Comparative analysis of convolutional network training for enhanced detection of PPE in automotive repair services

Álvaro Severo Marcus. Author[1], Luiz Alberto Pinto. Author[1], Gustavo Maia de Almeida. Author[1], Amanda Santos Silva. Author[1]

[1]*Instituto Federal do Espírito Santo*
Serra, Brasil
Av. dos Sábias 330, 29166-630, Serra, Brasil
alvaro.marcus@gmail.com, luiz.pt@ifes.edu.br
gmaia@ifes.edu.br, amandasantos010403@gmail.com

Abstract. This study explores enhancements in the detection of Personal Protective Equipment (PPE) in the automotive service industry through improved computer vision techniques. By comparing initial training using 600 images with subsequent training using 1,200 images, the study evaluates the impact on real-time detection accuracy. Utilizing a YOLOv8 convolutional neural network, approximately 2,000 high-resolution images from security cameras were analyzed to ensure compliance with safety gear requirements. The results demonstrate significant improvements in detecting smaller items like safety glasses and gloves, highlighting the potential of advanced image recognition technologies to enhance workplace safety and operational efficiency. This research sets the stage for new advances in smart security practices in industrial environments.
Keywords: Computer Vision, Personal Protective, Equipment, Automotive Industry.

1 Introduction

This research aims to explore a specific approach in computer vision, focusing on real-time detection of PPE usage in automotive service settings using a YOLOV8, "You Only Look Once" from Ultralytics [1] convolutional neural network model. The YOLOV8 model has demonstrated good results in PPE detection, as shown by Barlybayev et al. [2], where results of pretrained models are compared, highlighting the concern to reduce accidents as demonstrated by Bhana et al. [3], where computer vision is considered an intelligent option to help mitigate accidents caused by inadequate or lack of PPE. Inspired by successful implementations as demonstrated by Ferdous et al. [4], the model will be simplified to facilitate its embedding in a portable platform, a strategy similarly applied by Gallo et al. [5]. By enhancing the detection of PPE compliance, this study seeks to minimize administrative actions due to accidents and improve overall workplace safety and quality.

2 PPE Classification Process

Workers who consistently use PPE have a lower risk of accidents, as demonstrated by Sadat et al. [6]. Safety glasses are crucial for protecting your eyes from flying debris, while gloves offer protection from cuts and maintain the dexterity and grip needed to handle delicate materials. Furthermore, there are studies that seek to combine intelligent solutions in the area of worker protection, collecting information in real time, as observed by Rasouli et al. [7]. Figure 1 shows a visual example of the target objects in this study, including gloves, safety glasses, and a uniform that is a dark blue shirt. The uniform has a small logo, which is not highlighted in the study to protect company information.

Figure 1. PPE used by workers

Detecting workers in uniform not only ensures compliance with these policies but also helps control access to specific activities, ensuring that only authorized personnel engage in certain tasks, as highlighted by Karlsson [8].

3 YOLO Configuration

The pretrained model utilized is 'yolov8n', which demands less processing power. Previous studies have successfully employed this model, as evidenced by Shi et al. [9] where the same pretrained model was chosen. This indicates the potential to create a practical model suitable for deployment on portable solutions, as demonstrated by Khan et al. [10], where the approach has a different objective from PPE detection but proves the technique's wide applicability, as seen in other real-time camera usage scenarios with certain complexities presented by Karna et al. [11].

4 Methodology

The study was developed in the following stages: image acquisition, definition of PPE and their classes, image labeling, dataset division for training and validation, and model training. The methodology began with the acquisition of approximately 2,000 images from security cameras, as shown in Fig. 2, installed in an automotive repair shop. These cameras captured high-resolution images of 1280x720 pixels, providing a satisfactory level of detail suitable for human perception. The setup of these images included a controlled environment with regular artificial lighting, ensuring visibility throughout the operating hours. Additionally, the scene often involved interactions between employees and potential customers waiting for vehicle repairs, adding complexity to the visual data due to the presence of multiple subjects and activities.

