[1] Nurzada Amangeldy
[2] Alibek Barlybaev
[3] Nazerke Gazizova *
[4] Bekbolat Kurmetbek

Evaluation of YOLOv8 and YOLOv10 Models in PPE Recognition Tasks

Journal of Electrical Systems

Abstract: - This paper presents a detailed analysis and performance comparison of the You Only Look Once (YOLO) version 8 and 10 deep learning models based on a dataset designed for personal protective equipment recognition. The novelty of the work lies in the fact that for the first time a comprehensive comparison of the results of YOLOv8 and YOLOv10 models is conducted on this dataset, allowing for an objective assessment of the improvements in the YOLOv10 architecture compared to the previous version. The obtained results demonstrated that YOLOv10 outperforms YOLOv8 in all major performance metrics, especially in tasks requiring high accuracy and completeness of object detection. However, particularly the YOLOv8m version, due to its stability and efficiency at lower computational resources, remains a strong candidate for applications where computational efficiency and processing speed are critical.

Keywords: Deep Learning Models, Personal Protective Equipment (PPE), Object Detection.

I. INTRODUCTION

Our study addresses pressing issues in the recognition of personal protective equipment (PPE) in workplace environments, particularly under real-world conditions where lighting, viewing angles, and complex backgrounds can negatively impact detection accuracy. In autonomous driving, YOLO is employed to detect and recognize objects on the road, such as vehicles, pedestrians, traffic signs, and traffic lights. This capability enables autonomous driving systems to make real-time decisions and prevent accidents [1-4]. In surveillance systems, YOLO assists in detecting suspicious objects or individuals in real time, enhancing safety and enabling prompt responses to potential threats [5-8]. YOLO is also utilized in facial recognition, playing a key role in biometric identification systems, ensuring access control and security [9-12].

In agriculture, this algorithm is applied to monitor crop conditions, detect pests, and automate harvesting processes [13-15]. In the medical field, YOLO is used for analyzing medical images, such as X-rays, contributing to early disease diagnosis, including the detection of pathologies such as tumors [16-18]. Its high accuracy and real-time processing capability make it an ideal tool for tasks requiring real-time analysis.

One of the crucial applications of deep learning algorithms is enhancing workplace safety. Personal protective equipment plays a key role in preventing injuries in industrial and construction sites. Automating the process of recognizing and monitoring PPE has become a priority for reducing risks and increasing workplace safety. In this domain, YOLO algorithms [19-21] have been widely adopted for helmet detection. Fan et al. [19] and Wang et al. [20] found ways to improve the YOLOv3 algorithm [24], particularly for helmet detection in the industry. Wang and his colleagues [20] created a high-quality dataset and used several YOLO versions to detect six object classes (four helmet colors, a person, and a vest), demonstrating the effectiveness of YOLOv5x for PPE detection. The COVID-19 pandemic necessitated stricter measures to ensure compliance with PPE usage, leading to various studies aimed at developing YOLO-based systems capable of detecting masks and gloves [25-27]. In 2018, Xie and his colleagues [18] evaluated different detection models on the same datasets. Results showed that the algorithm achieved the highest average precision (53.8%) and the fastest speed (10 FPS), surpassing SSD and Faster R-CNN in both metrics.

A challenge faced by current algorithms is their limited ability to detect such objects under variable and complex environmental conditions. This can result in missing critical elements, such as helmets or vests, which is crucial in industrial settings where employee safety directly depends on the effectiveness of PPE monitoring systems. Incomplete detection or frequent false positives can lead to reduced efficiency of these systems and, consequently, to threats to worker safety.

[1] L.N. Gumilyov Eurasian National University, Kazakhstan. nurzadaamangeldy@gmail.com
[2] L.N. Gumilyov Eurasian National University, Kazakhstan. frank-ab@mail.ru
[3] L.N. Gumilyov Eurasian National University, Kazakhstan. g.nazerke2755@gmail.com
[4] L.N. Gumilyov Eurasian National University, Kazakhstan. bigboss990930@gmail.com
* Corresponding Author Email: g.nazerke2755@gmail.com

Our work proposes an approach to address this problem through a detailed comparison of the performance of YOLOv8 and the latest YOLOv10 architecture, a comparison that has not been previously conducted in the literature. We use the high-quality SHEL5K dataset, specifically created for PPE recognition tasks, which includes diverse images with varying background conditions and objects. We focus on key metrics such as Precision, Recall, mAP50, and mAP50-95, which allows us not only to evaluate detection accuracy but also to gain deeper insights into how effectively the models perform in real-world conditions.

