

Article

Computer Vision for Safety Management in the Steel Industry

Roy Lan , Ibukun Awolusi * and Jiannan Cai

School of Civil & Environmental Engineering, and Construction Management, The University of Texas at San Antonio, San Antonio, TX 78249, USA; roy.lan@my.utsa.edu (R.L.); jiannan.cai@utsa.edu (J.C.)

* Correspondence: ibukun.awolusi@utsa.edu

Abstract: The complex nature of the steel manufacturing environment, characterized by different types of hazards from materials and large machinery, makes the need for objective and automated monitoring very critical to replace the traditional methods, which are manual and subjective. This study explores the feasibility of implementing computer vision for safety management in steel manufacturing, with a case study implementation for automated hard hat detection. The research combines hazard characterization, technology assessment, and a pilot case study. First, a comprehensive review of steel manufacturing hazards was conducted, followed by the application of TOPSIS, a multi-criteria decision analysis method, to select a candidate computer vision system from eight commercially available systems. This pilot study evaluated YOLOv5m, YOLOv8m, and YOLOv9c models on 703 grayscale images from a steel mini-mill, assessing performance through precision, recall, F1-score, mAP, specificity, and AUC metrics. Results showed high overall accuracy in hard hat detection, with YOLOv9c slightly outperforming others, particularly in detecting safety violations. Challenges emerged in handling class imbalance and accurately identifying absent hard hats, especially given grayscale imagery limitations. Despite these challenges, this study affirms the feasibility of computer vision-based safety management in steel manufacturing, providing a foundation for future automated safety monitoring systems. Findings underscore the need for larger, diverse datasets and advanced techniques to address industry-specific complexities, paving the way for enhanced workplace safety in challenging industrial environments.

Keywords: computer vision; steel manufacturing; hazard identification; site monitoring; deep learning



Citation: Lan, R.; Awolusi, I.; Cai, J. Computer Vision for Safety Management in the Steel Industry. *AI* **2024**, *5*, 1192–1215. <https://doi.org/10.3390/ai5030058>

Academic Editors: Giovanni Diraco and Arslan Munir

Received: 22 May 2024
Revised: 6 July 2024
Accepted: 16 July 2024
Published: 19 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Steel is an extremely versatile product commonly used in essential industries, including construction, transportation, energy, and manufacturing. It is considered the most important engineering and construction material. The process of steel manufacturing, however, presents a challenging and hazardous work environment. Workers frequently interact with heavy machinery and are exposed to toxic gases and numerous safety hazards. In 2022, data collected by World Steel from 55 organizations covering 60% of its membership reported 18,448 injuries and 90 fatalities among workers, including both employees and contractors. The primary causes of these incidents were identified as slips/trips/falls and interactions with moving machinery [1]. Injuries common in this industry include traumatic brain injuries (TBI), burns, amputations, spinal cord injuries, and carpal tunnel syndrome.

Numerous literature reviews underscore the importance of early hazard detection in significantly reducing these incidents [2,3]. The steel industry has proactively adopted various safety measures, such as comprehensive training programs, innovative maintenance and operation strategies, technological advancements for automating repetitive tasks, and consistent updates on safety protocols [4,5]. Despite these efforts, there remains an urgent need to advance the methods of hazard identification in steel mills by shifting from manual to automated processes, thereby reducing human error and enhancing safety interventions.

Computer vision, empowered by deep learning, offers robust capabilities for automating tasks like detection, tracking, monitoring, and action recognition, which can

be specifically tailored to the steel manufacturing sector [6–8]. For instance, identifying hazards such as proximity to heavy equipment can mitigate fatalities caused by moving machinery [9,10]. The use of personal protective equipment (PPE) is another critical area where computer vision can make a significant impact. According to the Occupational Safety and Health Administration (OSHA), proper PPE usage can prevent up to 37.6% of occupational injuries and diseases. Additionally, the failure to wear PPE contributes to 12–14% of occupational injuries leading to total disability [11]. The Centers for Disease Control and Prevention (CDC) note that among all industries, steelworkers are particularly susceptible to traumatic brain injuries (TBIs) often resulting from slips/trips/falls. Wearing a safety hard hat can reduce the likelihood of a TBI by 70% in such incidents [12]. In large steel manufacturing sites, manually ensuring comprehensive compliance with safety hard hat regulations is unfeasible. Here, computer vision's object detection capabilities, enhanced by deep learning, can automate and improve the precision of safety hard hat detection.

While significant progress has been made in applying automated hazard identification and safety monitoring in industries like construction and agriculture through computer vision and deep learning, a system specifically designed and tested for the steel manufacturing industry is yet to be realized. This gap highlights the potential for impactful advancements in this field. The objective of this research is to explore the feasibility and potential of implementing computer vision technologies for safety management within the steel manufacturing industry. This is executed through a pilot case study focused on the utilization of computer vision-based deep learning technology, specifically designed to automatically detect the use of hard hats by steelworkers. To achieve this objective, a review phase characterizing hazards with computer vision application was conducted, and then a multi-criteria decision model was deployed in selecting commercially available computer vision programs for application toward safety in steel manufacturing. This analysis is crucial to determine the most appropriate computer vision system for effective safety application within the steel manufacturing industry. Findings from this study could demonstrate the efficacy of CV in enhancing safety management in the steel industry, suggesting its broader applicability in high-risk sectors. This research exemplifies the practical use of CV in safety management at actual steel manufacturing sites and could pave the way for integrating more advanced technologies in industrial safety practices.

2. Background

2.1. Overview of the Steelmaking Process

Steel production predominantly employs either the integrated mill, known as the Basic Oxygen Furnace (BOF), or the mini-mill, termed the Electric Arc Furnace (EAF) method [13]. These methodologies diverge in their choice of raw materials and the inherent risks associated with each. The BOF process utilizes iron ore and coke, initially processed in a blast furnace and subsequently in an oxygen converter, where an exothermic reaction removes impurities like carbon, silicon, manganese, and phosphorus. In contrast, mini-mills primarily use scrap metal or direct reduced iron (DRI), processed in the EAF [14].

While both methods necessitate rigorous safety management, there are notable operational differences. Integrated mills, generally larger, surpass mini-mills in steel production [15]. EAF facilities, smaller and less economical, differ from the BOF by their intermittent operation, offering more flexibility [16]. Despite its lower scale and output, the safety management in mini-mills is critical and often underestimated when compared to BOF processes [14,15]. The increasing focus on steel recycling and greenhouse emissions reduction has amplified the adoption of EAFs, which are capable of using 100% scrap steel to produce carbon steel and alloys [17,18]. This trend underscores the importance of robust safety management in mini-mills.

2.2. Occupational Hazards in Steel Manufacturing

Steelworkers encounter numerous safety risks inherent to their profession [5]. Monitoring and reducing injury and fatality statistics are crucial. The steel industry primarily utilizes the Lost Time Injury Frequency Rate (LTIFR) and Fatality Frequency Rate (FFR) to gauge these metrics. LTIFR represents work-related incidents causing disability to employees or contractors, preventing them from performing their duties, expressed as the number of lost time injuries per million hours worked [19]. A notable decrease in LTIFR from 4.55 in 2006 to 0.81 in 2021, a reduction of 82%, has been reported by the World Steel Association [19]. The goal remains to further minimize these incidents toward a zero-incident rate. FFR, conversely, tracks fatalities among company and contractor employees. This index is critical as it represents loss of life and is often viewed as a lagging, rather than proactive, safety measure [20,21]. A proactive approach to reducing LTIFR can indirectly impact FFR. Recent data show an increase in FFR from 0.021 in 2019 to 0.03 in 2021, a 43% rise, highlighting the need for enhanced safety protocols [19].

The steelmaking process, with its complex operations, poses various hazards. These risks, akin to those in construction, agriculture, and general manufacturing, are categorized by the International Labor Organization [22] into physical, chemical, safety, and ergonomic hazards. Physical hazards encompass noise, vibration, extreme heat, and radiation. Chemical hazards include exposure to harmful gases and substances. Safety hazards, notably slips, trips, falls, proximity to heavy machinery, and risks from falling or flying objects, are predominant injury causes in the steel industry [19]. Kifle et al. [23] identified such incidents as the leading injury causes in this sector.

2.3. Current Safety Practices in the Steel Manufacturing Industry

The steel manufacturing industry has demonstrated significant progress in enhancing workplace safety, as evidenced by the decline in Lost Time Injury Frequency Rates (LTIFRs) over the past decade [19]. The National Institute for Occupational Safety and Health [24] identifies five hierarchical levels of hazard controls, which are ranked by effectiveness: personal protective equipment (PPE) controls, administrative controls, engineering controls, substitution, and elimination. These measures are increasingly being adopted in steel manufacturing sites.

