Jurnal Ilmiah
Teknologi Sistem
Informasi

Jitsi

Vol. 5 (2024) No.2

https://jurnal-itsi.org/index.php/jitsi; E-ISSN: 2722-4600; ISSN: 2722-4619
DOI: 10.62527/jitsi.5.2.238

YOLO-Based Personal Protective Equipment Monitoring System for Workplace Safety

Emin Guney[#], Husna Altin[*], Ayse Esra Asci[*], Oya Utuk Bayilmis[**], Cuneyt Bayilmis[*]

[#]Department of Computer Engineering, Sakarya University of Applied Sciences, 54050, Sakarya, Turkey
[*]Department of Computer Engineering, Sakarya University, 54050 Sakarya, Turkey
***Sakarya University of Applied Sciences, Sakarya Vocational School, 54290, Sakarya, Turkey*
E-mail: eminguney[at]subu.edu.tr

ABSTRACTS

Occupational health and safety are of paramount importance in industrial environments and various work fields. In this context, tracking personal protective equipment (PPE) is highly necessary. This article investigates the performance and application of deep learning-based object detection models to enhance the accuracy and speed of tracking personal protective equipment for ensuring occupational health and safety. These models detect personal protective equipment in images, enabling monitoring of their correct usage and intervention when necessary. The study aims to minimize damage resulting from accidents through the use of protective equipment and to prevent possible accidents.In our study, a dataset consisting of 2581 images, encompassing different workplace environments and workers, was prepared. This dataset was evaluated for performance using deep learning models. Popular deep-learning models such as YOLO-NAS, YOLOv8, and YOLOv9 were utilized in the comparisons. During the training of the models, the number of epochs was kept consistent for fair comparison. Upon examining the results, it is observed that the YOLO-NAS and YOLOv9 models generally exhibit similar and high performance.

Keywords — *YOLO-NAS, YOLOv8, YOLOv9, PPE, Safety Helmet Detection*

Manuscript received May 24, 2024 revised May 29, 2024. Accepted Jun 01, 2024 Date of publication Jun 30, 2024 International Journal,

CORRESPONDING AUTHOR

Emin Guney
Department of Computer Engineering, Sakarya University of Applied Sciences, 54050, Sakarya, Turkey
Email: eminguney[at]subu.edu.tr

1. INTRODUCTION

Today, the industrial sector is a critical sector constantly growing with the rapid advancement of technology and economic changes. There has been a significant revival in the industrial sector in recent years, especially in developing countries. This field dramatically contributes to the country's sectoral economy, including production, manufacturing, construction, etc. It provides a large amount of employment worldwide in these fields and causes an increase in the demand for workers. This increase in demand, especially in relatively heavier branches of work such as construction, infrastructure, and production industries, also causes work accidents.

While this rapid development in the sector causes similar accidents to existing ones, it also creates new types of accidents when we consider today's work areas. The increase in the use of large and dangerous machines and the number of workers working in high-rise buildings can increase the number of accidents and damage. While safety equipment support cannot be provided to some work sites, even if personal protective equipment is allocated to some work sites, not using them can make accidents inevitable. Personal protective equipment monitoring is of utmost importance to prevent accidents. According to International Labor Organization (ILO) reports, approximately 270 million work accidents and 160 million occupational diseases occur annually worldwide. Nearly two million people die every year due to work accidents or occupational diseases. These deaths constitute 3.9% of the annual number of deaths worldwide, and approximately 15% of the world's population faces such problems. This situation reaches more serious levels in developing countries. For example, while Turkey ranks first among the European Union countries in terms of the number of fatal work accidents, considering the data of 2017, it is seen that the rate of fatal work accidents per hundred thousand is 4.5 times higher than the EU average [1]. The types of injuries in the study conducted by European Work Accident Statistics (ESAW) on fatal and non-fatal work accidents in the European Union in 2021 were discussed. According to the data obtained from the study, it is seen from Fig. 1 that the majority of injuries occur in the head and trunk regions [2]. These results show us how important occupational health and safety is in workplaces. Occupational health and safety are fundamental in national and international regulations.

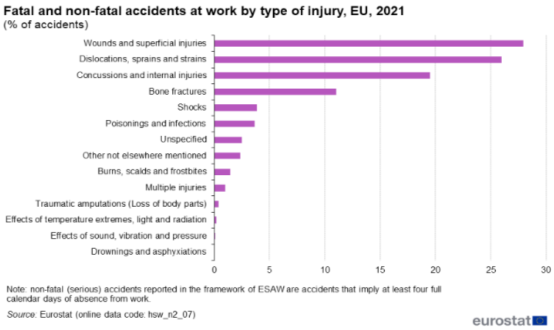

FIGURE 1. Fatal and non-fatal occupational accidents by ESAW injury types

In this context, occupational health is the health precautions that must be taken against risk factors that threaten work-related health in the working and living environment. Occupational safety covers the technical measures required to be taken against accident risks that threaten the employee's life. Within the scope of the employer's obligations in the Occupational Health and Safety Law No. 6331, employers must take all precautions to ensure the health and safety of employees and check whether the occupational health and safety measures taken are complied with. It is responsible for monitoring, auditing, and eliminating non-conformities. [3]