We also propose using our results to optimize monitoring systems in industrial settings, which could significantly improve safety and reduce risks. Experimental results will help to understand which improvements in the YOLOv10 architecture may contribute to better object recognition and how these improvements can be applied to enhance workplace safety.

II. Methods

The research method was aimed at a detailed analysis and comparison of the performance of YOLOv8 and YOLOv10 models in object detection tasks using data on personal protective equipment (PPE). The study was based on the following stages:

A. Data Collection and Preparation:

In this study, the SHEL5K dataset was used, which was detailed in the review section. This dataset includes 5000 images annotated across six classes and is widely applied for personal protective equipment (PPE) recognition tasks. For comparison, results from our previous research, published in the work by the authors, were used, where the YOLOv8 architecture was employed for PPE recognition on various benchmark object detection datasets (Cogent Engineering, 11(1)). In this study, the YOLOv10 model was additionally trained on the same SHEL5K dataset, allowing for a direct comparison with YOLOv8 results. Performance analysis was conducted based on key metrics such as Precision, Recall, mAP50, and mAP50-95 to assess improvements in the new YOLOv10 architecture compared to the previous YOLOv8 version.

In the context of PPE recognition tasks, the SHEL5K dataset (Safety Helmet Detection with Extended Labels 5K) plays a crucial role. This dataset, created by Munk-Erdene Otgonbold, Munkhjargal Gochoo, Fadi Alnajjar, and others, was published in the Sensors journal in 2022. SHEL5K includes 5000 images annotated across six classes and provides researchers and developers with high-quality data for training and testing deep learning models, such as YOLO and SSD [28].

In the article "Automatic Detection of Construction Workers’ Helmet Wear Based on Lightweight Deep Learning," authors Han Liang and Suyoung Seo proposed a helmet-wearing recognition method by combining two datasets: SHWD (Safety Helmet Wearing Dataset) and SHEL5K. The new combined dataset, which includes over 12000 annotated images, achieved high accuracy in helmet detection while reducing computational costs. This solution proved effective for real-world industrial applications where helmet monitoring is necessary [29].

In the article "Research on Small Target Detection Method for Industrial Safety Helmets Based on Improved YOLOv8," an improved version of YOLOv8 is proposed for detecting small objects, such as helmets, under challenging conditions. Using SHEL5K for model training demonstrated high accuracy during validation, highlighting the effectiveness of this dataset in industrial safety tasks [30].

The article "Industrial Safety Helmet Detection: A Comprehensive Review and Evaluation of State-of-the-Art Models Using SHEL5K Dataset" assesses various models for helmet recognition using SHEL5K. The study examines different approaches, including deep learning-based methods such as YOLO and SSD, analyzing their performance and accuracy in real-world conditions [28].

The study "Enhanced Detection of Industrial Safety Helmets Using Deep Learning Techniques: Insights from SHEL5K Dataset" focuses on improving helmet detection methods using new deep learning architectures. Utilizing SHEL5K and integrating its data with advanced methods significantly improves accuracy and speed in helmet detection within industrial environments [31].

Additionally, in the article "Deep Learning Based Human Robot Interaction with 5G Communication," authors Mücahid Barstuğan and Zeynep Osmanpaşaoğlu mention the use of SHEL5K as part of a helmet detection system in the context of occupational health and safety. They discuss the application of YOLO methods for object detection, such as protective helmets, across various industries, emphasizing its high accuracy and real-time capability [33].

Thus, SHEL5K represents a key dataset actively used in research for automating PPE recognition. It supports the development and testing of effective deep learning models aimed at enhancing safety and preventing workplace injuries.

B. Conducting Experiments:

Experiments were conducted using the SHEL5K dataset, which includes 5000 images annotated with categories of personal protective equipment (PPE). To evaluate the performance of the YOLOv8 and YOLOv10 models, key metrics such as Precision (P), Recall (R), mAP50, and mAP50-95 were selected. These metrics provide a comprehensive assessment of the models' accuracy and completeness. Precision measures the proportion of correct predictions among all positive predictions, indicating the model's ability to accurately classify objects. Recall, on the other hand, evaluates the proportion of correctly identified objects among all true objects, helping to understand how effectively the model detects all target objects in the dataset. Additionally, for a more in-depth performance analysis, metrics such as mAP50 (mean average precision with an IoU threshold of 0.50) and mAP50-95 (mean average precision at varying IoU levels from 0.50 to 0.95) were used to assess the model's accuracy based on the degree of overlap between predicted and actual objects.