For instance, administrative controls like the Lock-out/Tagout/Tryout (LOTO) procedures are crucial in preventing accidents during equipment maintenance. The rigorous use of PPE, a fundamental safety practice, has been evolving with innovations in the Internet of Things (IoT) and wearable technologies and safety incentives [25]. This evolution has introduced smart PPEs, such as helmets, bracelets, and belts, providing real-time safety feedback, a potential boon for the steel industry [23,26–29]. Additionally, workers training in hazard identification and safe working practices has advanced, with interactive simulators and the use of virtual and augmented reality technologies proving effective safety measures [30,31]. NIOSH posits that hazard elimination or substitution represents the most effective control mechanisms, as they remove or replace the hazard entirely. However, given the nature of steel manufacturing, complete elimination or substitution of certain hazards is impractical without ceasing specific operations. Thus, the focus shifts to engineering controls that leverage technology to distance workers from hazards. Implementing these controls might include modifying existing equipment or introducing advanced technologies like computer vision to automate certain tasks, thus enhancing worker safety.

2.4. Computer Vision Applications in Safety Management and Review of YOLO Models

Computer vision (CV), an interdisciplinary domain, capitalizes on advanced technologies for analyzing visual data, such as images and videos, to extract meaningful insights in real time. This technology, particularly beneficial in securing work environments by isolating workers from hazards, is recognized as a potent form of engineering control [32,33]. Object detection is a critical task in computer vision, aiming to identify and locate objects within an image or video. Among the various approaches developed, the You Only Look Once (YOLO) family of models has garnered significant attention for its real-time object detection capabilities. The original YOLO model, introduced by Joseph Redmon et al. [34], marked a significant shift by framing object detection as a single regression problem rather than a classification problem. This novel approach enabled real-time detection, distinguishing YOLO from other object detection frameworks such as R-CNN and Fast R-CNN.

YOLOv2, also known as YOLO9000, introduced enhancements, including batch normalization, a high-resolution classifier, and multi-scale training, allowing it to detect over 9000 object categories using a joint training algorithm that combined detection and classification datasets [35]. YOLOv3 further advanced the architecture by adopting a multi-scale prediction strategy with residual connections and predicting bounding boxes at three different scales, which improved its capability to detect smaller objects. YOLOv4, developed by Alexey Bochkovskiy et al. [36], optimized the model for single GPU systems by integrating advanced techniques such as CSPDarknet53 as the backbone, PANet path aggregation, and the use of the Mish activation function, setting new benchmarks in speed and accuracy. YOLOv5, released by Ultralytics, introduced major improvements in usability and performance, incorporating features like auto-learning bounding box anchors, mosaic data augmentation, and integrated hyperparameter evolution. YOLOv5 is available in various versions (small, medium, and large), catering to different resource constraints and accuracy requirements [37]. Subsequent versions, with minor distinctions from version 5, such as YOLOv6 and YOLOv7, focused on efficiency and accuracy improvements, while YOLOv8 introduced architectural changes that further refined detection capabilities, particularly in challenging scenarios with occlusions and complex backgrounds. YOLOv8 provided five scaled versions: YOLOv8n (nano), YOLOv8s (small), YOLOv8m (medium), YOLOv8l (large), and YOLOv8x (extra-large). YOLOv8 supports multiple vision tasks such as object detection, segmentation, pose estimation, tracking, and classification [37]. YOLOv9, on the other hand, leverages hybrid neural network architectures, enhanced data augmentation techniques, and improved loss functions to deliver unprecedented accuracy and robustness [38].

The architecture of YOLO models has evolved significantly, incorporating various innovations to enhance their detection capabilities. The backbone network, responsible for extracting feature maps from the input image, has seen a progression from Darknet-19 in earlier versions to deeper and more complex architectures like CSPDarknet53 [37,38]. The neck component, which aggregates features from different stages of the backbone, often uses PANet and FPN architectures to enhance detection. The head of the YOLO model generates the final predictions, including bounding boxes and class probabilities, with YOLOv3 introducing multi-scale predictions refined in subsequent versions. Figure 1 shows the architecture of modern object detectors, including the backbone, the neck, and the head. Figure 2 shows an example of different tasks that could be performed on YOLO models, a scenario in the object detection task where the non-compliant worker (no_hardhat) is highlighted with a red bounding box. These models use a combination of localization, confidence, and classification loss functions, with advances in loss functions like Complete Intersection over Union (CIoU) and Distance Intersection over Union (DIoU), improving model convergence and accuracy.

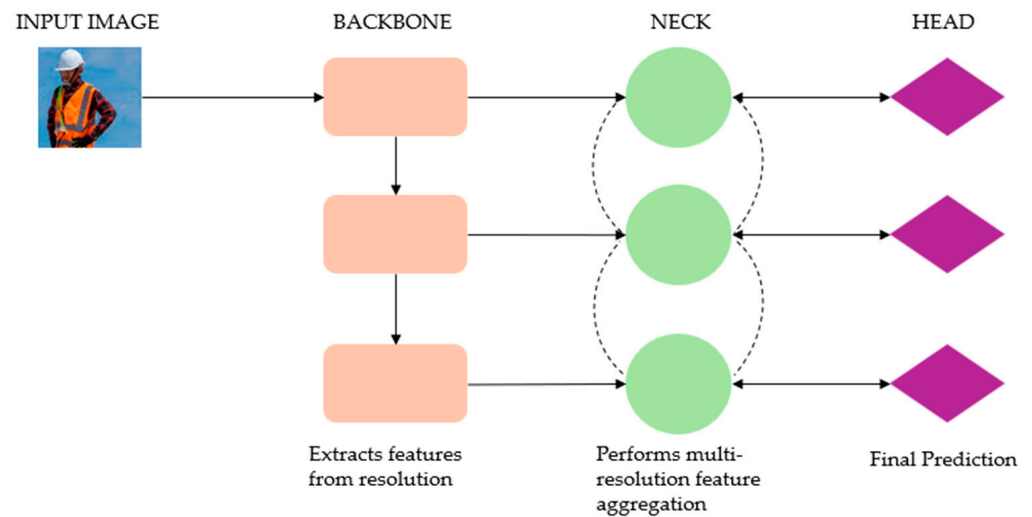


Figure 1. The architecture of advanced object detectors depicting the backbone, the neck, and the head.

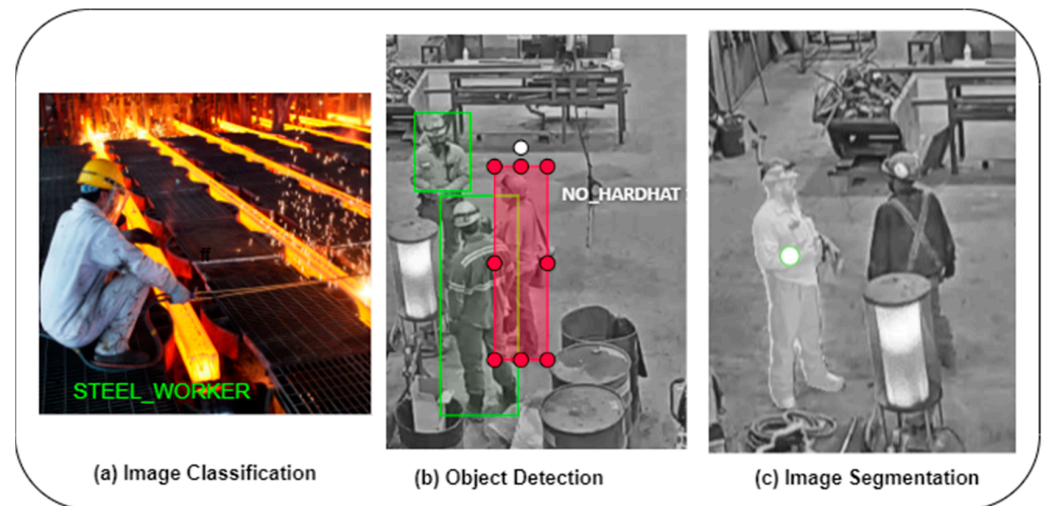


Figure 2. CV task examples.

YOLO networks have been applied across various domains, demonstrating their versatility and effectiveness. In autonomous driving, YOLO models are used for real-time object detection, enabling the detection of pedestrians, vehicles, traffic signs, and obstacles [39]. In security and surveillance, they facilitate real-time monitoring and threat detection [40]. In healthcare, YOLO networks assist in medical imaging tasks such as tumor detection, organ segmentation, and surgical tool localization [41]. In agriculture, they are employed for crop monitoring, pest detection, and yield estimation, enhancing precision farming practices [42]. Table 1 shows a summary of other domains in which YOLO networks have been implemented.

Table 1. Summary of YOLO applications in various domains.

Use Case	Sector	Classes	YOLO Network	Number of Images	Performance	Source
Detection of construction equipment	Construction	4	YOLO v5 & v8	4800	mAP@0.5–0.951	[43]
Detection of sidewalk cracks	Construction	4	YOLO v2, v3 and v4-tiny	4000	Accuracy 0.94	[44]
Detection of cherry fruits	Agriculture	3	YOLO v4	400	F1-score 0.947	[45]
Heat staking process inspection	Automobile industry	3	YOLO v5	3000	mAP@0.95–0.95	[46]
Assembly component identification	Aerospace	150	YOLO v5	9450	mAP@0.5–0.99	[47]
Surface defects on steel surface	Steel	6	YOLO v3	4057	mAP@0.95–0.72	[48]
Warehouse robot detection	Logistics	2	YOLO v8	335	mAP@0.95–0.86	[49]

2.5. Research Need Statement

The application of CV systems in the steel manufacturing industry, particularly for safety management, remains underexplored in the existing literature. Current studies predominantly focus on enhancing productivity through CV applications in areas like quality control, monitoring of continuous casting processes, defect detection on steel surfaces, and breakout prediction [50–52]. These studies, while valuable, do not specifically address safety management within the steel manufacturing industry. Given the dynamic and hazardous nature of steel manufacturing, along with the variety of risks present in its processes, there is a clear need for dedicated research on the potential of CV systems to improve industrial safety in this sector. This paper aims to bridge this gap by presenting pioneering work that could serve as a foundation for future studies focused on developing CV applications for safety management in steel manufacturing.