The presented study contributes to the employer's fulfillment of its responsibilities. It aims to prevent possible accidents and minimize the damage that may occur as a result of the accident by ensuring the correct and complete control of the use of personal protective equipment. In this study, since it is tough and time-consuming to track personal protective equipment in areas with many employees manually, we used deep learning-based object detection models to speed up this process. This way, we aimed to ensure faster and more effective equipment tracking. YOLO (You Only Look Once), which stands out among these models, was chosen by the project's objectives due to its ability to quickly detect multiple objects and its real-time nature, which is advantageous in providing instant warning and feedback.

In recent years, many studies based on computer vision (CV) and deep learning (DL) have been carried out to ensure workers' safety in heavy and dangerous lines of work. Deep learning-based methods, such as self-learning, stand out with their strengths. Among these studies, four different data sets used in the study by GİRGİN were selected (SODA, CHV, GDUT-HWD, Harward Hardhat), and the models trained with YOLOv5, YOLOv7, and YOLOv8 algorithms were compared on these chosen data sets. As a result of this comparison, it was seen that the YOLOv8 architecture generally gave better results than the YOLOv5 and YOLOv7 architectures. At the same time, the YOLOv7 model was more successful in high resolution and fast detection [4]. In another article published in 2022, a deep learning-based solution is proposed for automatically detecting helmets. A data set containing 5000 hard hat images was used in the study. The dataset is divided into 60%:20%:20% for training, testing, and validation. YOLOv5x architecture achieved the best (mAP) rate (92.44%) compared to YOLOv3 and YOLOv4 [5]. In another study, a convolutional neural network (CNN)-based deep learning method is proposed to control construction workers' wearing of helmets automatically. This method was applied using the transfer learning technique on models such as YOLO V4, V5, and Faster R-CNN, and the results were examined. In the study, it was determined that the models obtained by applying transfer learning had higher accuracy and success rates [6].In the study published in 2021, eight deep-learning detectors based on YOLO architectures were trained and evaluated. The CHV dataset containing 1330 images from real construction sites was created. YOLO v5x has the best mAP (86.55%), while YOLO v5s offers the fastest processing speed (52 FPS). YOLO v5 models perform superior to other deep learning approaches [7]. A paper published in 2020 presented research on deep learning and computer vision to detect construction workers' safety compliance with PPE in real-time. The dataset, consisting of 2509 images, uses the CNN model with YOLOv3 deep learning. The model achieved an F1 score of 96% on the test data set [8].N.K. Anushkannan and his colleagues aimed to compare deep learning models through different metrics using a data set of 5000 images taken from the Kaggle website. In this study, where R-CNN and YOLOv3 models were compared based on mAP, recall, and precision values, the YOLOv3 model achieved a mAP rate of 97.12% and demonstrated a better success than the R-CNN model with which it was compared, proving its effectiveness in detecting helmet use [9].

Zhou et al. addressed the need for safety helmet detection in industrial environments and presented a digital system based on YOLOv5. They worked on four different YOLOv5 models with the 6045 datasets they collected. The experimental results stated that the average detection speed of YOLOv5s is 110 FPS, which fully meets the real-time detection requirements. However, they said that when the pre-training weights of the trainable target detector were used, the mAP of YOLOv5x reached 94.7%, which proved the effectiveness of helmet detection based on YOLOv5 [10]. Ferdous and Ahsan developed an automatic detection system that detects various PPEs to prevent accidents in the industry in 2022. Comparing the detection algorithm of the YoloX-m model, it was concluded that YoloX-S performed better than the YoloX-I and YoloX-X models and achieved the highest mAP value of 89.84% [11]. Armstrong Aboah and his team aimed to develop a robust model with less labeling using computer vision techniques to identify helmet use violations in traffic and used the YOLOv8 model to detect helmet violations in real-time from video frames together with the data processing technique called few-shot data sampling. Although the mAP of YOLOv8 was measured as 0.5861 and the detection speed was measured as 95 FPS, it was seen that it surpassed all other models with these measurements [12]. Chen and his colleagues developed a system called YOLO-Face based on YOLOv3, which was based on the problem of detecting different facial scales. In comprehensive experiments conducted on the WIDER FACE dataset, while YOLOv3 was found to be more successful than the YOLOv2 model, it was more successful than the YOLO-face model developed only in some areas and fell behind the YOLO-face model in terms of general performance [13]. Nipun D. et al. presented three deep learning models built on YOLOv3 architecture to verify workers' PPE usage status. Although the first approach is the fastest among these three approaches, it has been observed to process 13 frames per second and have a mAP value of 63.1%.

