Vol. 15, No. 43, Summer 2025, pages 43-54

Journal of Analytical and Numerical Methods in Mining Engineering

Journal home page: http://anm.yazd.ac.ir/

Research article

Optimized YOLO Model for Accurate and Real-Time Detection of Machinery Around Shovels in Copper Mining

Mohaddeseh Ghiasi[1], MasoudReza Aghabozorgi[1]*

1- Dept. of Electrical Engineering, Yazd University, Yazd, Iran

*Corresponding author: E-mail: *aghabozorgi@yazd.ac.ir*

(Received: *January 2025*, Accepted: *May 2025*)

DOI: 10.22034/ANM.2025.22650.1659

Keywords

Copper mining

Object detection

Shovel

YOLO

Abstract

Shovels are among the most important equipment in open-pit mining operations, widely used for loading minerals. These heavy machines play a crucial role in operational efficiency, but due to the operator's visibility limitations, particularly in the shovel's blind spots, they pose significant safety risks. In these situations, operators may face challenges in detecting vehicles around the shovel, increasing the likelihood of accidents and incidents. This study proposes an enhanced version of the YOLO model for the precise and rapid detection of vehicles around the shovel in copper mining environments. The proposed model, using real-time processing, is capable of detecting vehicles in four directions around the shovel and preventing collisions. To evaluate this model, real-world data collected from four cameras installed around the shovel in a copper mine under various lighting conditions, including day and night, were used. The proposed method was evaluated on a new dataset of shovels under real working conditions. The results, with an average accuracy of 94.2% and a rate of 159 fps, demonstrate a significant improvement in detection accuracy and an increase in the speed of the recognition process, meeting the requirements for accurate and real-time detection of vehicles around the shovel. The findings show that the proposed model can act as an effective collision avoidance system, preventing collisions between the shovel and surrounding vehicles, which directly enhances the safety of the work environment and personnel. Furthermore, this system can help reduce accidents and injuries caused by collisions between shovels and surrounding vehicles, thereby improving the overall productivity of mining operations in copper mines.

1. INTRODUCTION

Although mines are divine treasures on the earth's surface and must be exploited to elevate the economic level of human societies, they have always been a lurking danger for miners. Many workers have lost their lives in this pursuit. Mining accidents, regardless of their cause, have an adverse effect on the workforce and the productivity of mining units. While financial losses resulting from accidents may be recoverable, the psychological damage to workers' morale is often irreparable. Mining is the backbone of the economy in many countries, and without it, their economies would face severe disruption. For example, Chile derives 14% of its GDP and 59% of its total exports from mining, Australia relies on mining for 10% of its GDP and over 260,000 jobs, and South Africa depends on mining for 8.5% of its GDP and 30% of its foreign exchange earnings. However, it is accompanied by numerous risks. Safety has always been a challenging issue not only for workers but also for mining companies and labor unions. Various studies and research around the world indicate that despite efforts to reduce accidents, the number of serious incidents and fatalities in mines remains high. Despite advancements in modern technology [1-3], mining remains one of the industries with the highest rates of accidents [4-8]. For instance, Spain reports 12.7 accidents per 1,000 workers annually due to inadequate safety protocols [4], Indonesia attributes 34% of mining accidents to gaps in supervision and training [5], and Turkey identifies 58% of severe injuries as stemming from equipment failures [7].In recent years, despite technological advancements, the number of incidents caused by mining equipment in certain regions has risen [9]. Many workers seem unable to identify hazards or interpret and recognize dangerous situations in the context of necessary actions [10,11].

A particularly promising application of AI in mining is the development of collision avoidance systems powered by computer vision. These systems leverage advanced sensors such as cameras and LIDAR [12,13], combined with complex algorithms and deep learning models, to detect and track the movement of workers and objects within underground mines [14-16]. By providing real-time warnings and alerts to operators and personnel, this technology plays a crucial role in preventing accidents and ensuring worker safety.

The technological advancements, collectively referred to as Mining 4.0 [17–19], have played a pivotal role in this transformation, driving a fundamental shift in how mining operations are conducted. However, despite these technological improvements, the mining industry still faces challenges such as elevated rates of accidents and occupational illnesses, particularly when large, expensive machinery is employed to meet high production demands. Nevertheless, the transition from Mining 4.0 to Mining 5.0 has triggered a significant transformation, leading to more efficient collaboration between human workers and autonomous systems.

Numerous studies [20-23] indicate that mining equipment is the primary cause of injuries in the mining industry. The open-pit mining system, driven by the need for increased production, is advancing toward the use of larger machinery, which has raised concerns about the safety of these machines [20,22]. Blind spots in construction machinery refer to areas around heavy vehicles that are not directly visible through windows or mirrors, rendering the operator visually impaired in those zones [21,23]. These blind spots pose significant safety risks as they can obscure the presence of workers, pedestrians, or other vehicles near the equipment, leaving operators unaware of potential hazards [22]. To prevent accidents and ensure the safety of both operators and those working around construction machinery, identifying and minimizing these blind spots is crucial [20,23]. Recognizing blind spots in construction machinery requires a thorough understanding of each machine's design and its visibility limitations [21,22].

One of the fields where technological advancements have made a profound impact is artificial intelligence (AI). Across the entire mining lifecycle-from exploration and planning to reclamation and closure-AI can be utilized at various stages, including exploration, mine planning, mobile equipment operations, drilling and blasting, and mineral processing [24,25]. In recent years, AI has significantly advanced the automation of machinery and vehicle operations, enhancing both efficiency and safety.