This research contributes significantly to the existing body of knowledge by evaluating commercially available CV programs suitable for safety management in the steel industry. The findings offer a comprehensive overview of safety-oriented CV systems that can be applied in this sector, providing a template for steel manufacturers interested in integrating CV technology into their safety management protocols. Additionally, the paper showcases a practical implementation of CV for detecting personal protective equipment (PPE) in a steel manufacturing context, with a case study focusing on safety hard hat detection. This practical application demonstrates the feasibility of using CV for safety management in steel manufacturing.

3. Materials and Methods

This study aimed to test the feasibility of CV application for safety management in steel manufacturing by conducting a pilot case study that leverages the approach of computer vision-based deep learning technology to automatically detect hard hats on steelworkers. To achieve this objective, a review phase characterizing hazards with computer vision application was conducted, and then a multi-criteria decision model was deployed in selecting commercially available computer vision programs for application toward safety in steel manufacturing. Figure 3 shows the research process for this study.

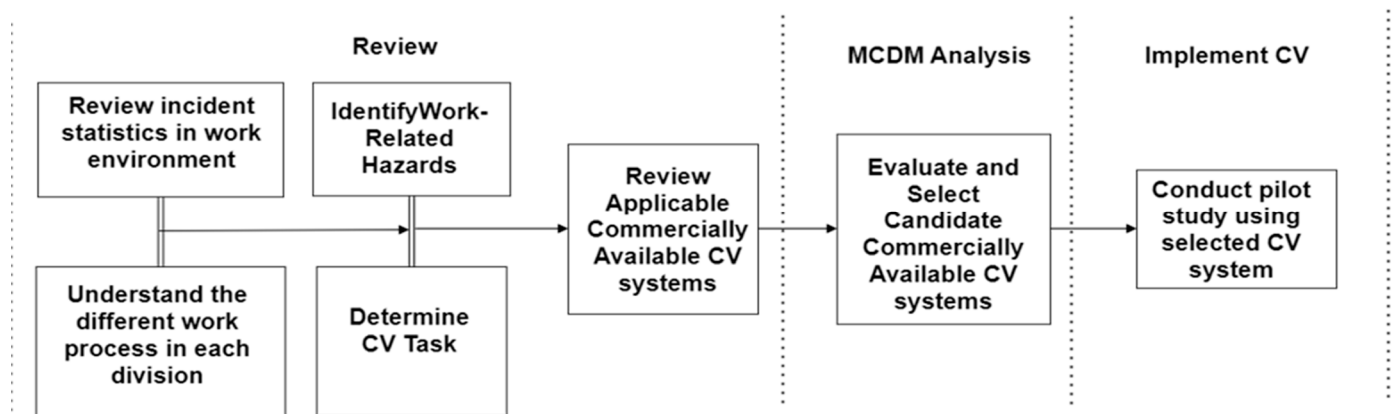


Figure 3. Research process.

3.1. Characterization of Computer Vision Applications for Safety in Steel Manufacturing

To effectively characterize the application of computer vision (CV) for safety in steel manufacturing, especially in a mini-mill context, this study's approach encompassed a comprehensive understanding of the various work processes involved. This endeavor unfolded in phases: review, observation, and characterization. During the review phase, the research team scrutinized reports and academic papers on steel mini-mill operations. The team further delved into the mini-mills' operational standards, machinery types, operation modes, and key safety metrics like the Lost Time Injury Frequency Rate (LTIFR) and Fatality Frequency Rate (FFR). Significantly, the Association for Iron & Steel Technology (AIST) steel wheel application proved invaluable. It offered a systematic and visual overview of standard mini-mill operations, enhancing our conjectural understanding of these processes.

Building on this foundational knowledge, the research team sought an empirical, practical perspective through a live tour of a steel mini-mill. During this visit, the mill's management team led the research team through various departments and work processes. Delving deeper, group interviews, focusing on gathering narratives about current safety hazards, existing safety measures, and their shortcomings, were conducted as suggested by Bolderston [53]. The interview, lasting approximately 70 min, was a collaborative effort, with team members actively sharing notes and highlighting key points, aligning with the methodology outlined by Guest [54]. The full-day tour, combined with the initial review, provided a robust foundation for the research approach. It enabled a thorough understanding of each work process in the steel mini-mill and identified the inherent hazards faced by workers. With this understanding, each identified hazard could be effectively aligned with corresponding CV tasks.

3.2. Evaluation and Selection of Commercially Available CV Systems for Safety Management

To conduct this pilot case study of detecting workers' compliance with safety hard hats in the steel mini-mill, there was a need to determine the computer vision system that would serve as the detection system for this pilot study.

3.2.1. Computer Vision System Search

The research team conducted an extensive online search to identify commercially available computer vision (CV) systems suitable for the pilot case study in steel manufacturing safety. Utilizing Google, key search phrases such as "computer vision companies", "artificial intelligence in health and safety", "commercially available object detection companies", and "artificial intelligence and workplace safety" were employed for the search. This search strategy yielded a vast array of CV systems. To refine these results, the focus was exclusively on CV systems that demonstrate potential for use in safety applications. Additionally, to enhance the depth and relevance of the CV system search, recommendations from industry professionals were sought, following the approach recommended by

Creswell [55]. This dual strategy of combining an internet search with expert consultations allowed for the identification of the most suitable and effective CV systems relative to this study's scope, ensuring that the selected systems were both commercially available and directly applicable to workplace safety.

3.2.2. Computer Vision System Evaluation and Selection

The research identified eight operational off-the-shelf CV systems suitable for safety management in the steel industry. To select the most appropriate system for this pilot study, a multi-criteria decision-making (MCDM) method was utilized. MCDMs are essential tools for decision-makers, aiding in choosing the best option among multiple alternatives based on various criteria. Common MCDMs include the Analytical Hierarchy Process (AHP), Analytical Network Process (ANP), Multi-objective Optimization on the basis of Ratio Analysis (MOORA), and the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS), with the selection often depending on the decision maker's preferences and objectives [56].

For this study, TOPSIS, a method praised for its straightforwardness, ease of interpretation, and efficiency in identifying the best alternative, as well as visualizing differences between alternatives using normalized values, was used [57,58]. Originating from Hwang and Yoon in 1981 and further modified since TOPSIS's core principle is to select the best option by measuring Euclidean distances, it aims to minimize the distance to the ideal alternative (PIS) and maximize the distance from the non-ideal alternative (NIS). The optimal choice is the one closest to the PIS and farthest from the NIS. Figure 4 of this study illustrates the flow process of the TOPSIS model, outlining the systematic approach we employed to select the most suitable CV system for our pilot study in the steel industry.

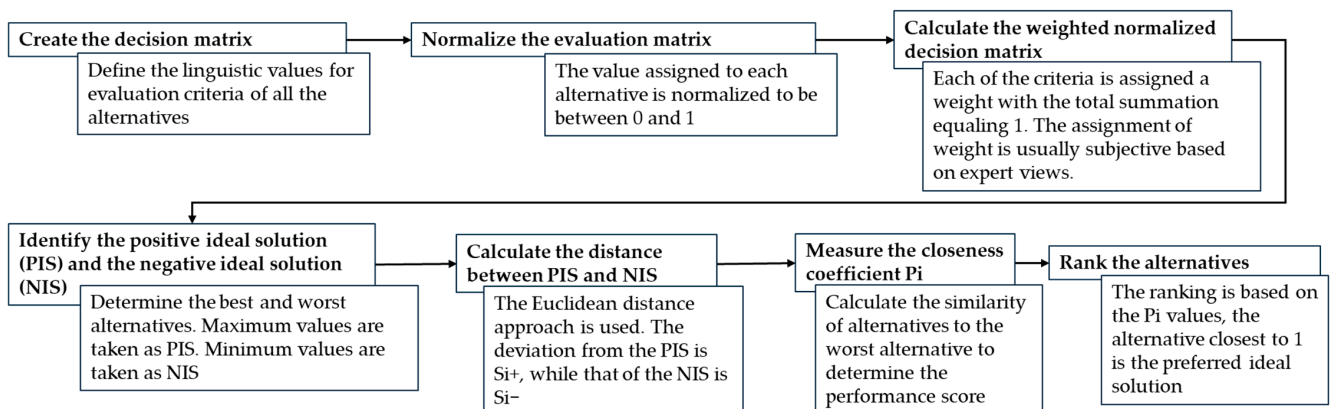


Figure 4. TOPSIS process for evaluating off-the-shelf CV program.