On the other hand, in the general evaluation, the second approach, with a mAP value of 72.3%, achieved the best performance and was able to process 11 frames per second. The closest alternative in terms of performance was the third approach, with a mAP value of 67.93% [14]. Gallo, smart systems were used for real-time PPE detection in industrial production environments. The Yolov4 object detection algorithm was used to capture visual recordings taken by the embedded system. As a result of the tests for the appropriate Yolov4 version, Yolov4 - Tiny v4.7 reached FPS and performed six times faster than Yolov4 [15]. In this study by Zhou et al., helmet detection was made with a yolov3-based deep learning model called AT-YOLO. As a result of the mAP value measured as 96.5%, AT-YOLO achieved a more successful outcome of 9.34% compared to Yolov3 [16]. In this study by Zeng et al., helmet tracking was carried out automatically to increase helmet use and prevent workers' accidental injuries. By developing Yolov4, the performance of the model they created has been increased to detect small targets. The new algorithm achieved an accuracy rate of 93.37%, which was 3.15% better than the existing Yolov4 [17]. In Yung et al.'s study, Yolov5s, Yolov6s, and Yolov7 models were trained, and their results were evaluated. According to the results, while all three models could not distinguish an ordinary hat from a hard hat, the Yolov7 model showed a better performance than Yolov5s and Yolov6s, albeit with low reliability, in monitoring the use of hard hats by dark-skinned people [18].

2. RESEARCH METHODOLOGY

2.1. Datasets

In the preparation part of the dataset, we first collected various images from web browsers containing employees from open and closed work sites. In this section, we ensured that the equipment worn by the employees included a variety of colors. Fig. 2 shows sample images from our dataset. We collected images of helmets in 5 different colors: yellow, orange, white, red, and blue, and images of protective vests in 4 different colors: yellow, orange, red, and blue. We carried out the labeling process consisting of 9 classes according to the colors and types of the obtained images in the Roboflow environment. The images were examined, and the data set underwent a rigorous labeling process to achieve a more effective and accurate performance using the object detection algorithm. Items such as helmets and vests in each image were manually labeled so the model could correctly recognize these objects. This labeling process increased the reliability of the dataset and allowed the object detection algorithm to work successfully in real-world conditions.

FIGURE 2. Sample images from our dataset

The dataset containing 2581 images was divided into training, validation, and testing via Roboflow. 2202 visuals were trained in the Google Colab environment, including 223 for training and 156 for testing. The object detection process was carried out using a deep learning-based approach. We used YOLO series to recognize work safety equipment such as helmets and vests. The YOLO model stands out as the main component of our project. Trained with various weights and hyperparameter settings, this model can quickly and accurately identify objects in images

2.2. YOLO Algorithms

YOLO series was used as the object detection algorithm [15]. YOLO, which is frequently preferred in object detection tasks, provides effective results as a deep learning model. It divides the image into small squares and performs object detection by analyzing these frames simultaneously [16]. With its high sensitivity, YOLO provides successful results for situations that require detecting a large number of objects simultaneously [17]. YOLOv8 is a state-of-the-art object detection model developed by Ultralytics to improve object detection accuracy and speed by leveraging previous versions' strengths. YOLOv8 significantly increases small objects' detection sensitivity using a more extensive and deeper neural network architecture. It also combines features from different neural network layers with a new feature fusion technique called Path Aggregation Network (PAN), which improves detection accuracy. Evaluations on the COCO dataset show that YOLOv8 provides a significant increase in average precision and faster rendering speed on a single GPU. As a result, YOLOv8 offers a substantial improvement over previous versions in accuracy and speed.YOLO-NAS is designed to detect small objects, increase localization accuracy, and improve performance per calculation ratio. Additionally, it has an open-source architecture that can be used for research purposes. AutoNAC, Deci's proprietary NAS technology, was used for automatic architecture design. The hybrid quantization method was used to balance latency and accuracy instead of standard quantization to quantize specific model parts selectively. YOLOv9, introduced by Wang and his team, is a new computer vision model architecture. Once released, the source code is available so anyone can train their YOLOv9 models. According to the research team, the YOLOv9 model architecture achieves a higher mAP value than the existing popular YOLO models (YOLOv8, YOLOv7, and YOLOv5) in the MS COCO dataset benchmarks.

FIGURE 3. General flow chart

3. RESULTS AND DISCUSSION

3.1. YOLO-NAS

We conducted training with the YOLO-NAS model using the dataset generated through Roboflow in the Google Colab environment. In order to achieve high success rates in this training, we set the number of epochs to 200. After approximately 14 hours of training, the confusion matrix shown in Fig. 3 and Fig. 4 was obtained.

According to the results of the training conducted with the current dataset, the YOLO-NAS model achieved an f1 score of 0.9156, a mAP (mean Average Precision) value of 0.8949, a Precision value of 0.8949, and a Recall value of 0.9594.

3.2. *YOLOv8*

We trained the dataset we created with the YOLOv8 model, one of the latest versions of YOLO. In order to accurately compare with other models, the number of epochs was set to 200. During the approximately 4-hour training, the model achieved an f1 score of 0.85, a precision value of 0.965, an mAP (mean Average Precision) value of 0.862, and a recall value of 0.87. The performance outputs of the YOLOv8 model obtained after training are provided in Fig 5. The confusion matrix of the model is included in Fig 6.