Shovels are a type of mechanical excavator widely used in open-pit mining to drill and move large volumes of ore and waste materials. These large and powerful machines play a key role in the mining extraction process. However, one of the major challenges in their use is the presence of blind spots, which, due to the operator's limited visibility, can create serious risks for both workers and equipment. Blind spots typically include areas behind the machine, around it, or regions covered by parts of the machine, such as the shovel bucket.

For example, the rear swing radius and the presence of various equipment behind the operator can block their view, especially during rotation, which limits the ability to operate at full efficiency. Additionally, the bucket and boom areas can create significant blind spots, making it difficult to see objects close to the excavator.

To reduce the risks associated with these blind spots, our goal is to utilize advanced technologies such as object detection systems. These systems can accurately detect vehicles around the shovel and alert the operator about the surrounding environment. Specifically, we aim to install cameras around the shovel and use artificial intelligence algorithms for vehicle detection to significantly improve operational safety. These systems can notify the driver of potential hazards through audio alerts or visual signals displayed on mirrors and monitors.

Over several decades of research aimed at object detection and identification in images, various methods have been proposed. Among the earliest and most successful ones are techniques like Viola-Jones [26], HOG [27], and DPM [28]. These methods worked by manually extracting features designed by researchers and using sliding windows, resulting in slow speeds and poor performance in detecting objects in complex images. With the saturation of classical methods, the advent of deep learning and convolutional neural networks (CNNs) changed the landscape of visual perception, leading to the development and introduction of deep learning-based object detection models and both single-stage and two-stage CNN architectures.

Among two-stage detectors, Faster R-CNN [29] is a notable example. Despite advancements in two-stage detectors, their speed is still limited by the multi-stage process [30]. In contrast, single-stage detectors search for objects at specific locations and sizes, with both bounding box extraction and classification occurring in a single step.

This results in faster speeds compared to two-stage detectors. SSD [31], YOLO [32], Retina-Net [33], and Center-Net [34] are some of the prominent single-stage detectors. One of the pioneering works in single-stage object detection is YOLO.

YOLO employs a Convolutional Neural Network (CNN) to simultaneously predict multiple bounding boxes and their corresponding classes across the entire image by dividing the image into an S×S grid. Each grid cell predicts bounding boxes and object confidence scores. The authors of YOLO improved their model with newer versions; YOLO9000 [35] introduced multi-scale training, and YOLOv3 [36] used an enhanced backbone called Darknet53, incorporating multi-scale detection for better object recognition at varying sizes. While YOLOv3 was the last official version developed by the original authors, other researchers continued to advance the model and introduced improved YOLO-based architectures. For instance, in 2020, YOLOv4 [37] was introduced, featuring advancements like cross-stage partial connections, Mish activation, CIoU, and new image augmentation techniques, enhancing YOLOv3's performance.

One major advancement in YOLO-based object detectors is YOLOv5 [38], which shares structural similarities with YOLOv4 but includes several enhancements, such as generating more accurate anchor boxes using genetic algorithms and implementing the model in Python and the Pytorch framework [39]. This makes YOLOv5 less complex and faster than other YOLO-based models, making it a favorable choice for real-time object detection tasks. YOLOv5 is an ongoing and popular project, further improved with the introduction of C3 and SPPF modules and the SiLU activation function [40].

Several unofficial YOLO-based models were introduced in 2022 and 2023 by various researchers. YOLOv6 [41] includes a new backbone network called Efficient-Rep and a new neck network named Rep-PAN, with separate localization and classification heads. A few months later, Wang et al. introduced YOLOv7 [42], featuring a new backbone called ELAN and auxiliary detection heads to enhance accuracy, making it significantly different from other YOLO-based models. Additionally, the first version of YOLOv8 was released in January 2023 by the YOLOv5 authors, introducing anchor-free object localization, C2F blocks, and online image augmentation techniques. However, YOLOv8 is still under active development [43].

Compared to earlier models like YOLOv5, which is widely used today, the newer YOLO variants mentioned are rarely applied in practice and require ongoing improvements and evaluations for real-world applications [44]. While the performance of YOLO models has improved over time, it is worth noting that they often prioritize speed and efficiency over accuracy. This trade-off is essential, as it enables real-time object detection across various applications. Many real-world applications require real-time object detection, which makes achieving a balance between speed and accuracy a challenging task.

To address this challenge, this work focuses on one of the most prominent members of the YOLO object detector family [38]. In order to achieve precise and accurate detection of vehicles around the shovel, modifications have been made to the YOLOv5 object detector to enhance its detection capabilities. The selection of YOLOv5 as the foundation of this study, despite the availability of newer versions, stems from a rigorous engineering analysis of the operational demands in mining environments.

Inherent compatibility with non-ideal mining conditions-such as dense dust, abrupt lighting variations, and partially obscured objects-makes YOLOv5 a more reliable choice. Its training on custom datasets collected from real-world mining scenarios equips it with unparalleled robustness against environmental noise. In contrast, newer versions, primarily trained on generic datasets, require extensive retraining and complex adjustments to adapt to mining-specific challenges.

The simpler modular architecture of YOLOv5 allows seamless integration of advanced attention mechanisms without fundamental alterations to core layers. Meanwhile, the modified architectures of newer versions introduce complexity that complicates development and optimization.

Native optimization for real-time processing through lightweight model variants enables simultaneous analysis of video streams from multiple shovel-mounted cameras without relying on high-power hardware. This capability is critical in mining environments, where processing delays directly impact safety. Broad industrial ecosystem support for YOLOv5, including deployment tools for edge platforms and compatibility with multi-sensor integration, ensures the flexibility needed to evolve the system toward future architectures. These convergent advantages solidify YOLOv5 as an engineered, practical solution for vehicle detection in shovel blind spots.