The following steps explain the TOPSIS method implemented in this study:

Step 1: Create the decision matrix consisting of alternatives (the eight commercially available computer vision programs) and criteria (PPE detection, data privacy, proximity to heavy equipment, slips, trips, falls, etc.). Also, define the importance weights of the criteria. Linguistic values for “available” or “not available” were set for the criteria.

Decision matrix $X = X_{ij}$. Weight factor $W = [w_1, w_2, w_3, \dots, w_n]$

$$\text{And } w_1 + w_2 + w_3 + \dots + w_n = 1.$$

Step 2: Normalize the evaluation matrix.

This is to ensure all data are in the same unit. The formulas in Equations (1)–(3) [56,59,60] below are used to calculate the normalized values:

$$n_{ij} = \frac{X_{ij}}{\sqrt{\sum_{i=1}^m X_{ij}^2}} \quad (1)$$

$$n_{ij} = \frac{X_{ij}}{\max_i X_{ij}} \quad (2)$$

$$n_{ij} = \begin{cases} \frac{x_{ij} - \min_i x_{ij}}{\max_i x_{ij} - \min_i x_{ij}} \\ \frac{\max_i x_{ij} - x_{ij}}{\max_i x_{ij} - \min_i x_{ij}} \end{cases} \quad (3)$$

Step 3: Calculate the weighted normalized decision matrix.

This is the product of the weights with the normalized values given by Equation (4) [56,59,60].

$$v_{ij} = w_j n_{ij} \quad (4)$$

For $i = 1, \dots, m; j = 1, \dots, n$. Function criteria are either costs or benefits; in this case, “availability” is the benefit function yielding toward a better alternative.

Step 4: Identify the positive and negative ideal solution. This is carried out by determining the best and the worst alternatives for each evaluation criterion (i.e., the maximum and minimum values) among all the CV programs.

The positive ideal solution is denoted as V^+ in Equation (5) [56,59,60].

$$V^+ = (v_1^+, v_2^+, \dots, v_n^+) = [\max_i v_{ij} | j \in I], [\min_i v_{ij} | j \in J] \quad (5)$$

Meanwhile, the negative ideal solution is denoted as V^- in Equation (6).

$$V^- = (v_1^-, v_2^-, \dots, v_n^-) = [\min_i v_{ij} | j \in I], [\max_i v_{ij} | j \in J] \quad (6)$$

Step 5: Calculate the distance between the PIS and NIS.

The Euclidean distance is measured, the deviation from PIS is S_i^+ , while the deviation from NIS is S_i^- . This is shown in Equations (7) and (8) [56,59,60].

$$S_i^+ = \sqrt{\sum_{j=1}^n (v_{ij} - v_j^+)^2}, i = 1, 2, \dots, m \quad (7)$$

$$S_i^- = \sqrt{\sum_{j=1}^n (v_{ij} - v_j^-)^2}, i = 1, 2, \dots, m \quad (8)$$

Step 6: Measure the closeness coefficient p_i .

Using values obtained from p_i , the alternatives are ranked; in this case, the p_i value closest to 1 is the best alternative among the options. To determine p_i , Equation (9) is used [56,59,60]:

$$P_i = \frac{S_i^-}{S_i^- + S_i^+} \quad (9)$$

3.3. Pilot Study on Safety Hard Hat Detection

3.3.1. Pilot Study Context

There has been extensive research on the application of computer vision (CV) in various industries, as evidenced by the literature review. However, the steel industry remains underrepresented in studies focusing on CV applications for safety management. This gap is particularly significant given the unique challenges that the steel industry faces, such as a volatile working environment, a shortage of training data, and concerns about data privacy among industry stakeholders. This study aims to address these challenges by serving as an introductory feasibility pilot study on the application of computer vision for safety management in the steel manufacturing industry. The primary focus is on assessing the practicality and feasibility of implementing CV systems in this context rather than on optimizing performance metrics. To conduct this feasibility pilot study, the candidate CV system utilized was selected through an MCDM analysis. This study specifically

explores the detection of safety hard hats on steelworkers. Using the selected CV system, a quantitative analysis was performed to evaluate its effectiveness and feasibility in the steel manufacturing environment. By focusing on feasibility, this study provides a foundational outlook on the potential of CV applications in the steel industry, paving the way for more detailed and performance-oriented research in the future. The results and insights gained from this pilot study are crucial for understanding the practical implications and readiness of CV systems for enhancing safety management in steel manufacturing.

3.3.2. Detection Models

Three pre-trained YOLO variants—YOLOv5m, YOLOv8m, and YOLOv9c were utilized and compared for detecting safety hard hats in the steel manufacturing industry. YOLOv5m, known for its balance of speed and accuracy, incorporates CSPNet for efficient gradient flow, PANet for multi-scale feature fusion, and SPPF for enhanced feature extraction. YOLOv8m builds on this foundation with CSPDarknet53, combining PANet and FPN for improved multi-scale detection and advanced augmentation techniques like MixUp and CutMix to enhance generalization, while YOLOv9c integrates CSPResNeXt for robust feature extraction, BiFPN for optimal feature fusion, and adaptive detection methods for improved localization and reduced false positives. These models collectively demonstrate significant advancements in detection performance, offering speed, accuracy, and robustness, making them ideal for real-time safety applications in industrial settings. The comparison of the three models can be seen in Table 2.

Table 2. Comparison between YOLO architectures [37].

Features	YOLOv5	YOLOv8	YOLOv9
Network Type	Fully Connected	Fully Connected Custom	Fully Connected
Backbone for Feature Extraction	CSPDarknet53	CSPDarknet53 Backbone Cross-stage Partial Connection	Generalized Efficient Layer Aggregation Network (GLEAN)
Neck	Path Aggregation Network (PANet)	Path Aggregation Network (PANet)	Programmable Gradient Information (PGI)
Head	YOLOv5 Head (Three Detection Layers)	YOLOv8 Head (Improved Anchor-free)	YOLOv8 Head (Adaptive Anchor-free)

3.3.3. Dataset Collection and Processing

In conducting the quantitative analysis for this study, a dataset comprising 703 meticulously labeled images from a steel manufacturing site was utilized. The images were extracted from five hours of CCTV footage, specifically from the maintenance ('maint') area of a steel mini-mill, which also includes a storeroom section. This particular area was selected due to its high frequency of worker activity, rendering it an optimal site for data collection within the constrained observation window. The labeling process was executed using the Computer Vision Annotation Tool (CVAT), an interactive video and image annotation tool designed for computer vision applications.

To assess the impact of fine-tuning on the pre-trained primary OD model, it was retrained using the small yet diverse dataset of 703 images. Diversity was introduced to the dataset through various augmentation techniques, including rotation, horizontal flipping, and adjustments in grayscale brightness and saturation. This strategy was critical to avert the overfitting of the model to the dataset. The scarcity of instances depicting safety hard hat violations in the CCTV data necessitated the generation of artificial images embodying such infractions. This was accomplished through the implementation of the stable diffusion inpainting technique, as expounded by Rombach et al. [61], with the result exhibited in Figure 5. A K-fold distribution (K = 5) was employed to enhance the robustness of the model evaluation. For each fold, the dataset was split into training (80%) and validation

(20%) subsets to facilitate comprehensive model assessment. Figure 6 shows a sample of the training dataset.



Figure 5. Representation showing dataset augmentation with stable diffusion inpainting, including (A) original data, (B) augmented data, (C) original data, and (D) augmented data.



Figure 6. Training dataset from steel mini-mill site.

3.3.4. Evaluation Metrics

To evaluate the performance of the object detection (OD) models, the following metrics were used: precision, recall, F1-score, average precision (AP), specificity, and area under the curve (AUC) [62,63]. To detect workers' compliance with wearing safety hard hats, not wearing a hard hat was considered the positive class. Where TP represents correct predictions of a person not wearing a hard hat, FP represents incorrect predictions of a person wearing a hard hat, and FN represents incorrect predictions of a person not wearing a hard hat. An Intersection over Union (IoU) threshold of 0.5 was used to determine the identification of these parameters in the confusion matrix.

4. Results and Discussion

The results of this study are systematically outlined in the subsequent sections, following the research methodology. These sections detail hazard characterization, the application of the TOPSIS technique, and the analysis of data from this pilot case study. This structured approach offers a comprehensive view of this study’s findings, from identifying and categorizing hazards to applying TOPSIS for their assessment and culminating in the empirical insights gained from this case study.

4.1. Safety Hazard Characterization and TOPSIS Analysis

4.1.1. Safety Hazard Characterization for CV Applications in Steel Manufacturing

This study rigorously identifies hazards inherent in each stage of the mini-mill steel manufacturing process and demonstrates how CV can be strategically employed to monitor these hazards. This approach aims to provide early warnings, thereby mitigating the risks of injuries, illnesses, and fatalities. The hazard assessment begins at the shredding site, a phase where scrap metals are processed, segregating ferrous from non-ferrous metals. It is noteworthy that not all mini-mills include a shredding phase; however, this study encompasses it for comprehensive analysis. The process continues with the transportation of the ferrous metal to the Electric Arc Furnace (EAF) for melting. The subsequent stage involves purification at the Ladle Metallurgical Station (LMS), followed by the solidification of the molten steel into semi-finished forms like billets, slabs, or blooms at the continuous caster. These semi-finished products undergo further processing in the rolling mill, where they are transformed into finished steel products through various methods, including annealing, hot forming, cold rolling, pickling, galvanizing, coating, or painting. The manufacturing cycle concludes with the finishing and transportation of the final products.