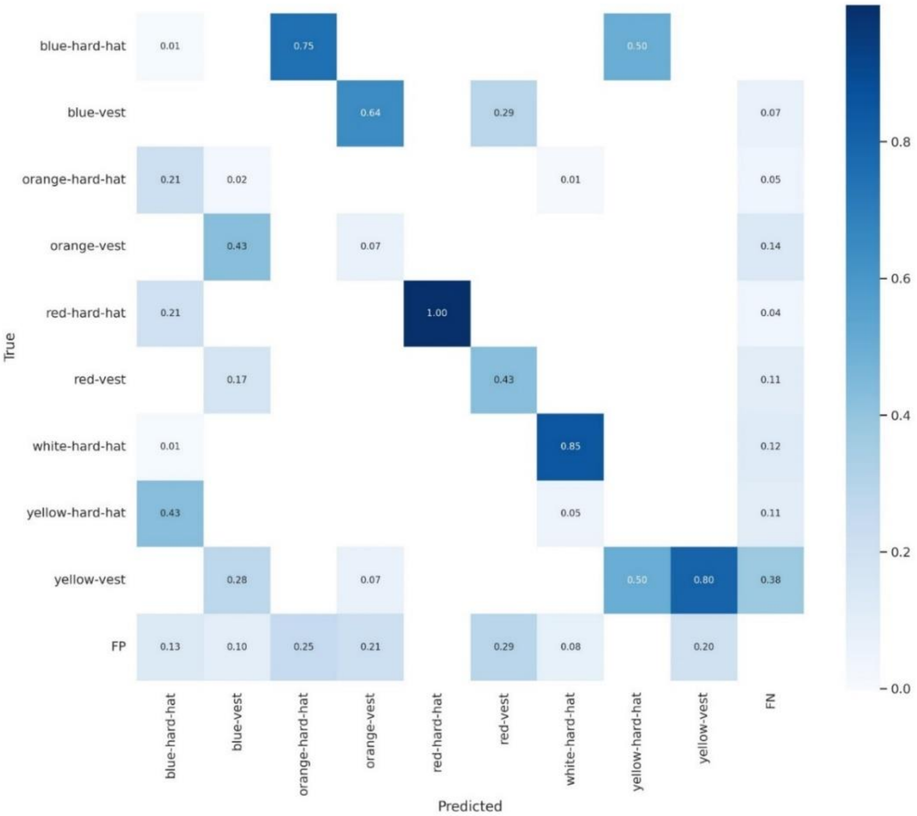

FIGURE 4. YOLO-NAS confusion matrix

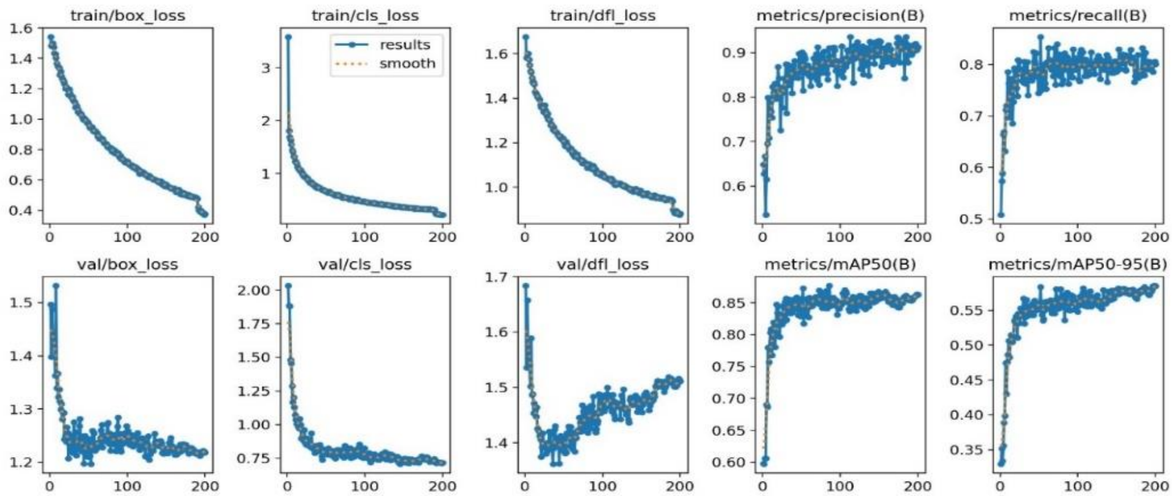

FIGURE 5. Performance outputs of YOLOv8 model

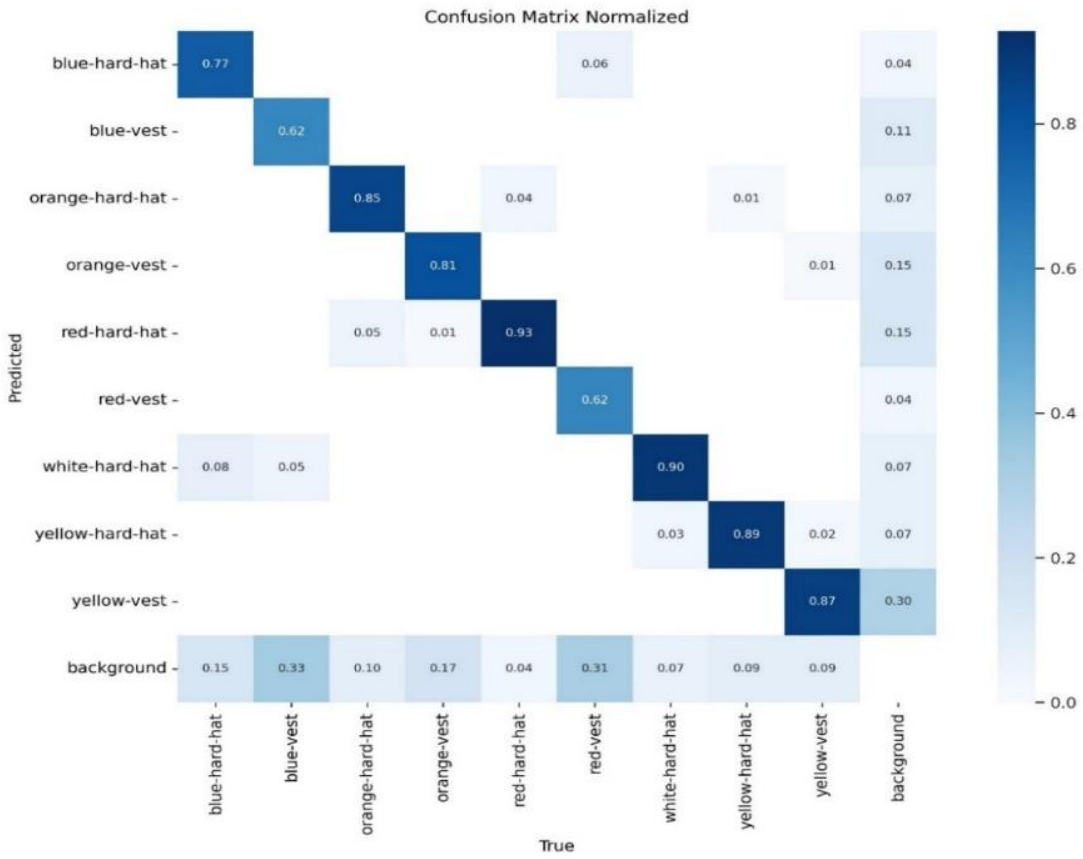

FIGURE 6. YOLOv8 confusion matrix

3.3. YOLOv9

Finally, the dataset was trained with the YOLOv9 model. In this training, the number of epochs was kept constant at 200 to enable comparison of results obtained from different models. After approximately 7 hours of training, the YOLOv9 model achieved an F1 score of 0.85, an mAP of 0.883, a Precision of 0.995, and a Recall of 0.89 on the current dataset. The performance outputs of the YOLOv9 model obtained after training are presented in Fig 7. The confusion matrix corresponding to the model is depicted in Fig8.

