona recognition and road user profiling play an essential role in the analytical and security sectors. Among the most important preventative actions to protect lives in and out of traffic accidents, helmets are a must-have. In order to greatly reduce the occurrence of serious head injuries, it is crucial to promptly identify individuals who are riding two-wheeled vehicles while wearing safety helmets. There would be no need for ground people if the relevant safety authorities and cities were immediately alerted. However, these methods need substantial investments of time and money, therefore automated detection is the way to go. Logical and effective data storage is one of several factors that would be required for such detection. As part of the suggested system, we must further think about the data intensity, which refers to the processing and transfer of data in real-time across devices, which encompasses a broad spectrum for ITS (Zabovsky et al., 2010; Bučko et al., 2021). Bučko et al. (2019) suggests that ontologies might be a processing strategy. There are a number of problems with visual identification that are associated with processing photos taken at a vast distance. These concerns include, for instance, the degree of illumination and the quantity of light in the visual scene, the degree to which certain parts of the head item resemble one another (for instance, a cap or hat might, from some perspectives, seem like a helmet), and so on. Several moving road users occupying overlapping frames, the

displacement of stationary road users, and the angle of capture of the camera are additional aspects to be taken into account (Tong et al., 2020). Furthermore, it is possible that details about the rider's helmet use will be lost due to the photographs' poor quality. The obtained picture cannot be used to detect the existence of a helmet under these conditions. The deep learning algorithm's detecting effect is still below the level of precision required for practical applications. Particularly when the frames are affected by illumination and angle variations, the safety helmet's detection performance in traffic drops. Convolutional Neural Networks (CNNs) are widely used for target identification with the advent of deep learning techniques. This is because CNNs are very good at feature extraction. In a series of studies, helmet detection has been approached by two-stage algorithms such as R-CNN, Faster R-CNN, Cascade R-CNN, Sparse R-CNN, and one-stage algorithms such as SSD and YOLO. Many of them identified that, due to its relatively greater inference time, YOLO is preferred in transportation safety helmet identification; examples include previous works such as Zhou et al. (2021), Jia et al. (2021), Zhang et al. (2022), and Otgonbold et al. (2022). The SAS-YOLOv3-tiny algorithm provided a better detection algorithm and enhanced the accuracy of helmet recognition. For example, Cheng et al. (2021) proposed scenario for the detection of helmets in security surveillance by upgrading the YOLOv3 backbone network to ResNet, introducing attention mechanisms, and others. Using YOLOv5, Zhou et al. (2021) came up with a method that was able to track when a wearer was wearing a helmet. Their efforts did contribute to the great improvement of precision and rapidity of the test. By modifying the segmentation structure of the YOLOv5 Dense-Block backbone network and adding an SE attention mechanism to the neck network, an improved YOLOv5s algorithm was proposed in Zhang et al. (2022), further enhancing the accuracy of target recognition. For a helmet detection task, Jia et al. (2021) proposed YOLOv5 with soft-NMS and improved triplet-attention fusion. Considering the large database, we had for safety helmets, in this study an improved object detector was employed, which is YOLOv7 Wang et al. (2022), combined with the Selectively Paying Attention Mechanism Woo S. et al. (2018), to accomplish real-time helmet detection. There are different sizes of YOLOv7 architectures, being the most recent generation. The performances of the suggested model have been compared to prior works of YOLOv3 from Redmon and Farhadi (2018) and YOLOv5 from Jocher (2020).
The outline of this article will thus be: An overview detailing identification of safety helmets in situations such as in-traffic in correspondence to a YOLOv7 second-to-latest single-stage object detector on grounds of state-of-the-art advancements from older previous improvements. This takes place within the third section as the sections carry measurements of a detailed experiment's setup, and the methodologies are related to those stated in a more detailed fashion regarding datasets used within work. Comparison takes place for acquired data from that of various others that exist toward performing their recognition process under safety helmets within Section 4.

2. OBJECT DETECTION METHOD

The YOLOv7, an improved version of the original YOLO object detection system, is the subject of this section because of its enhanced ability to identify tiny things. On top of that, the attention mechanism (AM) is used to make tiny object detection more accurate.