Table 3 in this study provides a detailed mapping of CV tasks and implementation processes to the identified hazards in each work process at the mini-mill. The goal is to establish a framework where CV can effectively track, detect, or monitor these hazards. Figure 7 in this study visually encapsulates the entire steel manufacturing workflow, highlighting the pivotal role of CV in augmenting safety throughout the process. By aligning CV tasks with specific hazards at each stage, this study offers a pragmatic approach to enhancing safety in the dynamic environment of steel mini-mills.

Table 3. Safety hazard characterization for CV applications in steel manufacturing.

Safety Hazard	CV Task	Data Collection Device	Hazard Sample Scenarios	Computer Vision Application	Source
Struck-by	Object tracking, object detection	RGB Cameras, RGBD Cameras, Stereo Vision Cameras	(1) In mini-mills with shredders on site, workers are exposed to flying objects as raw materials are shredded and crushed. (2) Moving equipment such as dump trucks or gantry cranes can strike workers. (6) Workers can be struck by finished products, shears, and bends rebars and by the trucks used for loading.	CV can track personnel in real time and issue alerts when workers are in proximity to the danger of being struck.	[64]
Caught in-between	Object tracking, object detection	RGB Cameras, RGBD Cameras, Stereo Vision Cameras	(1) In mini-mills with shredders on site, workers are exposed to the danger of being caught in between equipment, the shredder mill, and raw materials at the shredding yard. (2) There is a danger of workers being caught between loads transferred by a crane (within its swing radius) and stationary or moving objects in the steel shop. (6) Similarly, during the finishing and transportation, there is a hazard of workers being caught between trucks and objects, which could be the stacked rebars.	CV tasks can help in classifying, detecting, and tracking people and equipment, therefore alerting when workers are in proximity to these hazardous scenarios.	[65]

Table 3. Cont.

Safety Hazard	CV Task	Data Collection Device	Hazard Sample Scenarios	Computer Vision Application	Source
Fire	Object detection, object classification	RGB Cameras, RGBD Cameras, Infrared Cameras, Flash LIDAR	(2) There is a very high risk of fire hazards near furnaces. Molten steel is usually at very high temperatures > 2800 F. This increases the occurrence of fire hazards. (3) The LMS is a hotspot that includes molten steel undergoing its purification process, alloying, desulphurization, degassing, etc. (4) The solidification of liquid steel involves working at high temperatures, and there is a possibility of a fire hazard in this work process.	CV technologies can successfully detect fire at its incipient stage. Most of the time, it is earlier than smoke detectors due to its early detection using DL programs.	[66]
Spills	Object classification, segmentation, and detection	RGB Camera, RGBD Cameras	(2) Spills from the molten metal at the furnace can pose a hazard to workers. (3) Spill hazard is also prominent during the tapping process from EAF to the refining station. (4) There is the likelihood of spills of molten steel, especially during the transfer of molten steel from the Ladle to the Tundish.	CV using segmentation tasks can detect these spills and alert workers in real time when in proximity to the spills.	[67]
Slips/trips/falls	Object detection, action recognition	RGB Cameras, RGBD Cameras	(2) Slip, trip, and fall occurrences are high at the furnace due to the nature of activities during this process. There is usually reduced visibility at this location due to the excess heat from the furnace, which increases the possibility of trips and falls. (5) Workers may trip at the rolling mill due to improperly stacked rolling equipment. (6) Workers are exposed to the danger of slips, trips, and falls during the finishing and transportation work process; stacked rebars can cause this, too.	CV can detect workers' proximity to the hazards causing slips, trips, and falls and detect when workers fall; this is viable in real time when integrated with sensors that can be integrated into safety vests, smart watches, or hard hats.	[68]
Bad worker posture	Pose estimation/action recognition	RGB Cameras, RGBD Cameras, Stereo Vision Cameras	This hazard is recognized and characterized in all phases of the work process. Bad worker posture is observed in all work areas that involve workers' presence. There is an increasing need to observe human interaction with tools, equipment, and the environment.	CV techniques using images and videos can detect workers' postures in every work process and send alerts when workers' posture poses a risk to their safety.	[69]
No PPE	Classification, object detection	RGB Cameras, RGBD Cameras	Personal protective equipment (PPE) is required at all work processes in the mini-mill. These protective outfits are to be worn at all work process locations.	CV using images can detect workers not wearing their PPEs, including safety helmets, vests, gloves, and, in some cases, glasses.	[70]
Extreme temperature	Image classification, segmentation, and object detection	Infrared Cameras (IR)	The entire work process in the mini-mill steel process subjects and endangers workers to extremely high temperatures, which needs to be critically monitored.	CV using heat maps on face detection of workers can detect, based on varying color codes, when workers are experiencing heat stress.	[71]

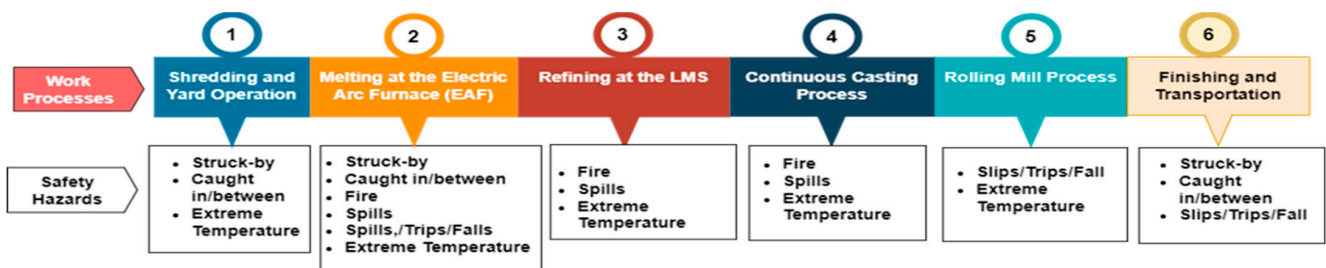


Figure 7. Identified hazards at work processes in steel mini-mill manufacturing.

4.1.2. TOPSIS Analysis

With the result of the characterization of hazards, the need to select a CV system that would be deployed in conducting the pilot case study was imperative. TOPSIS technique, as described in the methodology, was deployed for the analysis of the eight (8) computer vision programs (alternatives) considered for evaluation, with the objective of selecting one. The eight CV systems enlisted were Everguard, Intenseye, Cogniac, Protex, Rhyton, Chooch, Kogniz, and Matroid. These alternatives were appraised on eleven (11) criteria: PPE detection, data privacy, ergonomics, health, geofencing, proximity to heavy equipment, slips/trips/falls, application in steel manufacturing real-time processing, user-friendliness of graphical user interface (GUI), and versatility in other applications. These criteria were identified from the literature review and selected based on their respective influence in achieving satisfactory safety management in steel manufacturing. Tables 4–8 present the results of the TOPSIS MCDM approach used for analysis. Table 4 shows the defined linguistic values for each of the evaluation criteria, with 1 and 2 denoting “Not Available” and “Available”, respectively. Table 5 presents the elements of the matrix, showing the relationship between the criteria and the alternatives with their assigned weights. This assignment of weights was subjectively determined by the research team’s safety experience and the respective criteria’s impact on the research objective. Table 6 shows the normalized values of the evaluation matrix. Table 7 shows the weighted values of the evaluation matrix, including PIS and NIS.

Table 4. Definition of linguistic values for evaluation criteria.

Linguistic Value	PPED	DP	OA	EG	H	GF	PHE	STF	AS	RT	GUI
Available = 2	2	2	2	2	2	2	2	2	2	2	2
Not Available = 1	1	1	1	1	1	1	1	1	1	1	1

PPED = PPE detection; DP = data privacy; OA = other application; EG = ergonomics; H = health; GF = geofencing; PHE = proximity to heavy equipment; STF = slips, trips, and falls; AS = application in steel manufacturing; RT = real time; GUI = user-friendliness of graphical user interface.

Table 5. Elements of the evaluation.

CV System	PPED	DP	OA	EG	H	GF	PHE	STF	AS	RT	GUI
Everguard	2	2	1	2	2	2	2	2	2	2	2
Intenseye	2	2	1	2	1	1	2	2	1	2	2
Cogniac	2	2	2	1	1	1	1	1	1	2	2
Protex	2	2	1	1	1	1	1	1	1	2	2
Rhyton	2	1	1	2	1	1	2	1	1	1	1
Chooch	2	2	2	1	1	2	2	1	1	2	2
Kogniz	2	1	2	1	1	1	2	2	1	2	2
Matroid	2	2	2	1	1	2	1	1	1	2	2
Weights	0.1	0.05	0.1	0.1	0.05	0.1	0.1	0.1	0.15	0.1	0.05

Table 6. Normalized values of the evaluation matrix.