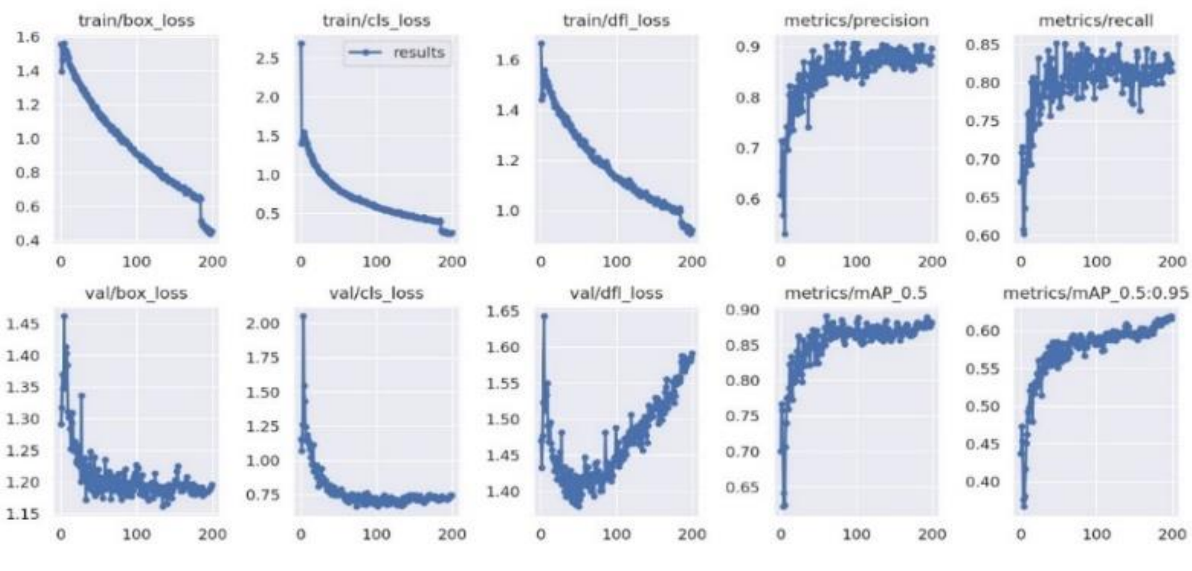

FIGURE 7. Performance outputs of YOLOv9 model

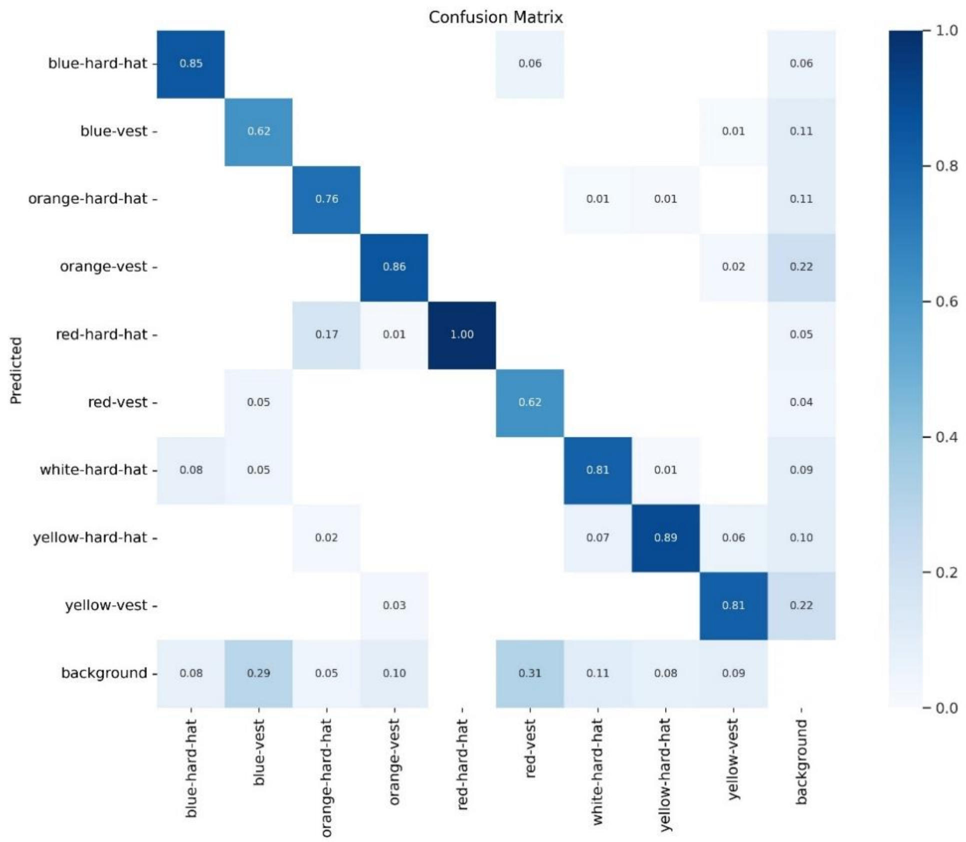

FIGURE 8. YOLOv9 confusion matrix

4. CONCLUSIONS

In this study, the dataset we curated was tested on the deep learning models YOLO-NAS, YOLOv8, and YOLOv9. The dataset, comprising 2581 images, was trained for 200 epochs to ensure a fair comparison of the models

TABLE 1. Comparison of model accuracy rates

Model name	Epoch time (h)	F1 Score	Precision	Recall	mAP
YOLO-NAS	13.8	91,56%	89,49%	95.94%	89.49%
YOLOv8	4.17	85%	96.5%	87%	86.2%
YOLOv9	7.4	85%	99.5%	89%	88.3%

The data for the three different models obtained after training are summarized in Table 1. Based on the obtained data, in the comparison made, it was observed that the training of the YOLO-NAS model took much longer compared to the other models, but it achieved higher recall, F1 and mAP values than the other models. When the data for the YOLOv8 model was examined, it was observed that the training time of the model was shorter than the other models, while its results were close to those of YOLOv9. When compared to the other models, the YOLOv9 model had the highest precision value at 99.5%, with recall and mAP values close to YOLO-NAS but slightly lower. Graphs illustrating the results for the YOLO-NAS, YOLOv8, and YOLOv9 models together are depicted in Fig 9.