2.1. YOLO

Since its introduction in 2015, YOLO has successfully improved detection accuracy while compensating for the two-stage detection network's slow inference time. A better version of the previous work was YOLOv3 (Redmon and Farhadi, 2018). The Feature Pyramid Network (FPN) architecture and the Darknet-53 residual module were its defining features; they allowed for multi-level fusion and item prediction on three different dimensions. Improved variants of YOLOv3 with upgraded backbone and neck architectures were introduced thereafter as YOLOv4 (Bochkovskiy et al., 2020) and YOLOv5 (Jocher, 2020). Outperforming previous architectures on the MC COCO detection challenge, YOLOv7 was designed in 2022 (Wang et al., 2022) and significantly boosts detection speed and accuracy. Each YOLO architecture has a spine that extracts features, a neck that gathers feature maps, and a head that does the detection. The YOLOv7 network uses an E-ELAN, or Extended Effective Layer Aggregation Network, instead of the Darknet. Using the cardinality of expand, shuffle, and merge, the E-ELAN research continuously improves the network's learning capability while keeping the gradient route intact. An FPN-PAN structure is created by combining the advantages of the FPN and the Path Aggregation Network (PAN) in the neck component of YOLOv7. YOLOv7 has a system with many heads. A secondary Auxiliary Head aids in training inside the intermediary layers, while the main Lead Head sorts the final output detections. As with YOLOv6, the creators of this version used compound model scaling to get their models ready for release. In addition to improved scalability, YOLOv7 features re-parameterization (RP). In order to achieve robustness, RP averages the modules. Certain parts of the model have their own unique approaches, and this is the subject of an ongoing study on module-level re-parameterization. The YOLOv7 model modules that need to be recalibrated are determined by the gradient flow propagation channels.

Table 1. Comparison between structures of YOLOv3, YOLOv5 and YOLOv7.

	YOLOv3	**YOLOv5**	**YOLOv7**
Network Type	Fully Connected	Fully Connected	Fully Connected
Backbone	Darknet-53	CSPDarknet53	E-ELAN

Neck	Feature Pyramid Network	Path Aggregation Network	FPN and PAN
Head	YOLO Output Layer	YOLO Output Layer	Multi-head Framework

2.2. Selectively Paying Attention Mechanism

Our experiment's AM further improves the accuracy of YOLO detection for tiny objects. AM, according to Yang 2020 and Guo et al. 2022 is a signal processing system similar in operations to the way the human brain works. In this way, the approach effectively works in the application of computer vision. An example of attention mechanism developed for dealing with issues brought about by redundancy in deep learning network information flow is Convolutional Block Attention Module. This is a type that was developed and introduced by Woo S. Et al. 2018 have applied the module in many computer vision applications. Using a set of AM in both the channel and spatial dimensions, the CBAM module can infer the importance of pixels within an input picture or feature map. The module first generates a channel attention map and a spatial dimension attention map. With channel dependencies, the channel attention map computes the importance of each feature channel, while the spatial attention map tracks the importance of each feature location internally within the feature map. Combining the two attention maps, the CBAM module generates a weighted feature map with a focus on the most important and informative characteristics. This method improves the precision of tiny item detection by removing unnecessary elements from the feature map and emphasizing the most important ones. Would you like more details on CBAM? Check out Woo S. et al. (2018).

3. EXPERIMENTATION

This research uses one publicly available dataset to propose a new architecture of protective helmet detection in two-wheeler safety using deep learning. More precisely, we employed the latest YOLOv7 object identification method in computer vision for identifying helmets in source photos. Figure 1 shows the standard approach for detecting helmets. Gathering data, cleaning it up, training the model, and finally evaluating it are all parts of the process.

Fig. 1. Detection of helmets using YOLO

3.1. Experimental Setup

The Windows operating system, an Intel Core i9 12900HX Alder Lake CPU, and an NVIDIA GeForce RTX 3080Ti 24GB GPU were all used in our studies. Python 3.8, CUDA 11, and PyTorch 1.8.1 were the frameworks used. The experimental training technique parameters are shown in Table 2. The images were automatically enlarged to 1080 pixels wide and modified to 1080 pixels tall, keeping the original aspect ratio in mind.

Table 2. Training parameters

Parameter	Value	Parameter	Value
Learning Rate	0.01	Weight Decay	0.0005
Batch Size	16	Momentum	0.8
Image Size	1080 × 1080	Epochs	200

3.2. Dataset Description

Researchers in this research combed through the Helmet Detection (2022) dataset in search of helmet-wearing bikers, scooter riders, and motorcyclists. The finished product includes more than 4000 photos annotated with bounding boxes. Because there are only two possible answers, "with helmet" and "without helmet," we may call this a binary classification issue. Figure 2 displays the example photos of various settings and helmets of varied sizes. In addition, there are three distinct categories for the dataset: training, testing, and validation. The three datasets consist of three hundred and one photographs each: the training set, the testing set, and the validation set.