CV System	PPED	DP	OA	EG	H	GF	PHE	STF	AS	RT	GUI
Everguard	0.354	0.392	0.224	0.485	0.603	0.485	0.417	0.485	0.603	0.371	0.371
Intenseye	0.354	0.392	0.224	0.485	0.302	0.243	0.417	0.485	0.302	0.371	0.371
Cogniac	0.354	0.392	0.447	0.243	0.302	0.243	0.209	0.243	0.302	0.371	0.371
Protex	0.354	0.392	0.224	0.243	0.302	0.243	0.209	0.243	0.302	0.371	0.371
Rhyton	0.354	0.196	0.224	0.485	0.302	0.243	0.417	0.243	0.302	0.186	0.186
Chooch	0.354	0.392	0.447	0.243	0.302	0.485	0.417	0.243	0.302	0.371	0.371
Kogniz	0.354	0.196	0.447	0.243	0.302	0.243	0.417	0.485	0.302	0.371	0.371
Matroid	0.354	0.392	0.447	0.243	0.302	0.485	0.209	0.243	0.302	0.371	0.371

Table 7. Weighted values of the evaluation matrix showing PIS and NIS.

CV System	PPED	DP	OA	EG	H	GF	PHE	STF	AS	RT	GUI
Everguard	0.035	0.020	0.022	0.049	0.030	0.049	0.042	0.049	0.090	0.037	0.019
Intenseye	0.035	0.020	0.022	0.049	0.015	0.024	0.042	0.049	0.045	0.037	0.019
Cogniac	0.035	0.020	0.045	0.024	0.015	0.024	0.021	0.024	0.045	0.037	0.019
Protex	0.035	0.020	0.022	0.024	0.015	0.024	0.021	0.024	0.045	0.037	0.019
Rhyton	0.035	0.010	0.022	0.049	0.015	0.024	0.042	0.024	0.045	0.019	0.009
Chooch	0.035	0.020	0.045	0.024	0.015	0.049	0.042	0.024	0.045	0.037	0.019
Kogniz	0.035	0.010	0.045	0.024	0.015	0.024	0.042	0.049	0.045	0.037	0.019
Matroid	0.035	0.020	0.045	0.024	0.015	0.049	0.021	0.024	0.045	0.037	0.019
V ⁺	0.035	0.020	0.045	0.049	0.030	0.049	0.042	0.049	0.090	0.037	0.019
V ⁻	0.035	0.010	0.022	0.024	0.015	0.024	0.021	0.024	0.045	0.019	0.009

Table 8. Separation distance measure, closeness ratio, and ranking of alternatives.

CV System	S+	S-	Pi	Ranking
Everguard	0.022	0.071	0.760	1st
Intenseye	0.058	0.046	0.444	2nd
Cogniac	0.067	0.032	0.324	6th
Protex	0.071	0.023	0.246	8th
Rhyton	0.067	0.032	0.323	7th
Chooch	0.059	0.045	0.435	3rd
Kogniz	0.060	0.044	0.426	4th
Matroid	0.062	0.040	0.392	5th

Table 8 shows the separation distance measure from the PIS and the NIS, the closeness ratio, and the ranking of alternatives. The closeness ratio to the PIS was the determinant of which alternative was selected. Out of the eight CV systems, Everguard, Intenseye, and Chooch were ranked 1st, 2nd, and 3rd, respectively, with a closeness ratio of 0.760, 0.444, and 0.435. Therefore, based on the examined criteria, this MCDM analysis selected Everguard as the candidate selected CV system among the evaluated systems that could be deployed for safety management in steel manufacturing.

4.2. Pilot Case Study Results: Safety Hard Hat Detection Using Candidate CV System

4.2.1. Experimental Environment

The network model used for analysis was developed using the Python programming language in conjunction with the PyTorch deep learning library. PyTorch provided the necessary framework to input the annotated data into the Python interface, facilitating the process of updating weights and performing calculations essential for the analysis. The specific configuration of the environment, including the CPU, GPU, and other relevant libraries and tools, is detailed in Table 9. All training images were resized to meet the 640-input size. These datasets underwent training for 50 epochs with batch sizes of 16. The learning rates were set at 0.001. Constant values for the momentum, box loss gain, and optimizer choice were maintained. Utilizing the Stochastic Gradient Descent (SGD) optimizer with a momentum of 0.937, based on Gupta et al. [72], who noted that although SGD converges slower and its gradients are uniformly scaled, its lower training error leads to better generalization, which is especially beneficial for test data. The training was executed using an NVIDIA 4090 ×1 GPU. This setup was instrumental in ensuring the smooth execution and accurate processing of the object detection task.

Table 9. Configuration of the modeling system and environment.

Parameter Name	Configuration
CPU	Intel i9-13900K × 1
GPU	NVIDIA 4090 × 1
RAM	64 GB
Language	Python
Operating System	Ubuntu 22.04

4.2.2. Detection Results across Models

Precision, Recall, F-1 Score, and mAP Results

To ensure a robust evaluation of the models’ performance, a five-fold cross-validation approach was employed. This method allows for a comprehensive assessment of the models’ ability to generalize to unseen data, providing a more reliable estimate of their real-world performance. Three state-of-the-art YOLO variants—YOLOv5m, YOLOv8m, and YOLOv9c were evaluated on a steel mill dataset comprising 703 labeled images.

The evaluation involved partitioning the dataset into five equally sized subsets, where each subset served as a test set once while the remaining subsets were used for training. This process was repeated five times, ensuring that every image was used for both training and validation. The performance metrics for each fold were averaged to provide an overall estimate of the models’ effectiveness. A visual comparison of the models’ predictions against the ground truth labels in an instance is shown in Figure 8. The ground truth images represent the actual safety compliance labels, serving as the benchmark. The detection results for each model highlight instances of workers wearing hard hats (“in_hardhat”), not wearing hard hats (“no_hardhat”), and cases where workers are not visible (“invisible”). High confidence scores are observed across various models in these instances, reflecting the models’ predictive capabilities.

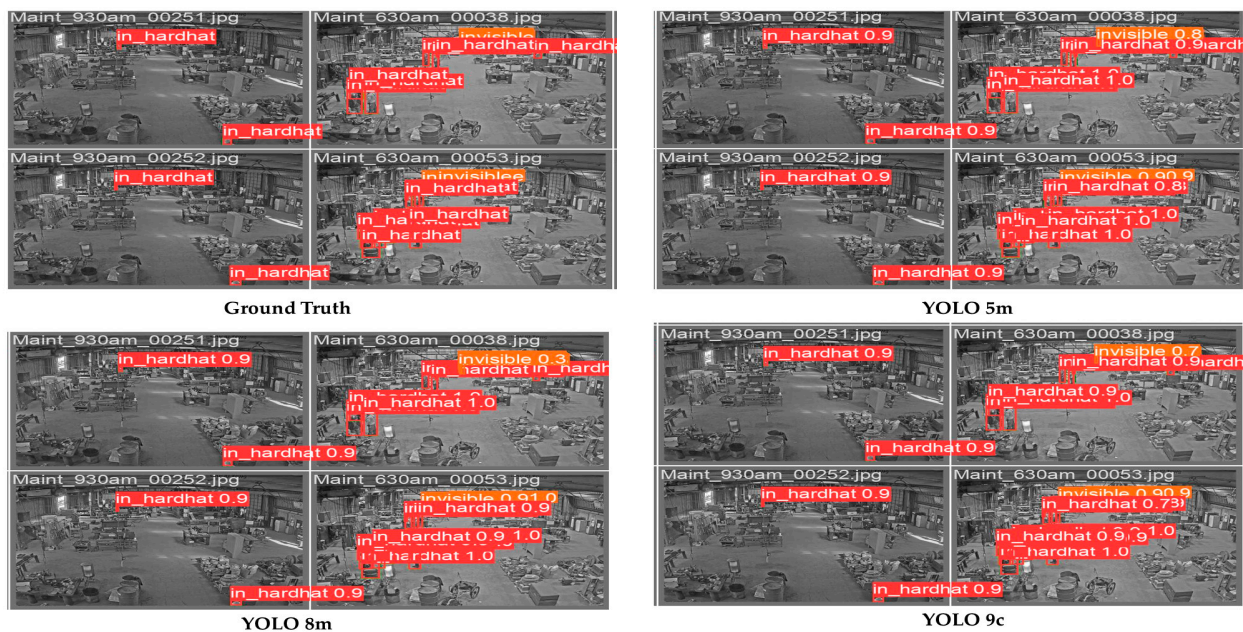


Figure 8. Ground truth compared with bounding box estimates of models.

Table 10 presents a comprehensive summary of the cross-validation results for each model across four key performance metrics: precision, recall, F1-score, and mean average precision (mAP). These metrics provide a detailed insight into the models’ effectiveness in detecting safety compliance and violations. High precision indicates the models’ ability to accurately identify non-compliant workers, minimizing false positives. High recall ensures

that most non-compliant instances are detected, minimizing false negatives. The F1-score provides a balance between precision and recall, reflecting the overall effectiveness of the models. Mean average precision (mAP) evaluates the models' performance across various recall levels, providing a comprehensive measure of accuracy table that illustrates the results of K-fold cross-validation for three different YOLO models, YOLOv5m, YOLOv8m, and YOLOv9c, highlighting their performance in terms of average precision, recall, F1-score, and mean average precision (mAP) across five folds. YOLOv5m demonstrated an average precision of 0.976, with individual fold values ranging from 0.96 to 0.98. Its recall was consistently high, averaging 0.974, indicating that the model effectively identifies true positive instances. However, its F1-score exhibited some variability, with a mean of 0.956, suggesting fluctuations in the balance between precision and recall. The mAP for YOLOv5m averaged 0.940, showing robust object detection capabilities across different classes despite some variability.