Analysis of the graphs (Figures 1-4) revealed the strengths and weaknesses of each model. For instance, YOLOv10x shows the highest values for Precision and mAP, making it preferable for tasks requiring maximum accuracy. On the other hand, YOLOv8 models, particularly YOLOv8n, although less accurate, maintain good performance for tasks where computational cost and speed are critical. The weaknesses of YOLOv8 are associated with a slight reduction in accuracy and completeness compared to the new YOLOv10 architecture.

YOLOv10 exhibits higher Precision values compared to YOLOv8, particularly in the YOLOv10x version, where Precision reaches 0.957, significantly surpassing YOLOv8x (0.920). It is evident that YOLOv10 performs better in tasks where minimizing false positives is crucial, which is critical for systems focused on industrial safety (Fig. 1).

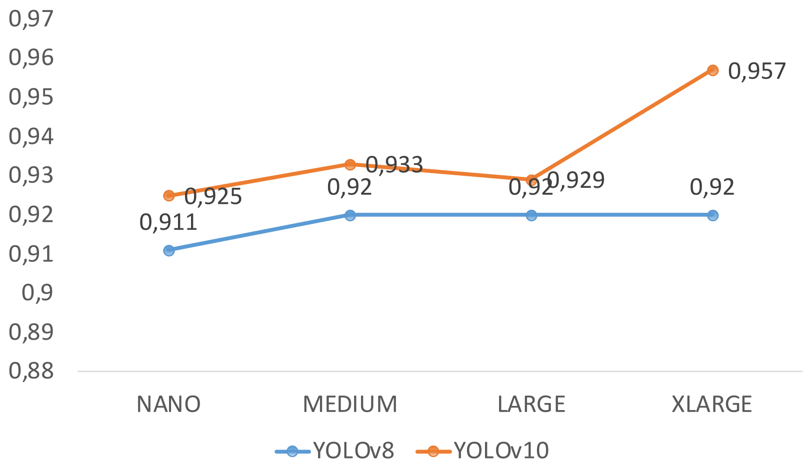

Fig. 1. Comparison of Precision values

For the Recall metric, YOLOv10 models also demonstrate better results. For instance, YOLOv10m has a Recall of 0.884, surpassing YOLOv8m (0.833). This indicates that version 10 is capable of detecting more objects, which is particularly important for safety-related tasks where missed objects can lead to significant consequences (Fig. 2).

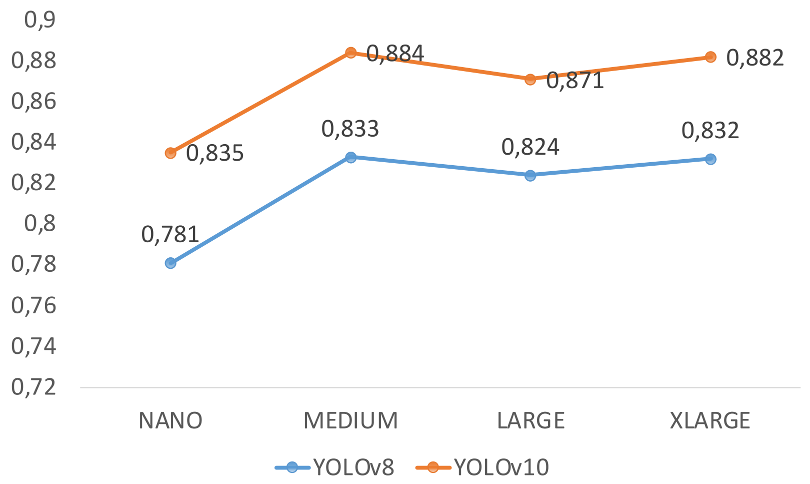

Fig. 2. Comparison of Recall values

In terms of the mAP50 metric, version 10 also outperforms, with a value of 0.936, which is higher than YOLOv8x (0.896) (Fig. 3).

