

Pedestrian Crosswalk Detection with Faster R-CNN and YOLOv7

Subjects: [Engineering](#), [Civil](#)

Contributor: Ömer Kaya , Muhammed Yasin Çodur , Enea Mustafaraj

Object detection techniques and object detection using deep neural networks are developing fields. Faster R-CNN (R101-FPN and X101-FPN) and YOLOv7 network models were used in the analytical process of a dataset collected. Based on the detection performance comparison between both models, YOLOv7 accuracy was 98.6%, while the accuracy for Faster R-CNN was 98.29%.

[object detection](#)[image processing](#)[traffic safety](#)[autonomization](#)

1. Introduction

The most important factor causing traffic accidents in a road network is stated as human. This factor comes to the fore, especially in urban transportation ^[1] because the pedestrian–passenger–driver–vehicle–road network equation becomes more complicated. In developing countries, the migration of the population from rural to urban areas has caused crowded city centers. As a result, unplanned cities, low-educated human profiles on traffic safety, and irregular movements inevitably cause negative incidents in traffic flow. This is a big problem, especially for underdeveloped and developing countries. In addition, vehicle–pedestrian conflict in city centers is exceptionally high. If the tendency to obey traffic rules decreases, it is clear that these conflicts will cause material and fatal accidents.

According to the latest Global Status Report of the World Health Organization, approximately 1.3 million people die from traffic accidents every year around the world ^[2]. About 93% of these deaths occur in low- and middle-income countries. Most of these deaths involve young people between the ages of 5 and 29. Although the main causes of these deaths are epidemic diseases, traffic accidents are shown as the eighth main cause. It is inevitable that these death rates will increase with the neglect of road safety and the trend of increasing number of vehicles ^[2]. If this trend stays the same, reaching a modern urban traffic structure will continue to be challenging. In recent years, the integration of autonomous vehicles with traffic has continued at full speed. This means that both road networks and human profiles must be ready. Autonomization of vehicles alone is not beneficial. Other components of road networks also need to be autonomous because warning systems should be given not only to vehicle drivers but also to other components using a road network. Evaluating vehicle–infrastructure–environment communication together is very important in achieving traffic safety.

Moreover, overcoming the difficulties of increasing traffic and micro-mobility solution proposals, which are released as a last-mile delivery, have created new challenges in traffic flow. For example, e-scooter users have many wrong

driving profiles, such as lousy parking, using pedestrian roads, and bad lane switching that will endanger traffic flow. To summarize, it is evident that disabled people, pedestrians, and autonomous vehicles use the same traffic flow. Therefore, solutions to some of the problems this union will bring should be proposed. The United Nations General Assembly has committed to halving global deaths and injuries from road traffic accidents by 2030 [2]. In line with this goal, researchers have to act to find solutions to problematic areas in traffic networks and adaptation to current and future technologies.

2. Automatic Detection of Pedestrian Crosswalk with Faster R-CNN and YOLOv7

Object detection techniques and object detection using deep neural networks are developing fields [3].

Crossing a street is not only difficult for visually impaired individuals but also has some difficulties for other users. Nevertheless, partial solutions to this problem can be offered with effective pedestrian crosswalk detection. Brief information about some of the studies in the literature is given below.

Se [4] first grouped crosswalk lines and proposed a crosswalk detection method using escape point constraint. However, the high computational complexity of this proposed model highly affects the efficiency of use. Huang and Lin [5] identified areas with alternating black and white stripes using bipolarity segmentation and achieved a good performance in standard lane zones. With this approach, they detected similar zebra crossings in [6][7]. Chen et al. [8] proposed a crosswalk detection method based on Sobel edge extraction and Hough Transform. This approach provides a good balance between model accuracy and speed. Cao et al. [9] achieved high accuracy in recognizing disabled roads and pedestrian crosswalks. Moreover, their study was planned to help visually impaired people facilitate orientation and perceive the environment. A lightweight semantic segmentation network, which was used to segment both pedestrian and disabled paths, combined with depthwise separable convolution, was used as a basic module to reduce the number of parameters of the model and increase the speed of semantic partitioning. In addition, an atrous spatial pyramid pooling module was used to improve the network accuracy. Finally, a dataset was collected from a natural environment to verify the effectiveness of the proposed method. As a result, it was observed that the proposed approach gives better or similar results when compared to other methods.

Ma et al. [10] emphasized the difficulties of people with disabilities in obtaining external information and provided some suggestions for overcoming these difficulties. Their main goal was to investigate the effectiveness of tactile paving at pedestrian crosswalks. Data were collected using unmanned aerial vehicles (UAV), such as drones and a three-axis accelerometer. A before–after comparative analysis of the quantitative index results revealed that the tactile coating helps people with visual impairments maintain a straight passageway, avoid directional deviations, reduce transit times, and improve gait patterns and symmetry.