Table 10. K-fold validation results across models.

K-Fold	Average Precision			Recall			F1-Score			mAP		
	YOLOv5m	YOLOv8m	YOLOv9c	YOLOv5m	YOLOv8m	YOLOv9c	YOLOv5m	YOLOv8m	YOLOv9c	YOLOv5m	YOLOv8m	YOLOv9c
1	0.96	0.97	0.94	0.96	0.98	0.98	0.97	0.97	0.98	0.93	0.93	0.94
2	0.97	0.98	0.98	0.99	0.98	0.97	0.958	0.98	0.97	0.94	0.92	0.94
3	0.99	0.98	0.98	0.98	0.97	0.98	0.94	0.98	0.98	0.95	0.95	0.95
4	0.98	0.98	0.99	0.97	0.97	0.97	0.98	0.98	0.98	0.95	0.94	0.95
5	0.98	0.98	0.98	0.98	0.98	0.98	0.93	0.98	0.98	0.93	0.94	0.95
Avg	0.976	0.978	0.974	0.976	0.976	0.976	0.9556	0.982	0.978	0.938	0.936	0.944

YOLOv8m, on the other hand, exhibited superior consistency across all metrics. Its average precision was 0.978, with little variation across folds. The recall averaged 0.974, matching YOLOv5m, but with less variability, indicating more reliable performance. The F1-score for YOLOv8m was consistently high, with an average of 0.978, reflecting its strong ability to balance precision and recall. The mAP for YOLOv8m averaged 0.938, demonstrating its capability to accurately detect objects across different classes with slight variability but still maintaining high performance.

YOLOv9c displayed high precision and recall values similar to YOLOv8m, with averages of 0.974 and 0.976, respectively. Its F1-score was the highest among the three models, averaging 0.982, indicating the best balance between precision and recall. YOLOv9c's mAP was the highest, with an average of 0.944, showcasing its superior performance in detecting objects across all classes consistently.

Overall, the high performance across all models (with all metrics above 0.93) demonstrates the viability of using these YOLO variants for hard hat detection in steel manufacturing environments. The consistent recall scores obtained are particularly important for safety applications, as they indicate a low probability of missing instances where workers are not wearing hard hats. However, the slight variations in performance metrics highlight the importance of model selection based on specific use-case requirements. For instance, if minimizing false alarms is a priority, YOLOv8m might be preferred due to its high and consistent precision. If adaptability to various scenarios is crucial, YOLOv9c could be the better choice, given its superior mAP.

The box plots in Figure 9 provide a visual representation of the performance of three YOLO models (YOLOv5m, YOLOv8m, and YOLOv9c) across multiple metrics: precision, recall, F1-score, and mean average precision (mAP). The precision plot shows that YOLOv5m has a wider spread compared to YOLOv8m and YOLOv9c, indicating more variability in its precision across different folds. The median precision of YOLOv5m is slightly lower than that of YOLOv8m and YOLOv9c, which have very similar and consistent precision values, as evidenced by their narrow interquartile ranges and few outliers.

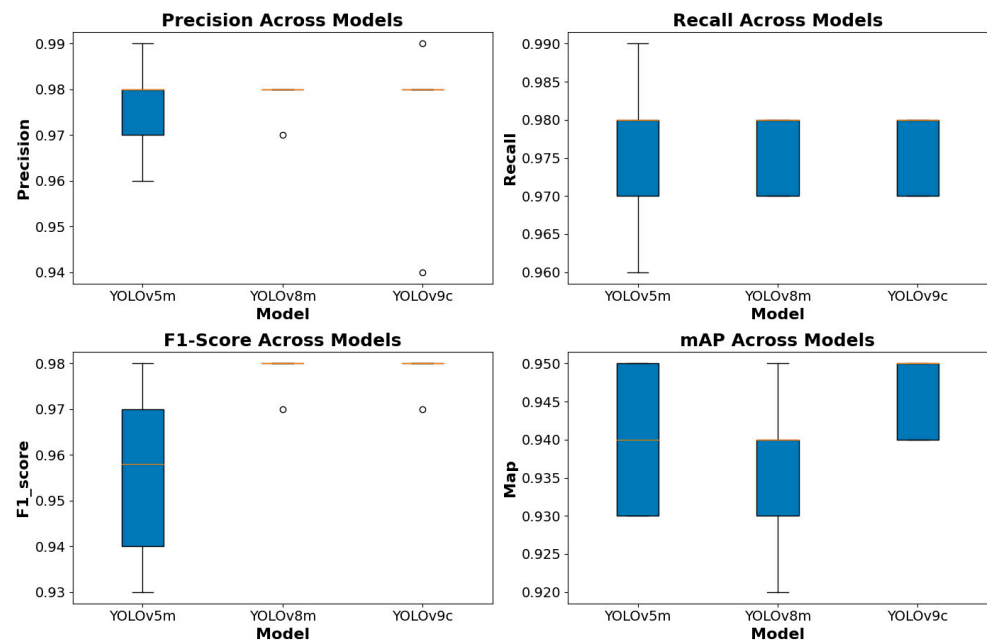


Figure 9. Box plots of performance metrics across models.

In the recall plot, YOLOv5m exhibits more variability compared to the other two models. The recall values for YOLOv8m and YOLOv9c are very consistent, with their medians and interquartile ranges nearly identical, highlighting their reliability in identifying true positive instances across different folds. The F1-score plot shows that YOLOv5m has a wider interquartile range and several outliers, indicating fluctuations in balancing precision and recall across folds. In contrast, YOLOv8m and YOLOv9c demonstrate very stable F1-scores, with narrow interquartile ranges and few outliers, reflecting their strong and consistent performance. The mAP plot reveals that YOLOv5m has a slightly wider spread compared to YOLOv8m and YOLOv9c, but its median mAP is comparable to the other two models. YOLOv8m shows a noticeable variability in mAP, suggesting some inconsistency in detecting objects across different classes. YOLOv9c, however, maintains a high and consistent mAP, as indicated by its narrow interquartile range and absence of outliers.

The high performance across all models (with all metrics consistently above 0.93) validates the feasibility of using these YOLO variants for hard hat detection in steel manufacturing environments. The consistently high recall scores are particularly crucial for safety applications, as they indicate a low probability of missing instances where workers are not wearing hard hats. However, the slight variations in performance metrics highlight the importance of model selection based on specific use-case requirements: For areas where minimizing false alarms is critical, YOLOv8m might be preferred due to its high and consistent precision and F1-score. For applications requiring adaptability to various scenarios or where overall detection performance is paramount, YOLOv9c could be the optimal choice given its superior mAP. In scenarios where computational resources are limited or where a balance between performance and model complexity is needed, YOLOv5m remains a viable option.

Specificity and AUC Results

The AUC and specificity analysis in Figure 10 revealed nuanced performance differences among YOLOv5m, YOLOv8m, and YOLOv9c for hard hat detection, with an additional challenge presented by the grayscale nature of the source videos. For the “in_hardhat” class, all models demonstrated high AUC values (>0.7), with YOLOv9c slightly outperforming the others ($AUC \approx 0.75$), indicating robust discrimination ability for compliant hard hat usage even in the absence of color information. However, the “no_hardhat” class presented significant challenges, with lower AUC scores across all models, particularly for

YOLOv5m (AUC ≈ 0.38), while YOLOv9c showed the best performance (AUC ≈ 0.6). This performance disparity likely stems from a class imbalance in the training data, limitations of synthetic data augmentation, and the reduced feature space inherent to grayscale imagery. The grayscale format potentially exacerbated difficulties in distinguishing subtle contrasts between hard hats and backgrounds or other headwear, especially in low-light areas of the steel mill.

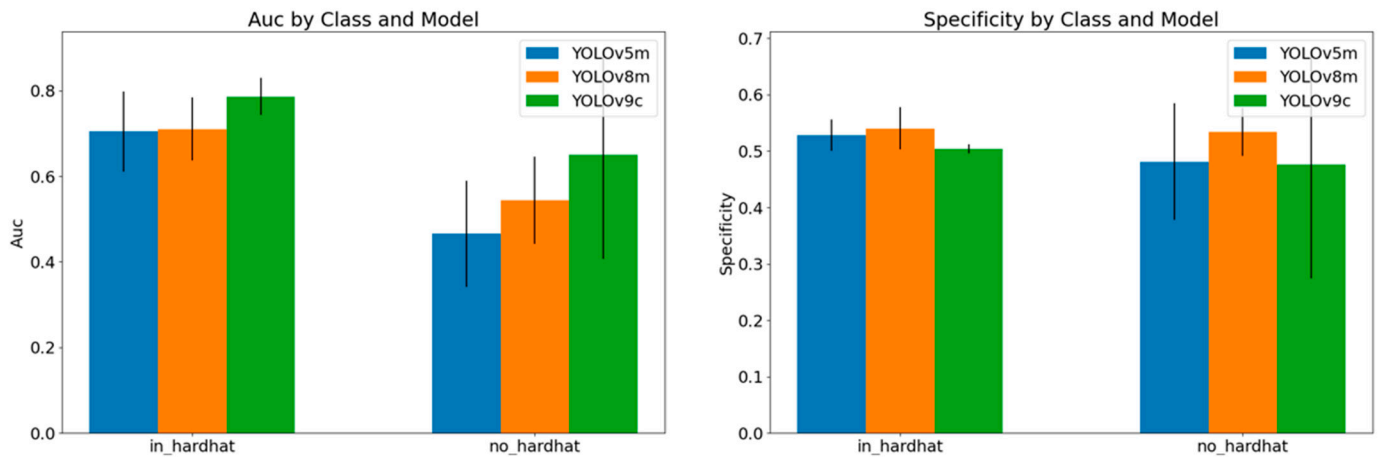


Figure 10. Bar plots with error bars of performance metrics across models.