A shovel dataset from a copper mine, collected under real-world working conditions, was used to train a deep learning model on the given dataset, with an average accuracy of up to 94%. In Section 2, the overall structure of the YOLOv5 network and the proposed modifications made to the base model to improve its performance are explained.

Then, in Section 3, the evaluation metrics and result analysis are presented, and finally, the last section is dedicated to the conclusion and future recommendations.

2. METHODOLOGY

In this section, we first provide an overview of the YOLOv5 architecture, explaining its key components and functionality. Then, we describe the proposed improved method based on YOLOv5, aimed at enhancing both the accuracy and speed of the model.

2.1. Yolov5 Network Architecture

Due to the challenges of accurately detecting vehicles around the shovel and the importance of preventing collisions in industrial environments, we improved the YOLOv5 model. However, in this section, we aim to explain the default structure of YOLOv5.

As previously mentioned, YOLO detectors are primarily deep learning networks developed for object detection tasks. These networks offer higher inference speeds compared to other models, making them more suitable for real-time object detection requirements. YOLO networks have been improved

over successive versions. Since YOLOv5 is faster and performs better than its predecessors, our proposed model was developed based on YOLOv5 v6.1, which was released in 2022.

In this section, we explain the default structure of this version of YOLOv5. The main difference between YOLOv5 and other YOLO models lies in its improved architecture, implementation, and advanced features. This model, as a faster and more optimized version of its predecessors, offers improved performance with advanced features.

YOLOv5 was updated in 2022 by Ultralytics, and with improvements in its architecture and implementation, it has become a suitable option for industrial applications such as mining. These enhancements have made YOLOv5 more accurate in object detection and faster in inference compared to previous models.

YOLOv5 is an improved version of the YOLO network, continuing the core idea of the YOLO series in algorithm design. YOLOv5 consists of four main parts: input, backbone, neck, and head, with its architecture shown in Fig. 1.

The image to be detected is processed through an input layer and then sent to the backbone for feature extraction. The backbone generates feature maps of varying sizes, which are then combined through the feature fusion network (neck) to ultimately produce three feature maps, P3, P4, and P5, with sizes 80×80, 40×40, and 20×20, respectively, for detecting small, medium, and large objects in the image.

The head section is responsible for object detection and classification. After the three feature maps are sent to the head, a multi-dimensional array is inferred, containing the object class, class probability, coordinates, and box width and height information. Then, a post-processing operation is applied to filter out irrelevant information, including a confidence threshold to select boxes with probabilities above the threshold and a non-maximum suppression algorithm to select the box with the highest probability from the chosen boxes.

The backbone consists of several modules, including (Conv+BatchNorm+SiLU) or ConBNSILU, C3 modules, and finally an SPPF module. The ConBNSiLU module is used to assist the C3 module in feature extraction, while the SPPF module, a fast spatial pyramid pooling layer, eliminates the constraint of fixed input size, meaning the network no longer requires a fixed-size image. Specifically, an SPPF layer is added on top of the final convolution layer. The SPPF layer combines features and produces fixed-length outputs that are then fed into fully connected layers.

In the backbone, YOLOv5's most important module is C3, whose core idea is derived from CSP-

Net. This layer ensures the capability to extract features in the backbone by removing redundant gradient information. The neck section uses the PAN method, a feature fusion path from bottom to top, which is employed to improve detection accuracy across different object scales.

The fundamental mechanism of all YOLO versions is the same: images are divided into cells of equal size, and each cell is responsible for detecting objects whose centers fall within the cell. The primary differences between various YOLO versions lie in the backbone and neck components.

The only distinction between the different versions of YOLOv5 lies in the number of layers and parameters. As these increases, they influence the training time and accuracy.

In Table 1, Framework, Backbone, and Year of Release are listed for each YOLO version. By understanding the overall structure of the YOLOv5 network, the next section demonstrates how we have modified and improved its real-time performance to accurately detect vehicles around the shovel.