Figure 2. High-definition camera used in this study

Given these conditions, precise identification of relevant objects—PPE in this case—was crucial to minimize false positives and enhance the effectiveness of the detection process. These images were then processed using the YOLOv8 algorithm, an advanced iteration of the "You Only Look Once" series, chosen for its proven efficacy in real-time object detection as detailed by Delhi et al. [12]. YOLOv8's capabilities in accurately classifying, detecting, and segmenting objects within images made it an ideal choice for this study as illustrated in Fig. 3, ensuring robust and efficient analysis of the complex visual data presented in the automotive service environment. The training stage was carried out in two parts: an initial training with a dataset of 600 images and a second training with 1,200 images, including those from the first training.

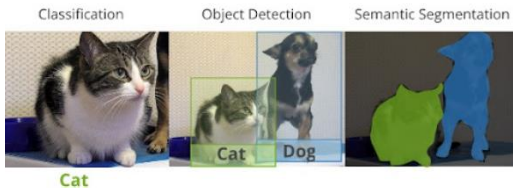

Figure 3. YOLOv8 applications

Following image acquisition, the next step involved defining the objects of interest, which had been previously identified as gloves, safety glasses, and uniforms. The classes important for the study were defined as follows: “uniform” for an employee wearing the uniform correctly, “glasson” for an employee correctly using safety glasses, “glassoff” for an employee with safety glasses placed on the head, indicating incorrect usage, “gloveson” for an employee correctly using gloves, “glovesoff” for an employee not using gloves, indicating non-compliance and “noglass” for an employee neither wearing safety glasses correctly nor on the head.

The dataset in the first training was then organized into training and validation sets, distributed at 55% and 45% respectively. In the second training, a more efficient distribution was achieved with 80% for training and 20% for validation. In the coding and parameterization stage, the Python programming language was used along with the Ultralytics libraries. The model was trained for 100 epochs with a batch size of 10. The entire labeling process was conducted without making the dataset publicly available due to the supplier's confidentiality requirements.

5 Results and Discussion

The preliminary results from the training of the dataset using the YOLOv8 model are illustrated in Tab. 1 and the results from second training in Tab. 2. These results highlight the relative detection frequencies of different classes of Personal Protective Equipment (PPE) in the dataset. The importance of a greater range of data in a dataset is evident, so that better results are possible, as seen in Tab. 2.

Table 1. Quantities per class detected in the first training.

Instances	Quantity
noglass	10
glasson	0
glassoff	78
gloveson	48
glovesoff	130
uniform	155

Table 2. Quantities per class detected in the second training.

Instances	Quantity
noglass	100
glasson	500
glassoff	390
gloveson	950
glovesoff	575
uniform	1480

Figure 4 shows a comparison of the two training sessions. It illustrates the metrics of Mean Average Precision (mAP) for the mAP50(B) metric across different epochs of training. The x-axis represents the number of epochs (0 to 100) and the y-axis indicates the mAP values. The mAP metric combines precision and recall providing an overall performance score, with mAP50(B) measuring the mean average precision at a specific intersection over union (IoU) threshold. It can be observed that the mAP50 reaches rates of 0.9 in the second training, while in the first training, it was close to 0.7 in the best cases. Precision is another important metric for model evaluation. In the second training, precision stabilized at around 0.9 after an upward curve, while in the first training, precision stabilized at around 0.6 after a downward curve. This confirms the hypothesis that a larger and more variable dataset tends to achieve better results in computer vision models.