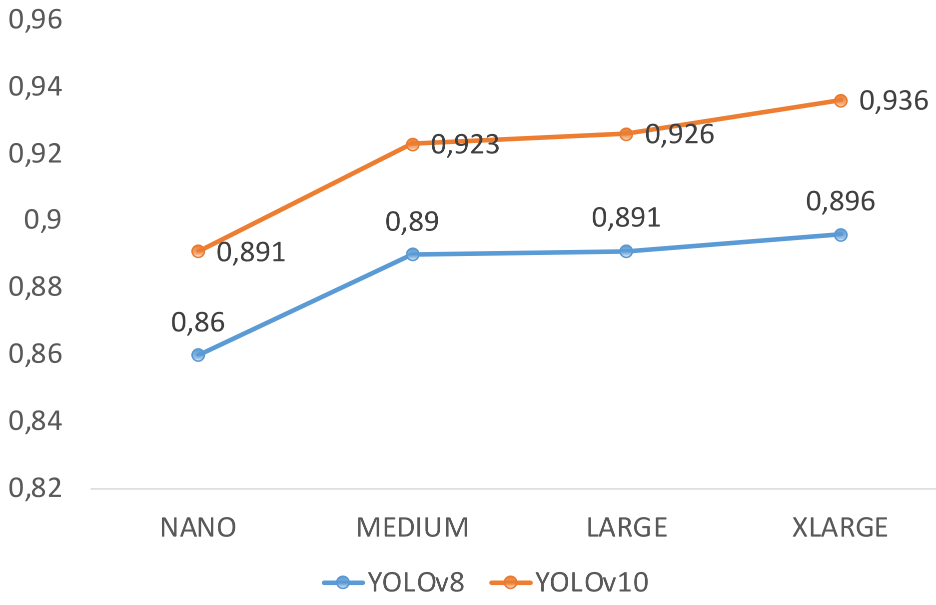

Fig. 3. Comparison of mAP50 values

For the mAP50-95 metric, YOLOv10 models demonstrate better performance. However, the YOLOv8m (Medium) version may be preferable to YOLOv10m in certain scenarios and for specific requirements. YOLOv8m is more suitable for tasks that require a balance between accuracy and speed, especially on devices with limited resources. Although YOLOv10m often outperforms YOLOv8m on most metrics, YOLOv8m retains its advantage in situations where not only accuracy but also resource efficiency, operational stability, and deployment speed are important (Fig. 4).

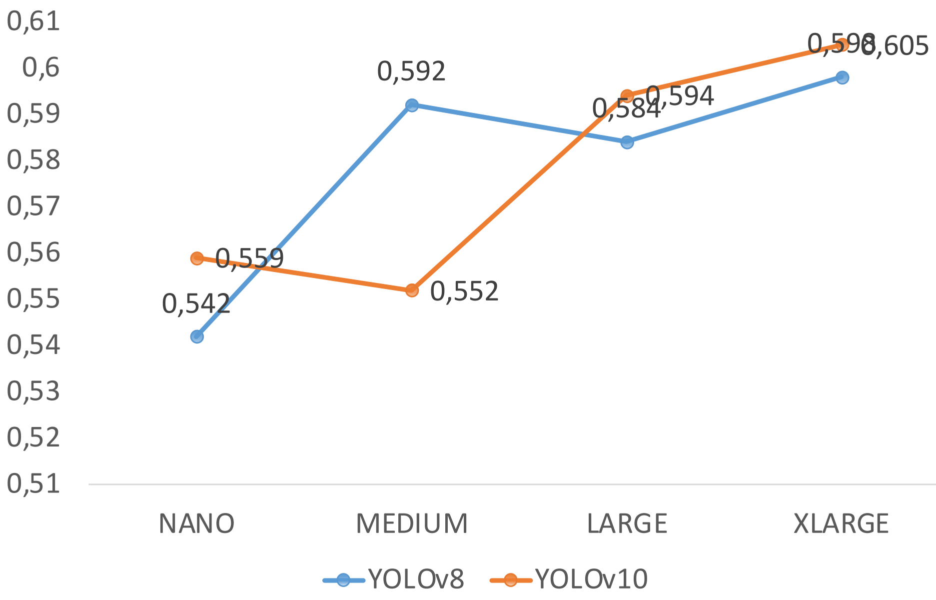

Fig. 4. Comparison of mAP50-95 values

To gain a deeper understanding of the performance differences between YOLOv8 and YOLOv10 models, it is not sufficient to rely solely on quantitative metrics. It is also crucial to examine the errors made by the models in object recognition and how they handle complex cases. For this purpose, confusion matrices were used to provide a visual representation of correct and incorrect classifications (Fig. 5, 6). Analyzing these matrices helps to identify where the models perform better or worse, especially under complex and ambiguous conditions.