Romić et al. [11] proposed a method based on an image column and row structure analysis to detect pedestrian crossings to facilitate crossing. The technique was also tested with real data input, and it was found that its performance depends on the image quality and resolution of the dataset.

Tian et al. [12] proposed a new system for understanding dynamic crosswalk scenes; detecting key objects, such as pedestrian crosswalks, vehicles, and pedestrians; and defining pedestrian traffic light status. The proposed system that was implemented on a device worn on the head of a person, which transmits scene information to disabled individuals via an audio signal, was proven to be beneficial.

Tümen and Ergen [13] emphasized that crossroads, intersections, and pedestrian crosswalks are essential areas for autonomous vehicles and advanced driver assistance systems because the probability of traffic accidents in these areas is relatively high. In this context, a deep learning-based approach over real images was proposed to provide instant information to drivers and autonomous vehicles. This approach used CNN-based VggNet, AlexNet, and LeNet to classify the data. As a result, high classification accuracy was achieved, and it was shown that the proposed method is a practical structure that can be used in many areas.

Dow et al. [14] designed an image processing-based human recognition system for a pedestrian crossing. The system aims to reduce accidents and accident possibilities and to increase the level of safety. In order to improve the accuracy of pedestrian detection and reduce the system's error rate, a dual-camera mechanism was proposed. The experimental results showed that the obtained prototype system performs well.

Pedestrian crosswalks are an essential component of urban transportation, mainly because these road sections are the regions where pedestrian and vehicle accidents occur more frequently. It is an undeniable fact that developing countries experience many problems in these regions [15]. Many parameters, such as driver–pedestrian behavior profile, mutual respect, low tendency to obey the rules, penalties, and flexibility in the rules, are shown as the reasons for these problems.

In addition to the detection of pedestrian crosswalks, object detection processes are also applied in many fields nowadays. Object detection plays an active role in sectors such as health [16], safety [17], transportation [18][19], and agriculture [20][21]. Many researchers enlarge their dataset to achieve high detection accuracy and experiment with differences in the network structure of their detection model. In fact, studies continue for the detection of two-dimensional materials [22]. A few of the existing studies in different fields are detailed in this section and an in-depth analysis framework is presented.

Researchers such as [23][24][25][26][27] have carried out many studies to detect helmets used for dangerous works, such as construction. SSD, Faster R-CNN, and other YOLO models were used in these studies. Among these studies, [23] had the highest mAP value of 96 percent. Even in recent years, object detection in the agricultural and livestock sector provides great ease of work. Fast counting of animals on a farm and detection of dead animals increase efficiency. Different detection models were used in duck counting by [28]. The YOLOv7 achieved a better detection rate than other models, with a 97.57 percent mAP value.

Detection of weeds and product counting in the agricultural sector provide great opportunities in the marketing part. In a study conducted by [20], the YOLOv7 model reached a mAP value of 61 percent during the detection of weeds

in a field. This value might vary depending on the dataset and the detected object. Since weeds are small objects and a training dataset is difficult to collect, detection accuracy decreases.

The works on driver-assistance systems continue. Using the phone and drinking beverages, which cause distraction while driving, cause accidents and endanger traffic safety. In a study conducted by [29], YOLOv7 was used to detect driving distraction behaviors. Four different detection results were obtained, including danger, drinking, phone usage, and safety. A mAP value of 73.62 percent was obtained. Creating a dataset by obtaining four different driver states from different drivers is a difficult and time-consuming process. It is clear that the accuracy rate will increase if the variety of data is increased.

As the use of object detection progresses as a sector, it offers very useful solution suggestions. An early smoke warning system for fires that cause ecological problems will protect forests. In order to prevent the spread of fires, a study was carried out for a smoke warning system. The dataset was defined with three different distance scales. With YOLOv5x, an accuracy rate of 96.8 percent was achieved. Despite the irregularity of the smoke distribution data, a high accuracy rate was obtained [30].

As stated above, object detection continues to be a solution to some existing problems. Different object detection studies, such as cancer polyp detection [16], internal canthus temperature detection in the elderly [31], construction waste detection [32], in situ sea cucumber detection [33], ship detection in satellite images [34], and citrus orchard detection [35], are available in the literature.

All the studies examined showed that the accuracy rate is directly related to the dataset used. The dimensional properties of a detected object, its state in the image, and its variability reduce the detection accuracy. In order to eliminate this situation, the number and diversity of data should be increased. However, this process is quite difficult. In addition, the most suitable model for the dataset should be selected by using different detection models. In this study, the dimensional properties of the detected object are large. Thus, the accuracy rate is high. In particular, local municipalities should use object detection applications more frequently. The use of these models, most of which are open source, especially in urban transportation applications, would be correct within the scope of smart cities. These systems, which are quick solutions to problems, such as parking areas, wrong parking violation penalties, and red light violations, should be supported by politicians.