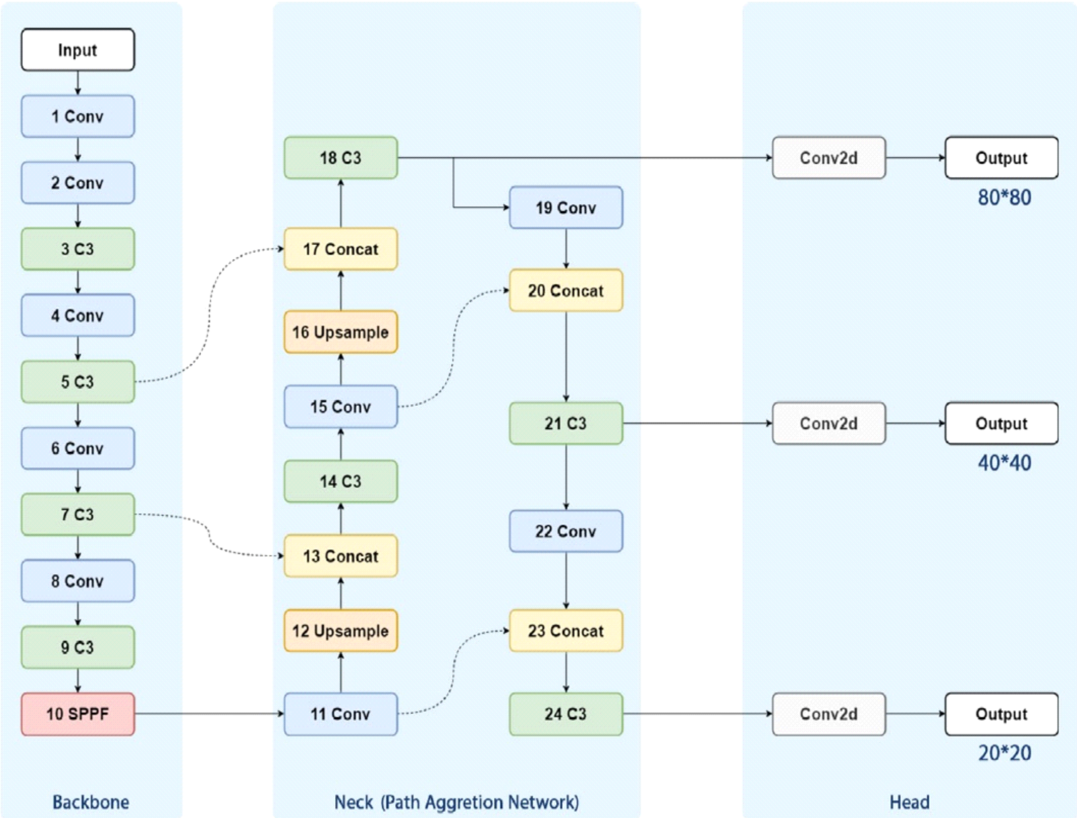


Fig. 1. YOLOv5 Block Diagram.

Table 1. Summary of YOLO architectures

Backbone	Framework	Anchor	Date	Model
Darknet24	Darknet	No	2015	YOLO
Darknet24	Darknet	Yes	2016	YOLOv2
Darknet53	Darknet	Yes	2018	YOLOv3
CSP Darknet53	Darknet	Yes	2020	YOLOv4
YOLOv5 CSP Darknet	Pytorch	Yes	2020	YOLOv5
EfficientRep	Pytorch	Yes	2022	YOLOv6
Extended ELAN	Pytorch	Yes	2022	YOLOv7
CSPDarknet (c2f)	Pytorch	No	2023	YOLOv8

2.2. Proposed Yolov5 Architecture

In this section, the proposed network based on the YOLOv5 architecture is explained in detail. The main goal of this research is to develop an accurate model for real-time vehicle detection around shovels in copper mines. The achievements of the proposed network could serve as an inspiration for accurate and real-time detection of similar objects.

To improve the trade-off between accuracy and speed, the ECA channel attention module has been used to modify the network's backbone. In the proposed architecture, all C3 modules in the backbone are replaced with the ECA attention module, and the SGD optimizer is used for training. The ECA channel attention module, as shown in Fig. 2, takes the W*H*C features extracted from the previous layer and transforms them into a 1*1*C tensor through a Global Average Pooling (GAP) operation.

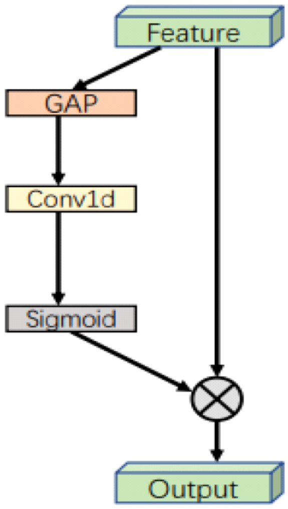


Fig. 2. Attention module architecture ECA [36].

This way, a global average of the features is obtained for each channel, which is then filtered by a 1D convolutional layer with learnable weights. Following this, a sigmoid activation function is applied to generate weights for the channels based on their importance. Finally, the obtained weights, which form a 1*1*C tensor, are multiplied by the input features. The output is a feature tensor with dimensions W*H*C, where the channels are weighted according to their importance. The main idea behind using ECA is that this module introduces an efficient attention mechanism that helps the network focus more effectively on relevant features of the input images while reducing unnecessary computations. The C3 module, which is the default in YOLOv5, is a simple convolutional block, while ECA adaptively adjusts the responses of channel features.

This allows the model to have a better ability to identify objects under various conditions, while also making more efficient use of computational resources. One of the main challenges in real-time object detection, especially in complex environments like around shovels in mines, is balancing accuracy and processing speed. The presence of dynamic and cluttered backgrounds, varying lighting conditions, and fast-moving vehicles makes it difficult for traditional models to maintain consistent performance without imposing excessive computational overhead. By replacing C3 with ECA, the network is able to focus more effectively on the important features of the scene, leading to improved detection accuracy, particularly for smaller vehicles or those partially obscured. This ultimately results in an overall improvement in the model’s performance. At the same time, ECA introduces a lightweight and low-cost attention mechanism, reducing the complexity compared to other attention-based models, which is crucial for maintaining the fast-processing speed required for real-time applications.

In mining environments, where shovels move frequently and rapidly, having a fast and accurate model is essential to ensure that no vehicle is missed and safety is maintained at all times. In summary, by integrating the ECA module into the YOLOv5 architecture, we have developed a more effective and efficient model for real-time vehicle detection around shovels. This improvement enhances both detection accuracy and processing speed, which are critical factors for timely vehicle identification and maintaining the safety of personnel in the mining environment. Fig. 3 shows the block diagram of the enhanced YOLOv5 network.