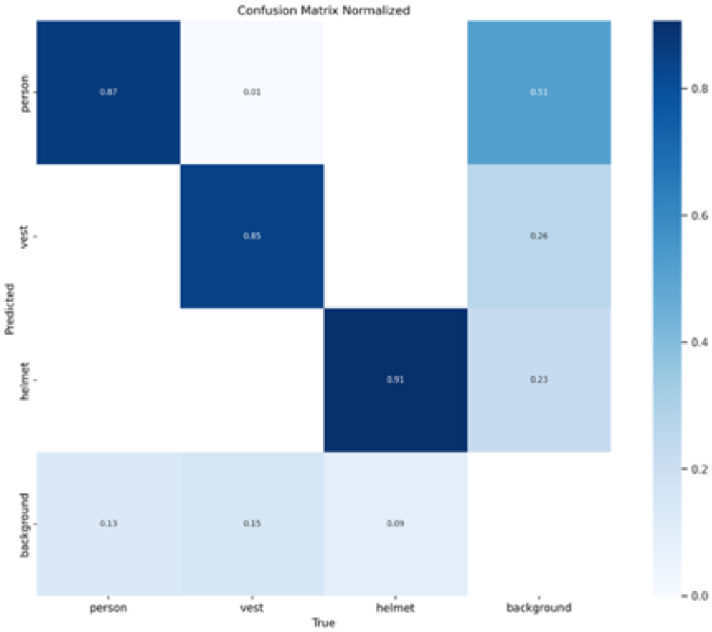

Fig. 5. Confusion Matrices for YOLOv8-X model

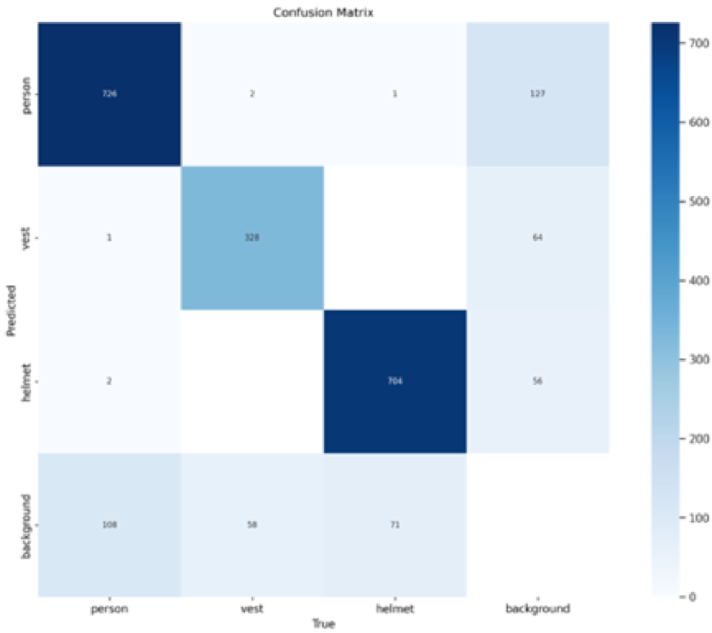

Fig. 6. Confusion Matrices for YOLOv10-X model

The analysis of the two confusion matrices clearly illustrates the differences in the performance of YOLOv8 and YOLOv10 models. In the first matrix (YOLOv8), the model shows high accuracy in recognizing objects such as helmets and heads, but makes more errors in recognizing backgrounds and people. This leads to an increase in false positives and missed detections, which reduces the overall accuracy of the model.

In version 10, significant improvements are observed in object recognition, with more accurate identification of key categories such as helmets and people, and fewer errors in classifying backgrounds and other classes. This indicates that YOLOv10 has a better ability to distinguish objects in complex conditions, which is particularly important for tasks where minimizing false positives is critical. The more precise recognition of personal protective equipment (PPE) highlights the advantage of YOLOv10 over YOLOv8 in the context of industrial safety, where accurate object classification can directly impact the effectiveness of monitoring and accident prevention.

Examining the distributions of helmets, vests, heads, and other elements in images allows for an assessment of how well the models recognize objects with different characteristics and under various conditions. Distribution diagrams provide valuable information about the frequency of different category objects in the dataset and the sizes they occupy (Fig. 7, 8).