References

1. Adanu, E.K.; Jones, S. Effects of Human-Centered Factors on Crash Injury Severities. *J. Adv. Transp.* 2017, 2017, 1208170.
2. World Health Organization (WHO). Global Status Report on Road Safety; World Health Organization (WHO): Geneva, Switzerland, 2018.

3. Zaidi, S.S.A.; Ansari, M.S.; Aslam, A.; Kanwal, N.; Asghar, M.; Lee, B. A survey of modern deep learning based object detection models. *Digit. Signal Process. A Rev. J.* 2022, 126, 103514.
4. Se, S. Zebra-crossing Detection for the Partially Sighted. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No.PR00662)*, Hilton Head, SC, USA, 15 June 2000; IEEE: Piscataway Township, NJ, USA, 2002.
5. Xin, H.; Qian, L. An Improved Method of Zebra Crossing Detection based on Bipolarity. *Sci.-Eng.* 2018, 34, 202–205.
6. Uddin, M.S.; Shioyama, T. Bipolarity and Projective Invariant-Based Zebra-Crossing Detection for the Visually Impaired. In *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)-Workshops*, San Diego, CA, USA, 20–26 June 2005; IEEE: San Diego, CA, USA, 2005.
7. Cheng, R.; Wang, K.; Yang, K.; Long, N.; Hu, W.; Chen, H.; Bai, J.; Liu, D. Crosswalk navigation for people with visual impairments on a wearable device. *J. Electron. Imaging* 2017, 26, 053025.
8. Chen, N.; Hong, F.; Bai, B. Zebra crossing recognition method based on edge feature and Hough transform. *J. Zhejiang Univ. Sci. Technol.* 2019, 6, 476–483.
9. Cao, Z.; Xu, X.; Hu, B.; Zhou, M. Rapid Detection of Blind Roads and Crosswalks by Using a Lightweight Semantic Segmentation Network. *IEEE Trans. Intell. Transp. Syst.* 2021, 22, 6188–6197.
10. Ma, Y.; Gu, X.; Zhang, W.; Hu, S.; Liu, H.; Zhao, J.; Chen, S. Evaluating the effectiveness of crosswalk tactile paving on street-crossing behavior: A field trial study for people with visual impairment. *Accid. Anal. Prev.* 2021, 163, 106420.
11. Romić, K.; Galić, I.; Leventić, H.; Habijan, M. Pedestrian Crosswalk Detection Using a Column and Row Structure Analysis in Assistance Systems for the Visually Impaired. *Acta Polytech. Hung.* 2021, 18, 25–45.
12. Tian, S.; Zheng, M.; Zou, W.; Li, X.; Zhang, L. Dynamic Crosswalk Scene Understanding for the Visually Impaired. *IEEE Trans. Neural Syst. Rehabil. Eng.* 2021, 29, 1478–1486.
13. Tümen, V.; Ergen, B. Intersections and crosswalk detection using deep learning and image processing techniques. *Phys. A Stat. Mech. Its Appl.* 2020, 543, 123510.
14. Dow, C.; Lee, L.; Huy, N.H.; Wang, K. A Human Recognition System for Pedestrian Crosswalk. *Commun. Comput. Inf. Sci.* 2018, 852, 443–447.
15. Alemdar, K.D.; Kaya, Ö.; Çodur, M.Y. A GIS and microsimulation-based MCDA approach for evaluation of pedestrian crossings. *Accid. Anal. Prev.* 2020, 148, 105771.
16. Karaman, A.; Karaboga, D.; Pacal, I.; Akay, B.; Basturk, A.; Nalbantoglu, U.; Coskun, S.; Sahin, O. Hyper-parameter optimization of deep learning architectures using artificial bee colony (ABC)