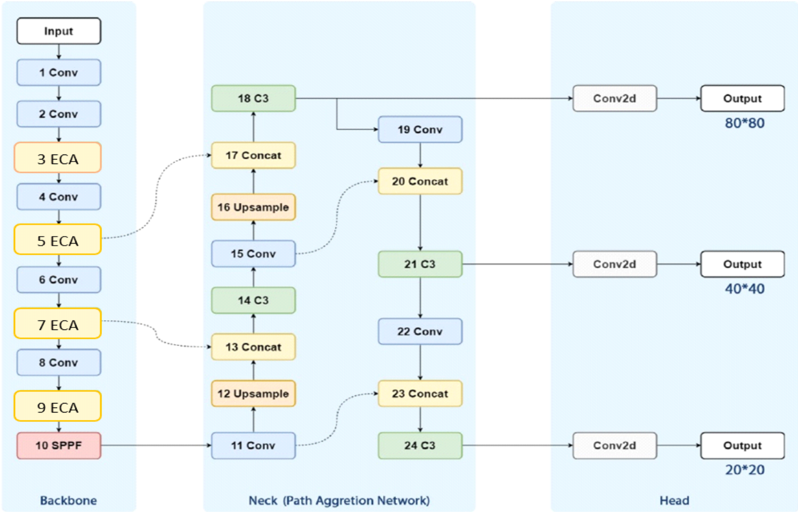


Fig. 3. Improved YOLOv5 block diagram.

3. EVALUATION OF RESULT

In this section, the dataset used to evaluate the proposed algorithm is first described. Subsequently, the evaluation metrics employed to assess the results are introduced. Finally, the obtained results are reported and analyzed.

The dataset used in this study initially included 600 raw images captured by four cameras positioned around a mining shovel. Techniques such as brightness reduction, contrast adjustment, and artificial blurring were applied to augment the training set, expanding it to 1,200 images. Seventy percent of the images were used as the training set, 20% as the validation set, and 10% as the test set.

The test set remained unprocessed to reflect real-world conditions. All machinery near the shovel was grouped into a single class, but future expansions will differentiate specific equipment types and introduce a human operator class.

A few examples of the images from the dataset are illustrated in Fig. 4. The simulations were performed on the Ubuntu 20.04 operating system using an Intel® Xeon(R) Silver 4210 CPU @ 2.20GHz × 20, 128 GB of RAM, and an NVIDIA GeForce RTX 3090 graphics card.

While the proposed system is primarily focused on detecting machinery around shovels, its flexible YOLOv5-based architecture and channel attention mechanisms inherently support generalization to other mining equipment, such as trucks or drilling rigs. This adaptability is achievable without structural redesign, through installing cameras in the blind spots of new equipment, collecting operational data from the target environment, and retraining the model.

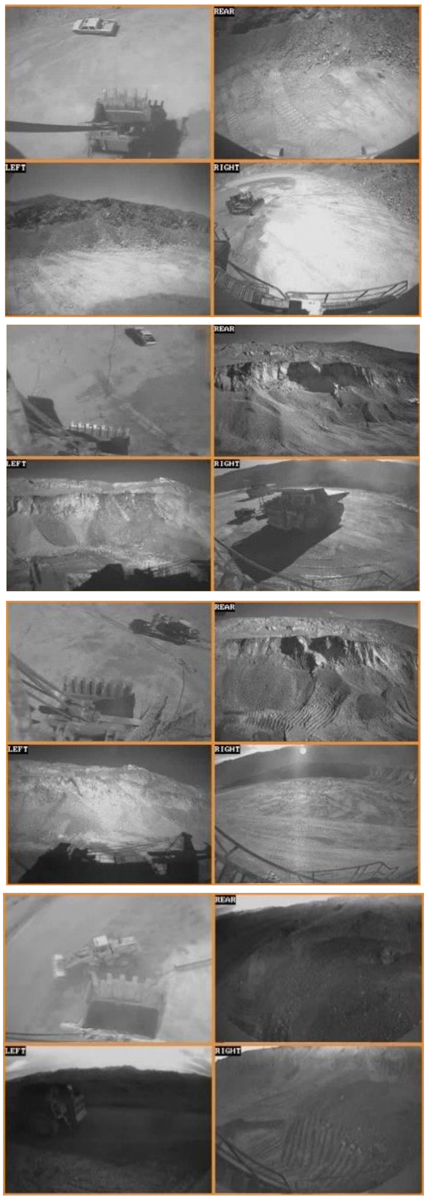


Fig. 4. Sample instances from the dataset used.

Although no specific empirical tests have been conducted on non-shovel machinery, the model's success in complex environments (e.g., dust, occlusion) indicates its readiness for broader applications. The cost and time required for this extension are minimized due to the reuse of the core system and the absence of architectural modifications.

One of the most critical steps after designing and developing a model or algorithm is evaluating its performance. Sensitivity (true positive rate) and specificity or detectability (true negative rate) are two key metrics for statistically assessing the performance of classification results. When data can be divided into two groups, positive and negative, the performance of an experiment that categorizes information into these two groups can be measured and described using sensitivity and specificity indices. To measure performance, four parameters are required[45].

True Positive (TP): Correctly identified.

False Positive (FP): Incorrectly identified.

True Negative (TN): Correctly rejected.

False Negative (FN): Incorrectly rejected.

One of the evaluation metrics for the proposed method is accuracy. This metric represents the ratio of the number of correct predictions made for samples of a specific class to the total number of predictions for samples of the same class.

High values of accuracy indicate a low number of data points incorrectly classified into a specific class. This metric is calculated using the formula in Eq. (1).

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

The next metric for evaluation is recall. This metric represents the ratio of the number of correctly classified data points in a specific class to the total number of data points that should have been classified into that class.