Fig. 2. Image detection for persons wear helmet

3.3. Evaluation Metrics

Computer vision contests and challenges are always testing the limits of accuracy in object recognition, which is a dynamic area of study (Everingham et al., 2010; Russakovsky et al., 2015). The output of a classification model is simple: the class of the item and a confidence value. While object classification models return simple outputs, object detection models return much more involved results: item classification is complemented with the bounding box and a set of confidence measures for every prediction. The coordinates of the top left and lower right corners define a rectangular bounding box that delineates the position of an item. That having been said, one of the most common ways to evaluate multi-class detectors on a very basic level would be by having a look at their mAP values. In this case, area underneath this curve basically means average precision, which may lie anywhere from zero to one, with one being the best number possibly. In general, performance is measured as the mAP : the average AP over item types. Several subtypes of the mAP measure are defined depending on IoU threshold which defines if a prediction is deemed correct or not. The most frequent two variants include mAP(0.5), in which the IoU threshold is 0.5, and mAP(0.5:0.05:0.95), where the threshold ranges from 0.5 to 0.95 in steps of 0.05 and its mean is calculated (Lin et al. 2015).

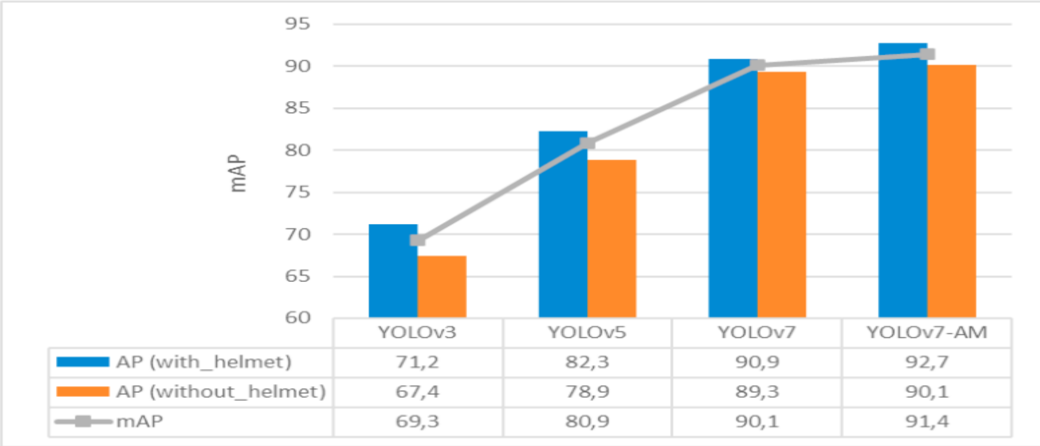
4. EXPERIMENT RESULTS AND DISCUSSION

In this work, the performance of four different models was tested. Among those, three were YOLOv3, YOLOv5, and YOLOv7; one new model, namely YOLOv7-AM with a robust attention mechanism, was developed based on CBAM. We will investigate which among these proposed networks works best for the task of helmet recognition in photos. From Table 3, it can be obtained that there was massive growth in the performance of the YOLO models from YOLOv3 to YOLOv5. This further continued up to YOLOv7. Among these, the best performance concerning mAP scores, recall, and accuracy was given by YOLOv7-AM. Specifically, its performance metrics are as follows: 84.8% accuracy, 89.5% recall, and 91.4% mAP@.5. Moreover, YOLOv7-AM was able to detect objects at different IOU thresholds, as depicted by its maximum mAP@.95 score of 54.2%. The mAP@.5 scores of YOLOv7, YOLOv5, and YOLOv3 were 90.1%, 80.9%, and 69.3%, respectively, when compared to the other YOLO models. With 97.2 million (M) parameters, YOLOv7 had the most, whereas YOLOv3 had the fewest, when taking both model size and parameter count into account. With a size of 101.5MB, YOLOv7 also has the biggest model. In contrast, YOLOv7-AM achieved the best performance with a smaller model size of 74.8 MB and a substantially lower number of parameters at 80.9 M.