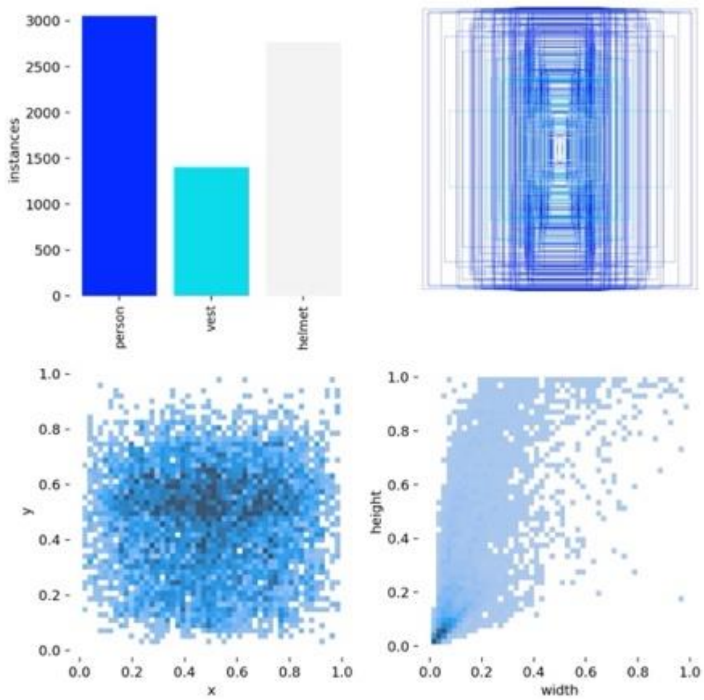

Fig. 7. Distribution diagrams for YOLOv8-M model

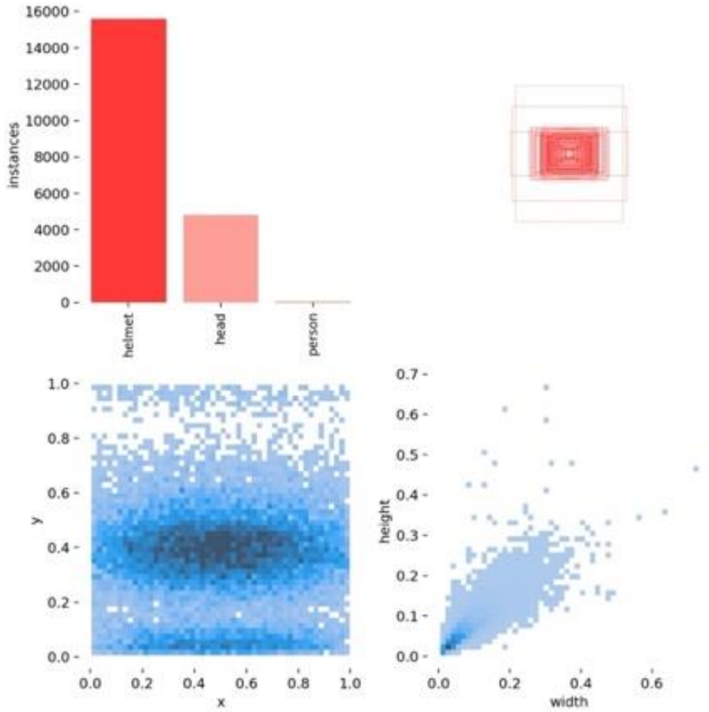

Fig. 8. Distribution diagrams for YOLOv10-M model

For example, the diagram reflecting the results of the YOLOv8 model shows the distribution of objects across categories. The "helmet" category is the most frequently represented, while the "head" and "person" categories are significantly less common. The object density plots indicate that "head" and "person" are typically located in the center of the image and have standard sizes. This simplifies the model's task of recognizing these objects but can lead to potential errors if the objects are located outside these standard areas or have non-standard sizes.

In the diagram related to the YOLOv10 model, a more even distribution of objects across the "person," "vest," and "helmet" categories is observed. The density plots for object width and height show a more varied distribution of object sizes and placements in the image, indicating that YOLOv10 is better able to adapt to different recognition conditions and objects of various sizes. This improvement makes YOLOv10 more resilient to changing conditions, which is crucial for PPE recognition tasks in real industrial scenarios.

The experiments testing YOLOv8 and YOLOv10 models were conducted on the same SHEL5K dataset, ensuring uniform conditions for comparing their performance. The results obtained for key metrics were carefully analyzed to identify performance differences between the models. The focus was on comparing versions of the same class, such as YOLOv8n and YOLOv10n, to evaluate the improvements in the new YOLOv10 architecture and their impact on accuracy, completeness, and overall performance metrics. This approach not only highlighted the advantages of YOLOv10 over the previous version but also provided recommendations for selecting the most suitable model for various PPE recognition applications.