High values for this metric indicate a low number of data points that were incorrectly excluded from that specific class, and it is calculated using the formula in Eq. (2).

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

The primary metric for evaluating the accuracy of object detection models is the mean Average Precision (mAP), which is calculated based on the Average Precision (AP) across various classes, as shown in Eq. (3). Here, N represents the number of classes. In general, APi indicates the average precision of class i over different IoU thresholds.

These thresholds typically range from 0.5 to 0.95, increasing in steps of 0.05. We utilize the mAP50 metric (threshold = 0.5) to assess the accuracy of the proposed model.

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i \tag{3}$$

Table 2 compares the frame rate (FPS) between the proposed model and the baseline YOLOv5 models. As shown in the table, the proposed method outperforms YOLOv5M, YOLOv5L, and YOLOv5X with a frame rate of 159 FPS. This increase in processing speed offers several important advantages for industrial applications. A higher frame rate allows the system to process images faster and send necessary alerts to the operator. This is especially crucial in dynamic environments like mines, where shovels and vehicles are constantly moving. Additionally, industrial systems need to minimize the time between detection and action to prevent accidents.

The 159 FPS rate of the proposed model minimizes processing delays and reduces the likelihood of accidents caused by such delays. Furthermore, the increased processing speed enables simultaneous analysis of multiple cameras installed around the mining site, which contributes to better coverage of blind spots and improved safety.

Table 2. Comparison of the proposed model with several YOLOv5 baseline models in terms of speed

FPS	Model
80	YOLOv5M
65	YOLOv5L
48	YOLOv5X
159	Proposed model

Table 3 presents the evaluation results for the baseline YOLOv5 models and the proposed model on the dataset of images created in this research, using input images with dimensions of 640×640. The results are shown based on various metrics, including precision, recall, and mAP50.

As observed in Table 3, the proposed method achieves higher accuracy compared to all baseline models, with its mAP50 values being 4.7%, 4.4%, and 5.1% higher than YOLOv5M, YOLOv5L, and YOLOv5X, respectively. The results, with an average precision of 94.2%, indicate an improvement in the detection accuracy of the proposed model compared to the baseline YOLOv5 models.

Therefore, by making modifications to the baseline model, the proposed network's detection capability is enhanced, leading to more accurate detection in practical applications.

Table 3. Comparison of the proposed model with several YOLOv5 base3line models in terms of accuracy

R	P	mAP50	Model
93.4	84.2	89.5	YOLOv5M
92.3	85.7	89.8	YOLOv5L
93.2	84.1	89.1	YOLOv5X
96.1	90.6	94.2	Proposed model

As previously mentioned, one of the most critical challenges in industrial environments, including mines, is ensuring the prevention of collisions between shovels and surrounding machinery. The operator's limited visibility, particularly due to blind spots and the specific structure of the shovel, often causes nearby machinery to remain undetected.

This limitation significantly increases the risk of collisions, potentially leading to severe accidents. To mitigate these risks, a rapid detection system with a high frame rate (FPS) is essential. A high FPS ensures that the detection of machinery is performed quickly and without delay, allowing timely warnings to be issued to the operator.

Given the constant movement of shovels and surrounding machinery, even slight delays in detection can dramatically increase the likelihood of accidents. Our proposed method, based on the optimized YOLOv5 model, has demonstrated its ability to accurately detect machinery in all four directions around the shovel using real-world data collected from a copper mining environment.

This precise and rapid detection effectively eliminates the blind spots of the shovel, minimizing the probability of collisions. Furthermore, the proposed system directly enhances the safety of personnel and prevents damage to expensive mining equipment.

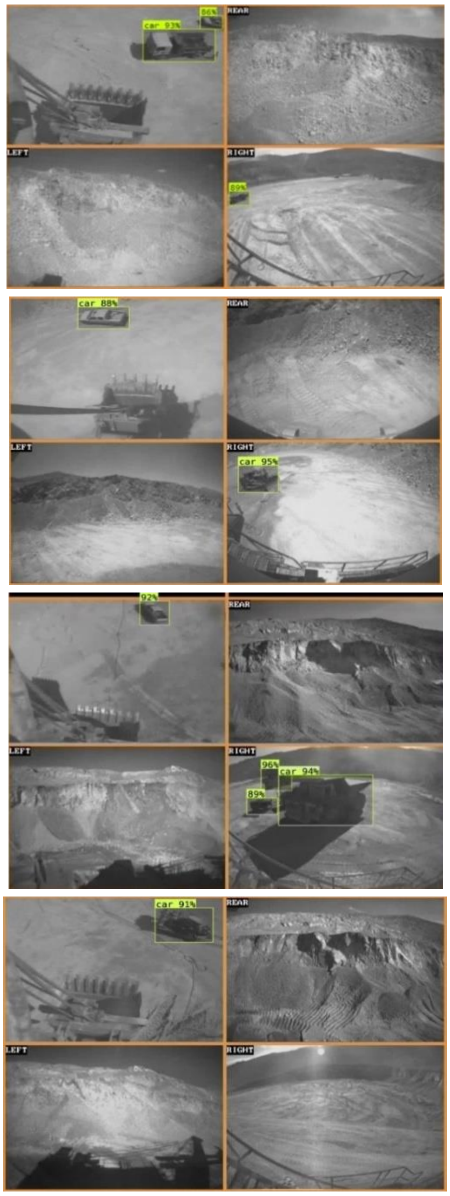
In the final evaluation section of the results, the proposed method was compared with other object detection methods. The results show that the mAP of the proposed method is 94.2, which is higher than that of other models. As shown in Table 4, this demonstrates its better performance in detecting vehicles around the shovel in the copper mine environment.