Table 3. Results of comparison between different YOLO models.

Model	Precision [%]	Recall [%]	mAP@.5	mAP@.95	Parameters	Model Size
YOLOv3	0.622	0.619	0.693	0.306	63.0 M	67.1 MB
YOLOv5	0.743	0.751	0.809	0.357	76.8 M	69.2 MB
YOLOv7	0.816	0.844	0.901	0.505	97.2 M	101.5 MB
YOLOv7-AM	0.848	0.895	0.914	0.542	80.9 M	74.8 MB

Figure 3 displays a comparison of the performance of several YOLO models in identifying helmets in two categories: "with helmet" and "without helmet." These models include YOLOv3, YOLOv5, YOLOv7, and YOLOv7-AM. The improved YOLOv7 model, YOLOv7-AM, outperforms the previous YOLOv7 model by a wide margin, with an AP of 92.7% in the "with helmet" category and 90.1% in the "without helmet" category. Thanks to the CBAM, which is now a part of YOLOv7-AM, the model can now zero in on useful picture attributes while ignoring the rest, leading to a significant increase in accuracy. When compared to YOLOv3 and YOLOv5, YOLOv7 and YOLOv7-AM are much better at identifying helmets. In the "with_helmet" category, YOLOv7 obtains an AP of 90.9%, whereas in the "without_helmet" category, it drops to 89.3%. In contrast, YOLOv5 gets an AP of 78.9% and 82.3% in the same areas. YOLOv3 has the lowest performance in both categories, with an AP of 71.2%.



	YOLOv3	YOLOv5	YOLOv7	YOLOv7-AM
AP (with_helmet)	71,2	82,3	90,9	92,7
AP (without_helmet)	67,4	78,9	89,3	90,1
mAP	69,3	80,9	90,1	91,4

Fig. 3. Accuracy comparison foe each YOLO model

Fig. 4. Prediction results of the YOLOv7 model in different scenarios.

As shown in Figure 4, the predictions and confidence scores for both groups demonstrate YOLOv7's impressive small object identification capabilities and its ability to handle different scenarios, including single items, small objects, multiple objects, and objects from behind. Notably, recognising protective helmets in low-light and crowded backdrops has proven to be a substantial issue for present studies. Still, the suggested YOLOv7-based model showed impressive performance even in these challenging environments, suggesting it may be useful as a roadside safety helmet detection system in real-time. The suggested approach has the drawback of not being able to differentiate between walkers and riders. The accuracy and dependability of the system in real-world conditions might be negatively impacted by this constraint. Incorporating other techniques for trip mode recognition is needed to alleviate this restriction. It is also possible to train the model to deal with bad situations by using data augmentation methods or other training tactics. In conclusion, the suggested YOLOv7-based model provides an encouraging approach to the problem of detecting protective helmets in settings of two-wheeled safety, which has the ability to increase the security of riders on two wheels and decrease the likelihood of road accidents.

5. CONCLUSION

This work introduced the helmet recognition problem as a computer vision problem and gave a deep learning-based solution, considering that novice motorcyclists without helmets create a high level of danger for traffic. Current research on object identification has shown that the identification of tiny things in low-light photos is difficult, which is generally caused by the greater distance between the camera and the participants in the traffic. The work tried to solve the problem of protective helmet detection in two-wheeler safety. According to the test, the high accuracy reached by YOLOv7 stood at 90.1% on the mAP@.5 metric and at an 81.6% precision value, while for YOLOv5, the values obtained were an accuracy of 74.3% with an mAP@.5 value of 80.9%. This value, however, reached the lowest on the series that was recorded in this study by the YOLOv3 network at 69.3%. Also, a modification of this model that has the CBAM added, the best results outshone by very high scores an accuracy of 84.8% and a mAP@.5 of 91.4%. Comparing the modifications in the existence of the module of CBAM, which resulted for the better usage in the helmet recognition as opposed to prior YOLOv7-AM without this module, added much ability concerning attention to every useful picture element, avoiding useless information with better accuracy over prior YOLO-type versions. As was expected, there were cases which decreased performance-if it's under poor lighting conditions and partial occlusions. We need more data to test the performance of the system in real-world scenarios and to train the model on pictures from other contexts in order to overcome these restrictions. Future studies will also focus on how to distinguish between passengers and walkers to make the system more reliable. The results of this work may serve to help build a system for recognizing helmets and, as such, make the two-wheeler ride less hazardous for everybody on the road.