The use of personal protective equipment (PPE) by workers in hazardous workplaces can greatly protect them from accidents. According to statistics, tens of thousands of workers are seriously injured in hazardous workplaces every year, facing lifelong difficulties due to the damages caused by these injuries. Therefore, it is crucial to inspect workers' protective equipment. For this inspection to be carried out properly, a correct and fast system is required. Based on these requirements, in this study, a dataset containing 2581 images was created. Of these images, 2202 were used for training, 156 for testing, and 223 for validation purposes. This dataset was tested comparatively on three different versions of YOLO (YOLO-NAS, YOLOv8, YOLOv9) with different parameters in terms of accuracy, speed, and other metrics. Upon examining the results, it is observed that the YOLO-NAS and YOLOv9 models generally exhibit similar performance. However, upon closer inspection, it can be observed that the YOLO-NAS model achieves high F1 Recall and MAP values, whereas YOLOv9 offers a higher precision value. The YOLOv8 model, on the other hand, generally shows slightly lower performance compared to the other two models. Overall, these results underscore the strong performance of YOLO models in object detection. Enhanced details and more precise results provide an important roadmap for future object detection applications. In the future stages, to increase the effectiveness of the models, the number of images in the dataset can be increased. Additionally, to provide workers with greater safety, the dataset can be expanded to include the detection of other protective equipment such as gloves and goggles, in addition to equipment like vests and helmets

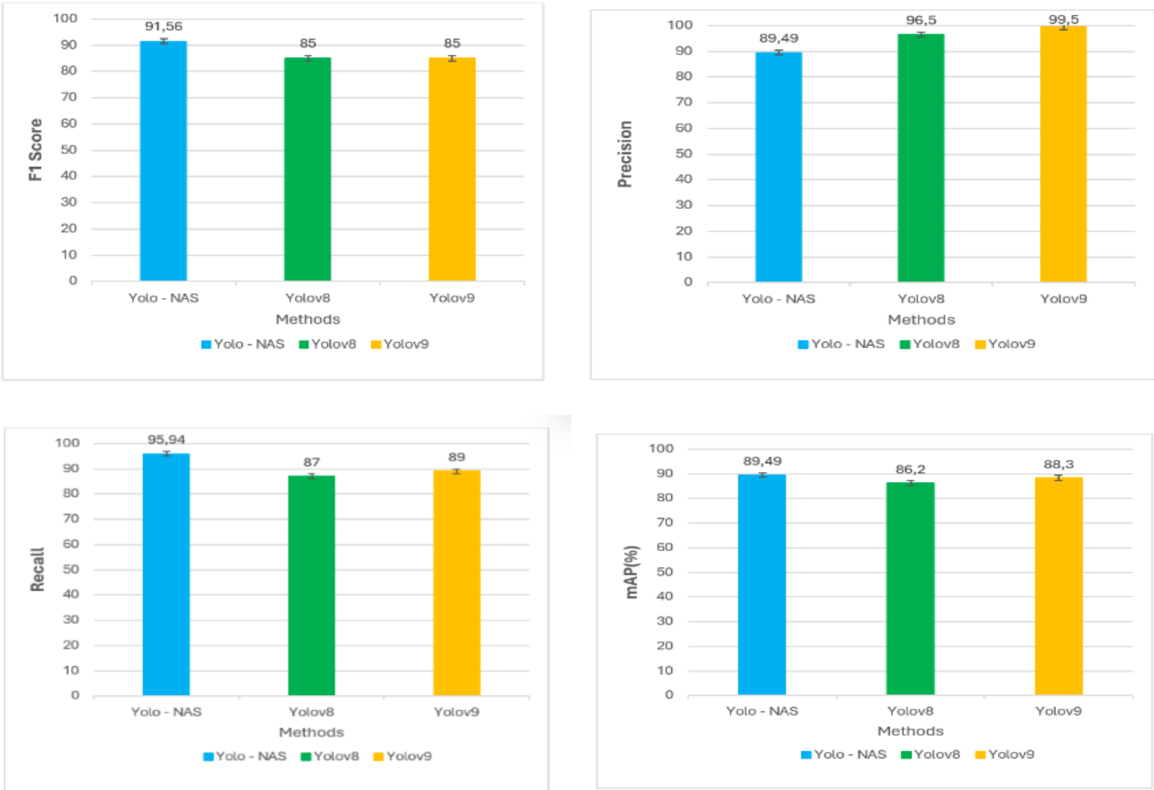

FIGURE 9. YOLO-series speed comparison curves

ACKNOWLEDGMENTS

The authors would like to thank the anonymous reviewers for their useful suggestions.

DISCLOSURE OF INTERESTS

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

REFERENSI

[1] International Labour Organization. (n.d.). "Events and Meetings - Safety and Health at Work". Retrieved March 6, 2024. [Online]. Available: https://www.ilo.org/global/topics/safety-and-health-at-work/events-training/events-meetings/WCMS_111296/lang--en/index.htm.