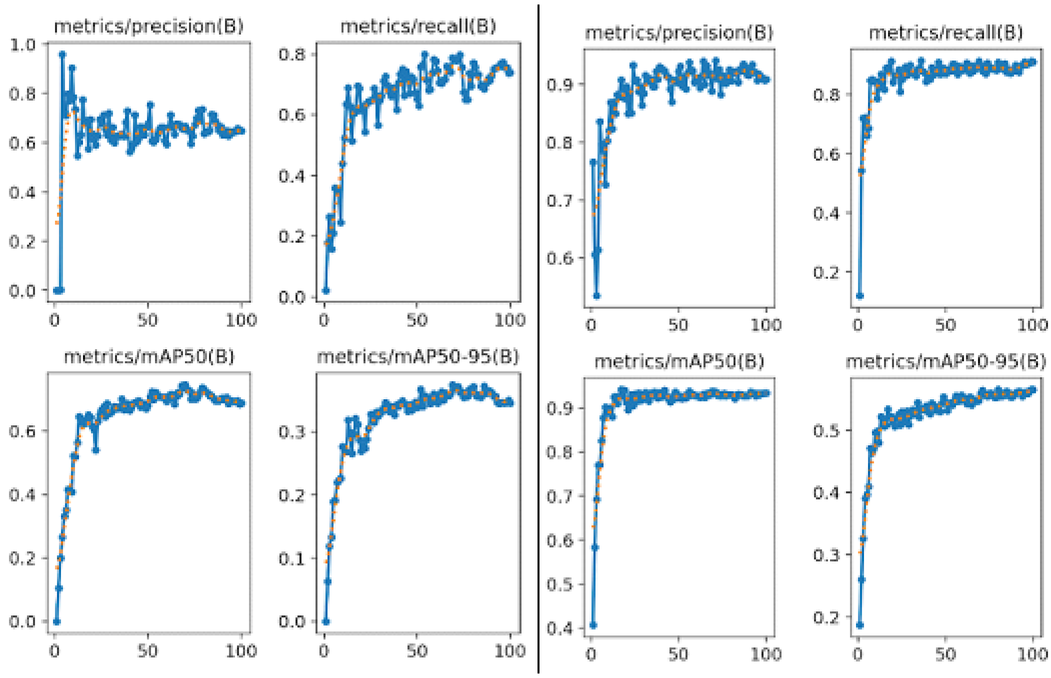

Figure 4. On the left, the results of the first training. On the right, the results of the second training.

Figure 5 is the clearest evidence of the effective detection achieved with the model implemented in the second training. The detections were made with good model confidence indicators, all above 0.6. The confidence indicator, obtained through YOLOv8, represents the model's certainty in its predictions. It quantifies the probability that a detected object belongs to a specific class and is often used to filter out low-confidence detections, thereby improving the overall accuracy of the detection process.

Figure 5. Detections obtained by applying the model obtained in the second training.

Safety glasses and gloves pose greater detection challenges and are encountered less frequently compared to uniforms. These items naturally demand more attention from an auditor when reviewing security camera footage, which means the results are in line with practical expectations. To improve detection of these smaller PPE items, it is necessary to not only increase the size of the dataset, as demonstrated by Lo et al. [13], but also integrate a dedicated testing phase, in addition to training and validation. This testing phase is crucial to evaluate the effectiveness of the trained network in the face of possible variations in the work environment, ensuring that the security and quality of the information obtained are not compromised. This rigor is essential as the system can support workplace safety inspections by offering a reliable tool for continuous monitoring.

Using techniques such as data augmentation – which involves manipulating images to create variations of existing images – can further improve the training process. This practice not only expands the volume of data, but also introduces a greater diversity of scenarios to the learning model, potentially increasing its ability to accurately identify EPI under different conditions. Data augmentation is an effective approach when combined with other techniques and can provide high accuracy rates, as reported by Riaz et al. [14].

5 Conclusions

This study demonstrates the effectiveness of YOLOv8 in environments with variable elements and diverse testing settings. The variability of the dataset is shown to be a key factor of utmost importance in increasing the effectiveness of the detection and classification performed by the trained model. The elements that are most challenging in terms of detection can be successfully classified, provided that the dataset offers a satisfactory number of variations in angles and sizes for the model to achieve effective prediction. Therefore, for future tests, it is understood that increasing the amount of information in the dataset will result in a higher probability of obtaining a more predictive model in terms of its accuracy.