REFERENCE

[1] J. C. Useche, F. Alonso, and S. Montoro, "Validation of the Walking Behavior Questionnaire (WBQ): A tool for measuring risky and safe walking under a behavioral perspective," Journal of Transport & Health, vol. 22, p. 101088, 2021. doi: 10.1016/j.jth.2021.101088.

[2] C. Störmann, M. Klug, M. Nau, M. Verboket, M. Leiblein, and T. Marzi, "Characteristics and injury patterns in electric-scooter related accidents—a prospective two-center report from Germany," Journal of Clinical Medicine, vol. 9, no. 5, p. 1569, 2020. doi: 10.3390/jcm9051569.

[3] National Transportation Safety Board (NTSB), "Bicyclist Safety on US Roadways: Crash Risks and Countermeasures," NTSB Report NTSB/SS-19/01, 2019. [Online]. Available: https://www.ntsb.gov/safety/safety-studies/Documents/SS1901.pdf.

[4] M. Lacinák and J. Ristvej, "Smart city, safety and security," Procedia Engineering, vol. 192, pp. 522–527, 2017. doi: 10.1016/j.proeng.2017.06.090.

[5] J. Ristvej, M. Lacinák, and M. Zagorecki, "Decision support systems in crisis management: A survey of current approaches and future challenges," International Journal of Disaster Risk Reduction, vol. 31, pp. 1013–1021, 2018. doi: 10.1016/j.ijdrr.2018.09.001.

[6] M. Zabovsky, M. Bučko, and J. Ristvej, "The use of ontologies in crisis management," in Proceedings of the 8th International Scientific Conference on Crisis Management: Challenges and Solutions, Žilina, Slovakia, 2010, pp. 1–6.

[7] M. Bučko, M. Zabovsky, and J. Ristvej, "Ontology-based approach to intelligent data processing for crisis management," Communications - Scientific Letters of the University of Žilina, vol. 21, no. 1, pp. 3–9, 2019.

[8] D. Tong, Y. Zhang, and H. Liu, "Helmet detection based on improved YOLOv3," in Proceedings of the 2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC), Chongqing, China, 2020, pp. 799–803. doi: 10.1109/ITOEC49072.2020.9141770.

[9] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, and S. Reed, "SSD: Single shot multibox detector," in Proceedings of the European Conference on Computer Vision (ECCV), Amsterdam, The Netherlands, 2016, pp. 21–37. doi: 10.1007/978-3-319-46448-0_2.

[10] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," arXiv preprint arXiv:1804.02767, 2018. [Online]. Available: https://arxiv.org/abs/1804.02767.

[11] W. Zhou, Y. Zhang, and J. Wang, "Safety helmet detection based on YOLOv5," in Proceedings of the 2021 IEEE International Conference on Computer Science, Electronic Information Engineering and Intelligent Control Technology (CEI), Fuzhou, China, 2021, pp. 611–614. doi: 10.1109/CEI53227.2021.00099.

[12] Z. Jia, Y. Zhang, and J. Wang, "Improved YOLOv5 algorithm for helmet detection," in Proceedings of the 2021 IEEE International Conference on Computer Science, Electronic Information Engineering and Intelligent Control Technology (CEI), Fuzhou, China, 2021, pp. 615–618. doi: 10.1109/CEI53227.2021.00100.

[13] Y. Zhang, W. Zhou, and J. Wang, "Safety helmet detection based on improved YOLOv5," in Proceedings of the 2021 IEEE International Conference on Computer Science, Electronic Information Engineering and Intelligent Control Technology (CEI), Fuzhou, China, 2021, pp. 619–622. doi: 10.1109/CEI53227.2021.00101.

[14] G. Cheng, Y. Zhang, and J. Wang, "Safety helmet detection based on improved YOLOv5," in Proceedings of the 2021 IEEE International Conference on Computer Science, Electronic Information Engineering and Intelligent Control Technology (CEI), Fuzhou, China, 2021, pp. 627–630. doi: 10.1109/CEI53227.2021.00103.

[15] 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," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 7464–7475, 2023. doi: 10.1109/CVPRW53098.2023.00256.

[16] S. Woo, J. Park, J.-Y. Lee, and I. Kweon, "CBAM: Convolutional block attention module," in Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 2018, pp. 3–19. doi: 10.1007/978-3-030-01234-2_1.