III. Analysis Of Results And Discussion

The performance results of YOLOv8 and YOLOv10 models were systematically organized and presented in the form of graphs (Figures 1-4) for clear comparison. These graphs include key metrics such as Precision (P), Recall (R), mAP50, and mAP50-95, allowing for a visual assessment of the performance changes in each model. These graphs make it evident how YOLOv10 models show improvements over YOLOv8 across nearly all metrics.

The results for YOLOv8m and YOLOv10m models demonstrate that, despite the clear advantage of YOLOv10m in the mAP50 metric, YOLOv8m may be a preferable choice in certain scenarios. The presented graphs indicate that, although YOLOv10m slightly outperforms YOLOv8m in Precision and Recall metrics, the latter shows a more stable reduction in loss during both training and validation phases. This suggests that YOLOv8m is more computationally efficient with lower resource consumption. In cases where a trade-off between accuracy and speed is necessary, particularly on devices with limited computational resources, YOLOv8m may be a more suitable model. This model is also preferable in conditions where deployment stability and operational speed are more critical than absolute accuracy.

Thus, despite YOLOv10m's superiority in several key metrics, YOLOv8m remains a competitive model, as evidenced not only by our current experiments but also by previous studies. This is particularly relevant for tasks where resource efficiency is of primary importance.

IV. Conclusion

Models YOLOv10 exhibit clear superiority over YOLOv8 in terms of mAP50, Precision, and Recall metrics, making them preferable for tasks related to industrial safety and PPE (Personal Protective Equipment) recognition, where accuracy and completeness are critical. However, YOLOv8, particularly the YOLOv8m version, remains a strong candidate for tasks where computational efficiency and processing speed are important. The choice between YOLOv8 and YOLOv10 should be based on the specific requirements of the task: if the priority is maximum accuracy and completeness, YOLOv10 will be the best choice. However, for tasks with less stringent accuracy requirements or in resource-constrained environments, YOLOv8 may be a more optimal solution.

References

[1] Cao, Yining, et al. "MCS-YOLO: A multiscale object detection method for autonomous driving road environment recognition." IEEE Access 11 (2023): 22342-22354.

[2] Liu, Jun, et al. "BiGA-YOLO: a lightweight object detection network based on YOLOv5 for autonomous driving." Electronics 12.12 (2023): 2745.

[3] Bao, Decheng, and Renjing Gao. "YED-YOLO: An object detection algorithm for automatic driving." Signal, Image and Video Processing 18.10 (2024): 7211-7219.

[4] Ren, Hongge, Fangke Jing, and Song Li. "DCW-YOLO: Road Object Detection Algorithms for Autonomous Driving." IEEE Access (2024).

[5] Zhou, Wentao, et al. "ASSD-YOLO: a small object detection method based on improved YOLOv7 for airport surface surveillance." Multimedia Tools and Applications 83.18 (2024): 55527-55548.

[6] Mostafa, Salama A., et al. "A YOLO-based deep learning model for Real-Time face mask detection via drone surveillance in public spaces." Information Sciences (2024): 120865.

[7] Zhang, Dawei. "A Yolo-based Approach for Fire and Smoke Detection in IoT Surveillance Systems." International Journal of Advanced Computer Science & Applications 15.1 (2024).

[8] Dalal, Surjeet, et al. "Improving smart home surveillance through YOLO model with transfer learning and quantization for enhanced accuracy and efficiency." PeerJ Computer Science 10 (2024): e1939.

[9] Anusudha, K. "Real time face recognition system based on YOLO and InsightFace." Multimedia Tools and Applications 83.11 (2024): 31893-31910.

[10] Ali, Nashwan Saleh, et al. "An Effective Face Detection and Recognition Model Based on Improved YOLO v3 and VGG 16 Networks." Journal homepage: http://iieta. org/journals/ijcmem 12.2 (2024): 107-119.