Acknowledgements. It is important to acknowledge the company that provided the footage for the advancement of the proposed study. The entire image labeling process was conducted in a way that preserved the integrity and privacy of all the content in the dataset, without causing any harm to the material's provider.

References

[1] Ultralytics, "YOLO by Ultralytics," [Online]. Available: https://github.com/ultralytics/ultralytics. [Accessed: Mar 25, 2024].
[2] A. Barlybayev, N. Amangeldy, B. Kurmetbek, I. Krak, B. Razakhova, N. Tursynova, and R. Turebayeva, "Personal protective equipment detection using YOLOv8 architecture on object detection benchmark datasets: a comparative study," Cogent Engineering, vol. 11, no. 1, 2024.
[3] R. Bhana, H. Mahmoud, and M. Idrissi, "Smart industrial safety using computer vision," in 2023 28th International Conference on Automation and Computing (ICAC), Birmingham, United Kingdom, 2023.
[4] Md Ferdous and S.M.M. Ahsan, "PPE detector: A YOLO-based architecture to detect personal protective equipment (PPE) for construction sites," PeerJ Computer Science, vol. 8, p. e999, 2022.
[5] G. Gallo et al., "A smart system for personal protective equipment detection in industrial environments based on deep learning at the edge," IEEE Access, vol. 10, pp. 110862–110878, 2022.
[6] L. A. Sadat, I. 'A. Najie, and A. R. Afrizal, "The effect of compliance with wearing appropriate protective personal equipment with exposure risk on the incidence of occupational accidents," Journal Health Sains, vol. 4, no. 12, 2023.
[7] S. Rasouli, Y. Alipouri, and S. Chamanzad, "Smart personal protective equipment (PPE) for construction safety: A literature review," Safety Science, vol. 170, 2024, Art. no. 106368. [Online]. Available: https://doi.org/10.1016/j.ssci.2023.106368.
[8] J. Karlsson et al., "Visual Detection of Personal Protective Equipment and Safety Gear on Industry Workers," 2022, unpublished.
[9] C. Shi, D. Zhu, J. Shen, Y. Zheng, and C. Zhou, "GBSG-YOLOv8n: A model for enhanced personal protective equipment detection in industrial environments," Electronics, vol. 12, p. 4628, 2023. [Online]. Available: https://doi.org/10.3390/electronics12224628.
[10] F. Khan, N. Zafar, M.N. Tahir, M. Aqib, H. Waheed, and Z. Haroon, "A mobile-based system for maize plant leaf disease detection and classification using deep learning," Frontiers in Plant Science, vol. 14, p. 1079366, May 2023. doi: 10.3389/fpls.2023.1079366. PMID: 37255561; PMCID: PMC10226393.
[11] N. B. A. Karna, M. A. P. Putra, S. M. Rachmawati, M. Abisado, and G. A. Sampedro, "Toward Accurate Fused Deposition Modeling 3D Printer Fault Detection Using Improved YOLOv8 With Hyperparameter Optimization," IEEE Access, vol. 11, pp. 74251- 74262, 2023. doi: 10.1109/ACCESS.2023.3293056.
[12] V. Delhi, S. Lal, and A. Thomas, "Detection of personal protective equipment (ppe) compliance on construction site using computer vision based deep learning techniques," Frontiers in Built Environment, vol. 6, September 2020.
[13] J.-H. Lo, L.-K. Lin, and C.-C. Hung, "Real-time personal protective equipment compliance detection based on deep learning algorithm," Sustainability, vol. 15, p. 391, December 2022.
[14] M. Riaz, J. He, K. Xie, H.S. Alsagri, S.A. Moqurrab, H.A.A. Alhakbani, and W.J. Obidallah, "Enhancing Workplace Safety: PPE_Swin—A Robust Swin Transformer Approach for Automated Personal Protective Equipment Detection," Electronics, vol. 12, p. 4675, 2023. [Online]. Available:https://doi.org/10.3390/electronics12224675.