This improvement in accuracy and performance reflects the high reliability of the proposed model in the complex and variable environmental conditions of the mine, making it an effective tool for monitoring and safety in such environments.

v. 5 shows the output of the proposed model on images from the copper mine dataset. As observed, the proposed model has successfully detected the vehicles around the shovel, indicating the high efficiency of the proposed method in accurately identifying vehicles in the mining environment.

Table 4. Comparison of mAP values for different object detection methods

mAP50	Model
57.3	Faster R-CNN
58.4	Retinanet
54.9	SSD
82.3	YOLOv3
75.8	YOLOv4
89.4	YOLOv5
94.2	Proposed model



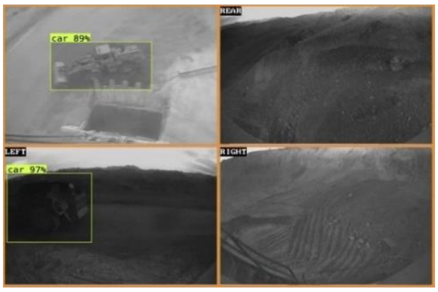


Fig. 5. Detection results of the proposed model on the test dataset.

4. CONCLUSION

The findings of this research demonstrated that our proposed method, based on an improved YOLOv5 model, achieved high accuracy in detecting machinery around shovels in copper mining environments. The achieved accuracy of 94% highlights the effectiveness of the model in accurately identifying machinery and its robust performance under complex and real-world conditions. With its rapid and precise image processing capability, the system can identify machinery around the shovel and prevent collisions, thereby directly enhancing workplace safety and protecting personnel.

The proposed model can serve as a vital safety support system by detecting surrounding machinery and alerting operators to potential hazards. This capability is particularly beneficial in scenarios where the operator's visibility is limited, significantly reducing risks and preventing accidents. To further improve system performance under challenging environmental conditions, such as fog, rain, or low-light scenarios, it is recommended to integrate LiDAR sensors with the object detection model. LiDAR can provide accurate distance measurements and create detailed 3D maps of the surrounding environment, complementing the vision-based system. This integration can enhance the system's accuracy and reliability, making it more effective in adverse conditions. In the future, this technology can be extended to other industrial environments, such as other mining operations or transportation industries. These systems not only improve workplace safety but also enhance operational efficiency and reduce costs associated with accidents, contributing to better overall industrial performance.

REFERENCES

[1] Kusi-Sarpong, S., Bai, C., Sarkis, J., and Wang, X. (2015). Green supply chain practices evaluation in the mining industry using a joint rough sets and fuzzy TOPSIS methodology. Resources Policy, 46, 86–100.

[2] Holmberg, K., Kivikytö-Reponen, P., Härkisaari, P., Valtonen, K., and Erdemir, A. (2017). Global energy consumption due to friction and wear in the mining industry. Tribology International, 115, 116–139.

[3] Ranängen, H., and Lindman, Å. (2017). A path towards sustainability for the Nordic mining industry. Journal of Cleaner Production, 151, 43–52.

[4] Sanmiquel, L., and Rossell, J. M., and Vintró, C. (2015). Study of spanish mining accidents using data mining techniques. Safety Science, 75, 49–55.

[5] Permana, H. (2012). Risk assessment as a strategy to prevent mine accidents in Indonesian mining. Revista de Minería. 18(4): 43–49.

[6] Kizil, M. (2003). Virtual reality applications in the Australian minerals industry. Application of computers and operations research in the minerals industries, South African. 569-574.

[7] Kazan, E., and Usmen, M. A. (2018). Worker safety and injury severity analysis of earthmoving equipment accidents. Journal of safety research, 65, 73-81.

[8] Squelch, A. P. (2001). Virtual reality for mine safety training in South Africa. Journal of the Southern African Institute of Mining and Metallurgy, 101(4): 209-216.

[9] Kia, K., Fitch, S. M., Newsom, S. A., and Kim, J. H. (2020). Effect of whole-body vibration exposures on physiological stresses: Mining heavy equipment applications. Applied ergonomics, 85, 103065 .

[10] Domingues, M. S., Baptista, A. L., and Diogo, M. T. (2017). Engineering complex systems applied to risk management in the mining industry. International journal of mining science and technology, 27(4): 611-616.

[11] Tichon, J., and Burgess-Limerick, R. (2011). A review of virtual reality as a medium for safety related training in mining. Journal of Health and Safety Research and Practice, 3(1): 33-40.

[12] Shahmoradi, J., Talebi, E., Roghanchi, P., andHassanalian, M. (2020). A comprehensive review of applications of drone technology in the mining industry. Drones, 4(3): 34.

[13] Patrucco, M., Pira, E., Pentimalli, S., Nebbia, R., and Sorlini, A. (2021). Anti-collision systems in tunneling to improve effectiveness and safety in a system-quality approach: A review of the state of the art. Infrastructures, 6(3): 42.

[14] Kim, H., and Choi, Y. (2021). Autonomous driving robot that drives and returns along a planned route in underground mines by recognizing road signs. Applied Sciences, 11(21): 10235 .

[15] Backman, S., Lindmark, D., Bodin, K., Servin, M., Mörk, J., and Löfgren, H. (2021). Continuous control of an underground loader using deep reinforcement learning. Machines, 9(10): 216.