[2] Eurostat. (n.d.). "Fatal and non-fatal accidents at work by type of injury." Statistics Explained. European Commission. Accessed: [Insert access date]. URL: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=File:Fatal_and_non-fatal_accidents_at_work_by_type_of_injury.png

[3] Kilkiş, İ.: İş Sağlığı ve Güvenliği [Occupational Health and Safety]. (4th ed.).: Dora Yayıncılık, Bursa, Turkey (2022)

[4] Girgin, İ.: Makine Öğrenmesi ile İş Sağlığı ve Güvenliği Uygulamaları (Applications of Machine Learning in Occupational Health and Safety] (master's thesis). Ankara Üniversitesi [Ankara Universitiy] (2023)

[5] Hayat, A., Morgado-Dias, F.: Deep learning-based automatic safety helmet detection system for construction safety. Applied Sciences, 12.16: 8268, (2022)

[6] Türkdamar, M., Taşyürek, M., Öztürk C.: Transfer öğrenmeli ve transfer öğrenmesiz derin ağlar ile inşaat alanında kask tespiti[Helmet detection in the construction site with deep networks with and without transfer learning].Niğde Ömer Halisdemir Üniversitesi Mühendislik Bilimleri Dergisi[Niğde Ömer Halisdemir University Journal of Engineering Sciences], 12.1: 39-51, (2023)

[7] Wang, Z., Wu, Y., Yang, L., Thirunavukarasu, A., Evison, C., Zhao, Y.: Fast personal protective equipment detection for real construction sites using deep learning approaches. Sensors, 21.10: 3478, (2021)

[8] Delhi, V., Sankarlal R., Thomas, A.: Detection of personal protective equipment (PPE) compliance on construction site using computer vision based deep learning techniques. Frontiers in Built Environment. 6: 136, (2020)

[9] Anushkannan, N., Kumbhar, V., Maddila, S., Kolli, C., Vidhya, B., Vidhya, R.: YOLO algorithm for helmet detection in industries for safety purpose. In 3rd International Conference on Smart Electronics and Communication (ICOSEC) 2022, pp. 225-230. IEEE (2022)

[10] Zhou, F., Zhao, H., Nie, Z.: Safety Helmet Based on YOLOv5. In International Conference on Power Electronics, Computer Applications (ICPECA) 2021, pp. 6-11. IEEE (2021)

[11] Ferdous, M., Ahsan, SMM.: PPE detector: a YOLO-based architecture to detect personal protective equipment (PPE) for construction sites. PeerJ Computer Science 8(999), (2022)

[12] Aboah, A., Wang, B., Bagci, U., Adu-Gyamfi, Y.: Real-time Multi-Class Helmet Violation Detection Using Few-Shot Data Sampling Technique and YOLOv8. In Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) 2023, pp. 5350-5358. IEEE/CVF (2023)

[13] Chen, W., Huang, H., Peng, S.: YOLO-face: a real-time face detector. Vis Comput 37, 805–813 (2021)

[14] Nath,N., Behzadan, A., Paal,S.: Deep learning for site safety: Real-time detection of personal protective equipment. Automation in Construction 112, 103085(2020)

[15] Gallo, G., Di Rienzo, F., Ducange, P., Ferrari, V., Tognetti, A., Vallati, C.: A Smart System for Personal Protective Equipment Detection in Industrial Environments Based on Deep Learning. In: International Conference on Smart Computing (SMARTCOMP 2021, pp. 222-227. IEEE (2021)

[16] Zhou, Q., Qin, J., Xiang, X.: Algorithm of Helmet Wearing Detection Based on AT-YOLO Deep Mode. Computers, Materials & Continua 69(1), 159–174 (2021)

[17] Zeng, L., Duan, X., Pan, Y.: Research on the algorithm of helmet-wearing detection based on the optimized yolov4. Vis Comput 39, 2165–2175 (2023)

[18] Yung, Yung, N., Juwono, F., Wong, W., Sim,Z.: Safety Helmet Detection Using Deep Learning: Implementation and Comparative Study Using YOLOv5, YOLOv6, and YOLOv7. In: International Conference on Green Energy, Computing and Sustainable Technology (GECOST) 2022, pp. 164–169., IEEE, (2023)

[19] Yung, Yung, N., Juwono, F., Wong, W., Sim,Z.: Safety Helmet Detection Using Deep Learning: Implementation and Comparative Study Using YOLOv5, YOLOv6, and YOLOv7. In: International Conference on Green Energy, Computing and Sustainable Technology (GECOST) 2022, pp. 164–169., IEEE, (2023)

[20] Guney, E., Sahin, I. H., Cakar, S., Atmaca, O., Erol, E., Doganli, M., and Bayilmis, C.: Electric shore-to-ship charging socket detection using image processing and YOLO. In 2022 International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), pp. 1069-1073. (2022)

[21] Güney, E., Bayılmış, C., Çakar, S., Erol, E., & Atmaca, Ö.: Autonomous control of shore robotic charging systems based on computer vision. Expert Systems with Applications, pp. 122116. (2023)

[22] Güney, E., & Bayılmış, C.: An implementation of traffic signs and road objects detection using faster R-CNN. Sakarya University Journal of Computer and Information Sciences, 5(2), pp. 216-224. (2022)