[11] Azatbekuly, Nurtugan, Aksultan Mukhanbet, and Samson Dawit Bekele. "Development of an Intelligent Video Surveillance System Based on YOLO Algorithm." 2024 IEEE 4th International Conference on Smart Information Systems and Technologies (SIST). IEEE, 2024.
[12] Badgujar, Chetan M., Alwin Poulose, and Hao Gan. "Agricultural object detection with You Only Look Once (YOLO) Algorithm: A bibliometric and systematic literature review." Computers and Electronics in Agriculture 223 (2024): 109090.
[13] Cui, Beibei, et al. "Exploring the YOLO-FT Deep Learning Algorithm for UAV-Based Smart Agriculture Detection in Communication Networks." IEEE Transactions on Network and Service Management (2024).
[14] Gao, Xiaomei, et al. "Improved YOLO v7 for Sustainable Agriculture Significantly Improves Precision Rate for Chinese Cabbage (Brassica pekinensis Rupr.) Seedling Belt (CCSB) Detection." Sustainability 16.11 (2024): 4759.
[15] Kaur, Tanvir, Shivani Kamboj, and Lovedeep Singh. "Advanced YOLO-NAS-Based Detection and Screening of Brain Tumors Using Medical Diagnostic." 2024 2nd International Conference on Artificial Intelligence and Machine Learning Applications Theme: Healthcare and Internet of Things (AIMLA). IEEE, 2024.
[16] Prinzi, Francesco, et al. "A yolo-based model for breast cancer detection in mammograms." Cognitive Computation 16.1 (2024): 107-120.
[17] Baranwal, Naman, Kedar Nath Singh, and Amit Kumar Singh. "YOLO-based ROI selection for joint encryption and compression of medical images with reconstruction through super-resolution network." Future Generation Computer Systems 150 (2024): 1-9.
[18] Xie, Z., Liu, H., Li, Z., & He, Y. (2018, December). A convolutional neural network based approach towards real-time hard hat detection. In 2018 IEEE International Conference on Progress in Informatics and Computing (PIC) (pp. 430-434). IEEE.
[19] Fan, Zheming, et al. "A deep learning-based ensemble method for helmet-wearing detection." PeerJ Computer Science 6 (2020): e311.
[20] Wang, H., Hu, Z., Guo, Y., Yang, Z., Zhou, F., & Xu, P. (2020). A real-time safety helmet wearing detection approach based on CSYOLOv3. Applied Sciences, 10(19), 6732.
[21] Xiong, R., & Tang, P. (2021). Pose guided anchoring for detecting proper use of personal protective equipment. Automation in Construction, 130, 103828.
[22] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).
[23] Redmon, J., & Farhadi, A. (2017). YOLO9000: better, faster, stronger. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 7263-7271).
[24] Redmon, Joseph. "Yolov3: An incremental improvement." arXiv preprint arXiv:1804.02767 (2018).
[25] Loey, M., Manogaran, G., Taha, M. H. N., & Khalifa, N. E. M. (2021). Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection. Sustainable cities and society, 65, 102600.
[26] Protik, A. A., Rafi, A. H., & Siddique, S. (2021, August). Real-Time personal protective equipment (PPE) detection using YOLOv4 and TensorFlow. In 2021 IEEE Region 10 Symposium (TENSYMP) (pp. 1-6). IEEE.
[27] Avanzato, R., Beritelli, F., Russo, M., Russo, S., & Vaccaro, M. (2020). Yolov3-based mask and face recognition algorithm for individual protection applications. In CEUR Workshop Proceedings (pp. 41-45).
[28] Otgonbold, M., Gochoo, M., Alnajjar, F., et al. (2022). SHEL5K: An Extended Dataset and Benchmarking for Safety Helmet Detection. Sensors. DOI: 10.3390/s22062315
[29] Liang, H., & Seo, S. (2023). Automatic Detection of Construction Workers' Helmet Wear Based on Lightweight Deep Learning. Sensors. DOI: 10.3390/s23084055
[30] Han, L., & Seo, S. (2023). Research on Small Target Detection Method for Industrial Safety Helmets Based on Improved YOLOv8. Sensors.
[31] SHEL5K Dataset. Enhanced Detection of Industrial Safety Helmets Using Deep Learning Techniques: Insights from SHEL5K Dataset.
[32] Barstuğan, Mücahid, and Zeynep Osmanpaşaoğlu. "DEEP LEARNING BASED HUMAN ROBOT INTERACTION WITH 5G COMMUNICATION." Konya Journal of Engineering Sciences 11.2 (2023): 423-438.
[33] Barlybayev, A., Amangeldy, N., Kurmetbek, B., Krak, I., Razakhova, B., Tursynova, N., & Turebayeva, R. (2024). Personal protective equipment detection using YOLOv8 architecture on object detection benchmark datasets: a comparative study. Cogent Engineering, 11(1). https://doi.org/10.1080/23311916.2024.2333209