[16] Kim, H., and Choi, Y. (2020). Comparison of three location estimation methods of an autonomous driving robot for underground mines. Applied Sciences, 10(14): 4831.

[17] Lööw, J., Abrahamsson, L., and Johansson, J. (2019). Mining 4.0—The impact of new technology from a work place perspective. Mining, Metallurgy and Exploration, 36, 701-707 .

[18] Sishi, M., andTelukdarie, A. (2020). Implementation of Industry 4.0 technologies in the mining industry-a case study. International Journal of Mining and Mineral Engineering, 11(1): 1-22.

[19] Sánchez, F., and Hartlieb, P. (2020). Innovation in the mining industry: Technological trends and a case study of the challenges of disruptive innovation. Mining, Metallurgy and Exploration, 37(5): 1385-1399.

[20] Groves, W. A., Kecojevic, V. J., andKomljenovic, D. (2007). Analysis of fatalities and injuries involving mining equipment. Journal of safety research. 38(4): 461-470.

[21] Amponsah-Tawiah, K., Jain, A., Leka, S., Hollis, D., and Cox, T. (2013). Examining psychosocial and physical hazards in the Ghanaian mining industry and their implications for employees' safety experience. Journal of safety research. 45: 75-84.

[22] Ruff, T., Coleman, P., and Martini, L. (2011). Machine-related injuries in the US mining industry and priorities for safety research. International journal of injury control and safety promotion. 18(1): 11-20 .

[23] Imam, M., Baïna, K., Tabii, Y., Ressami, E. M., Adlaoui, Y., Benzakour, I., and Abdelwahed, E. H. (2023). The future of mine safety: a comprehensive review of anti-collision systems based on computer vision in underground mines. Sensors. 23(9): 4294.

[24] Ali, D., and Frimpong, S. (2020). Artificial intelligence, machine learning and process automation: Existing knowledge frontier and way forward for mining sector. Artificial Intelligence Review. 53(8): 6025-6042 .

[25] Hyder, Z., Siau, K., and Nah, F. (2019). Artificial intelligence, machine learning, and autonomous technologies in mining industry. Journal of Database Management (JDM). 30(2): 67-79.

[26] Viola, P., and Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. In Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001, Vol. 1, pp. I-I .

[27] Dalal, N., and Triggs, B. (2005). Histograms of oriented gradients for human detection. In 2005 IEEE computer society conference on computer vision and pattern recognition. CVPR'05, Vol. 1, pp. 886-893 .

[28] Felzenszwalb, P., McAllester, D., and Ramanan, D. (2008). A discriminatively trained, multiscale, deformable part model. In 2008 IEEE conference on computer vision and pattern recognition. pp. 1-8.

[29] Ren, S., He, K., Girshick, R., and Sun, J. (2016). Faster R-CNN: Towards real-time object detection with region proposal networks. IEEE transactions on pattern analysis and machine intelligence. 39(6): 1137-1149.

[30] Zou, Z., Chen, K., Shi, Z., Guo, Y., and Ye, J. (2023). Object detection in 20 years: A survey. Proceedings of the IEEE. 111(3): 257-276.

[31] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., and Berg, A. C. (2016). Ssd: Single shot multibox detector. In Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14, Springer International Publishing. pp. 21-37.

[32] Redmon, J. (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.

[33] Lin, T., Goyal, P., Girshick, R., He, K., andDollár, P. (2017). Focal loss for dense object detection. In Proceedings of the IEEE International Conference on Computer Vision. pp. 2980–2988.

[34] Duan, K., Bai, S., Xie, L., Qi, H., Huang, Q., and Tian, Q. (2019). Centernet: Keypoint triplets for object detection. In Proceedings of the IEEE/CVF international conference on computer vision. pp. 6569-6578 .

[35] Redmon, J., and Farhadi, A. (2017). YOLO9000: better , faster, stronger. In Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 7263-7271.

[36] Redmon, J. (2018). Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767.

[37] Bochkovskiy, A., Wang, C. Y., and Liao, H. Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934 .

[38] Jocher, G., Chaurasia, A., Stoken, A., Borovec, J., Kwon, Y., Fang, J., Michael, K., Montes, D., Nadar, J. and Skalski, P. (2022). ultralytics/yolov5: v6. 1-

tensorrt, tensorflow edge tpu and openvino export and inference. Zenodo.

[39] Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L. and ;Desmaison, A. (2019). Pytorch: An imperative style, high-performance deep learning library. Advances in neural information processing systems, 32.

[40] Elfwing, S., Uchibe, E., andDoya, K. (2018). Sigmoid-weighted linear units for neural network function approximation in reinforcement learning. Neural networks, 107, 3-11.

[41] Li, C., Li, L., Jiang, H., Weng, K., Geng, Y., Li, L., Ke, Z., Li, Q., Cheng, M., Nie, W. and Li, Y. (2022). YOLOv6: A single-stage object detection framework for industrial applications. arXiv preprint arXiv:2209.02976.

[42] Wang, C. Y., Bochkovskiy, A., and Liao, H. Y. M. (2023). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 7464-7475.

[43] Jocher, G., Chaurasia, A., and Qiu, J. (2023). Ultralytics YOLOv8 (version 8.0. 0).

[44] Zhang, S., Yang, H., Yang, C., Yuan, W., Li, X., Wang, X., Zhang, Y., Cai, X., Sheng, Y., Deng, X. and Huang, W. (2023). Edge device detection of tea leaves with one bud and two leaves based on shuffleNetv2-YOLOv5-lite-E. Agronomy, 13(2), 577.

[45] Powers, D. M. (2020). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. arXiv preprint arXiv:2010.16061.