- algorithm for high performance real-time automatic colorectal cancer (CRC) polyp detection. *Appl. Intell.* 2022, 1–18.
17. Yung, N.D.T.; Wong, W.K.; Juwono, F.H.; Sim, Z.A. Safety Helmet Detection Using Deep Learning: Implementation and Comparative Study Using YOLOv5, YOLOv6, and YOLOv7. In *Proceedings of the International Conference on Green Energy, Computing and Sustainable Technology (GECOST)*, Miri Sarawak, Malaysia, 26–28 October 2022.
 18. Van Der Horst, B.B.; Lindenbergh, R.C.; Puister, S.W.J. Mobile laser scan data for road surface damage detection. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.-ISPRS Arch.* 2019, 42, 1141–1148.
 19. Pandey, A.K.; Palade, V.; Iqbal, R.; Maniak, T.; Karyotis, C.; Akuma, S. Convolution neural networks for pothole detection of critical road infrastructure. *Comput. Electr. Eng.* 2022, 99, 107725.
 20. Dang, F.; Chen, D.; Lu, Y.; Li, Z. YOLOWeeds: A novel benchmark of YOLO object detectors for multi-class weed detection in cotton production systems. *Comput. Electron. Agric.* 2023, 205, 107655.
 21. Wu, D.; Jiang, S.; Zhao, E.; Liu, Y.; Zhu, H.; Wang, W.; Wang, R. Detection of *Camellia oleifera* Fruit in Complex Scenes by Using YOLOv7 and Data Augmentation. *Appl. Sci.* 2022, 12, 11318.
 22. Zenebe, Y.A.; Xiaoyu, L.; Chao, W.; Yi, W.; Endris, H.A.; Fanose, M.N. Towards Automatic 2D Materials Detection Using YOLOv7. In *Proceedings of the 19th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*, Chengdu, China, 16 December 2022.
 23. Kamboj, A.; Powar, N. Safety Helmet Detection in Industrial Environment using Deep Learning. *The 9th International Conference on Information Technology Convergence and Services (ITCSE 2020)*, Zurich, Switzerland, 21–22 November 2020; pp. 197–208.
 24. Huang, L.; Fu, Q.; He, M.; Jiang, D.; Hao, Z. Detection algorithm of safety helmet wearing based on deep learning. *Concurr. Comput. Pract. Exp.* 2021, 33, 1–14.
 25. Li, Y.; Wei, H.; Han, Z.; Huang, J.; Wang, W. Deep Learning-Based Safety Helmet Detection in Engineering Management Based on Convolutional Neural Networks. *Adv. Civ. Eng.* 2020, 2020, 9703560.
 26. Long, X.; Cui, W.; Zheng, Z. Safety helmet wearing detection based on deep learning. In *Proceedings of the 2019 IEEE 3rd Information Technology Networking, Electronic and Automation Control Conference, ITNEC 2019*, Chengdu, China, 15–17 March 2019; pp. 2495–2499.
 27. Chen, K.; Yan, G.; Zhang, M.; Xiao, Z.; Wang, Q. Safety Helmet Detection Based on YOLOv7. *ACM Int. Conf. Proceeding Ser.* 2022, 31, 6–11.

28. Jiang, K.; Xie, T.; Yan, R.; Wen, X.; Li, D.; Jiang, H.; Jiang, N.; Feng, L.; Duan, X.; Wang, J. An Attention Mechanism-Improved YOLOv7 Object Detection Algorithm for Hemp Duck Count Estimation. *Agriculture* 2022, 12, 1659.
29. Liu, S.; Wang, Y.; Yu, Q.; Liu, H.; Peng, Z. CEAM-YOLOv7: Improved YOLOv7 Based on Channel Expansion and Attention Mechanism for Driver Distraction Behavior Detection. *IEEE Access* 2022, 10, 129116–129124.
30. Al-Smadi, Y.; Alauthman, M.; Al-Qerem, A.; Aldweesh, A.; Quaddoura, R.; Aburub, F.; Mansour, K.; Alhmiedat, T. Early Wildfire Smoke Detection Using Different YOLO Models. *Machines* 2023, 11, 246.
31. Ghourabi, M.; Mourad-Chehade, F.; Chkeir, A. Eye Recognition by YOLO for Inner Canthus Temperature Detection in the Elderly Using a Transfer Learning Approach. *Sensors* 2023, 23, 1851.
32. Zhou, Q.; Liu, H.; Qiu, Y.; Zheng, W. Object Detection for Construction Waste Based on an Improved YOLOv5 Model. *Sustainability* 2023, 15, 681.
33. Wang, Y.; Fu, B.; Fu, L.; Xia, C. In Situ Sea Cucumber Detection across Multiple Underwater Scenes Based on Convolutional Neural Networks and Image Enhancements. *Sensors* 2023, 23, 2037.
34. Patel, K.; Bhatt, C.; Mazzeo, P.L. Improved Ship Detection Algorithm from Satellite Images Using YOLOv7 and Graph Neural Network. *Algorithms* 2022, 15, 473.
35. Chen, J.; Liu, H.; Zhang, Y.; Zhang, D.; Ouyang, H.; Chen, X. A Multiscale Lightweight and Efficient Model Based on YOLOv7: Applied to Citrus Orchard. *Plants* 2022, 11, 3260.

Retrieved from <https://encyclopedia.pub/entry/history/show/115939>