Specificity analysis further elucidated these trends, with YOLOv8m exhibiting the highest specificity for the “in_hardhat” class (≈ 0.54), suggesting superior false positive mitigation even in grayscale conditions. For the “no_hardhat” class, YOLOv9c showed the highest specificity (≈ 0.55), albeit with greater variability across folds. These results underscore the need for targeted improvements in model architecture and training strategies, with particular attention to enhancing performance on grayscale inputs. Future work should focus on addressing class imbalance through more sophisticated data augmentation techniques, developing illumination-invariant features, and potentially exploring multi-modal approaches that can complement the limited information in grayscale imagery. While YOLOv9c demonstrates the most promise for real-world deployment in steel manufacturing safety monitoring, particularly for detecting safety violations, further refinement is necessary to improve the reliability of “no_hardhat” detection in challenging grayscale scenarios typical of steel manufacturing environments.

5. Contributions and Limitations of This Study

This study provides a detailed characterization of the specific hazards present in the steel manufacturing industry and explores how computer vision (CV) technologies can be applied to mitigate these risks. By identifying key risk factors and mapping them to CV applications, the research offers a strategic framework for enhancing workplace safety; this detailed characterization is currently lacking in existing research. This pilot study makes several significant contributions to the field of computer vision-based safety management in the steel industry while also acknowledging important limitations. Primarily, it pioneers the application of state-of-the-art YOLO models (YOLOv5m, YOLOv8m, and YOLOv9c) for safety management, specifically hard hat detection in steel manufacturing environments, addressing a critical gap in both the literature and industry practice. Also, this study provides a unique analysis of these models’ performance on grayscale imagery, which is common in industrial CCTV systems but underrepresented in computer vision research. The comprehensive comparison of these models, utilizing diverse metrics (precision, recall, F1-score, mAP, specificity, and AUC) across multiple classes (in_hardhat, no_hardhat, and invisible), provides valuable insights into their relative strengths and weaknesses in this specific context. Furthermore, this study identifies unique challenges in applying computer vision in steel manufacturing, such as extreme lighting conditions and dynamic

environments, establishing a performance baseline and methodological framework for future research.

However, several limitations must be considered when interpreting these results. This study utilized a relatively small dataset (703 images) from a specific area of a steel mini-mill, potentially impacting the generalizability of findings. Class imbalance, particularly the scarcity of “no_hardhat” instances, necessitated synthetic data augmentation, which may not fully capture real-world complexity. Additionally, this study’s focus solely on hard hat detection does not address the full spectrum of safety equipment required in steel manufacturing.

Despite these limitations, the high overall performance of YOLO models in hard hat detection, especially in the precision, recall, F1-score, and mAP, demonstrates the potential of computer vision for enhancing safety monitoring in steel manufacturing. These findings suggest that with further refinement, such models could significantly improve compliance with safety regulations and potentially reduce accidents in steel manufacturing environments. Future research should focus on expanding the dataset to include more diverse scenarios, developing techniques to improve model performance in detecting safety violations and handling occlusions, evaluating real-time performance, and extending the scope to include other critical safety equipment. Specifically, we recommend (a) collecting larger, more diverse datasets that better represent the full spectrum of steel manufacturing environments; (b) developing advanced data augmentation techniques to address class imbalance issues; (c) investigating multi-camera setups to mitigate occlusion problems; and (d) extending the study to include the detection of other safety equipment such as safety glasses, gloves, and protective clothing.

This pilot study provides crucial insights that will guide future research and development efforts in industrial safety, balancing promising results with a clear understanding of current limitations and areas for improvement. By addressing these challenges, future studies can further enhance the applicability and reliability of computer vision systems in the steel manufacturing environment.

6. Conclusions

The paramount importance of worker safety in steel manufacturing sites necessitates vigilant monitoring against unsafe practices and proactive identification of potential hazards. This research explored the feasibility and potential of implementing computer vision technologies for safety management within the steel manufacturing industry through a pilot case study focused on automatically detecting the use of hard hats by steelworkers using computer vision-based deep learning technology.

This study began with a comprehensive review phase characterizing hazards in the steel manufacturing environment and exploring their mitigation through computer vision applications. A multi-criteria decision model (TOPSIS MCDM) was then deployed to select commercially available computer vision programs suitable for safety management in this context. This approach involved a thorough assessment of hazards present in a steel mini-mill, accomplished through detailed analysis of mill reports, on-site observations, and stakeholder consultations. This comprehensive analysis provided the necessary insights for aligning potential hazards with computer vision capabilities, demonstrating the practicality of using computer vision for automated hazard recognition and active worksite surveillance. An extensive online search led to the evaluation of eight commercially available computer vision systems, with the Everguard system emerging as the most suitable candidate for this pilot study.

This pilot study demonstrated the feasibility of implementing computer vision systems for hard hat detection in steel manufacturing environments. We evaluated three state-of-the-art YOLO models (YOLOv5m, YOLOv8m, and YOLOv9c) using a dataset of 703 grayscale images from a steel mini-mill. A comprehensive comparison of these models across diverse metrics (precision, recall, F1-score, mAP, specificity, and AUC) provided valuable insights into their relative strengths and weaknesses in this specific context. All models showed

promising performance, particularly for the “in_hardhat” class, with high AUC values (>0.7) and YOLOv9c slightly outperforming the others. However, the “no_hardhat” class presented significant challenges, with lower AUC scores across all models, particularly for YOLOv5m. YOLOv9c demonstrated the best performance in detecting safety violations. The grayscale nature of the source videos added complexity to the detection task, potentially exacerbating difficulties in distinguishing subtle contrasts in low-light areas of the steel mill.

While this study affirms the feasibility of applying computer vision-based deep learning for safety management in steel manufacturing environments, it also highlights important challenges. These include class imbalance issues, the limitations of synthetic data augmentation, and the reduced feature space inherent to grayscale imagery. Future research would focus on addressing these challenges through more sophisticated data augmentation techniques, developing illumination-invariant features, and exploring multi-modal approaches to complement the limited information in grayscale imagery.

Looking ahead, this methodology holds promise for detecting additional safety equipment such as vests, gloves, and glasses, as well as for monitoring workers’ posture and proximity to heavy machinery. These areas represent valuable directions for future research, contributing to the continuous enhancement of workplace safety in the steel manufacturing industry.

In conclusion, this pilot study provides crucial insights that will guide future research and development efforts in industrial safety, balancing promising results with a clear understanding of current limitations and areas for improvement. By addressing these challenges, future studies can further enhance the applicability and reliability of computer vision systems in industrial safety management, particularly in the complex and dynamic environment of steel manufacturing.

Author Contributions: Conceptualization, I.A. and J.C.; methodology, R.L. and I.A.; formal analysis, R.L.; writing—original draft preparation, R.L.; writing—review and editing, I.A. and J.C.; supervision, I.A.; project administration, I.A.; funding acquisition, I.A. and J.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Association for Iron & Steel Technology (AIST) Foundation through the Digital Technologies for Steel Manufacturing (DTSM) Grant.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available upon reasonable request from the corresponding author.

Acknowledgments: The authors are thankful to the Association for Iron & Steel Technology (AIST) Foundation for funding this research through the Digital Technologies for Steel Manufacturing (DTSM) Grant. In addition, we would like to express our appreciation to our industry partners, particularly Paul Thurber, Yong Wu, and the Everguard Team, for their support of the pilot case study conducted in this project.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Khahro, S.H.; Khahro, Q.H.; Ali, T.H.; Memon, Z.A. Industrial Accidents and Key Causes: A Case Study of the Steel Industry. *ACM Int. Conf. Proceeding Ser.* **2023**, *66*–70. [[CrossRef](#)]
2. Sacks, R.; Perlman, A.; Barak, R. Construction safety training using immersive virtual reality. *Constr. Manag. Econ.* **2013**, *31*, 1005–1017. [[CrossRef](#)]
3. Zhang, S.; Sulankivi, K.; Kiviniemi, M.; Romo, I.; Eastman, C.M.; Teizer, J. BIM-based fall hazard identification and prevention in construction safety planning. *Saf. Sci.* **2015**, *72*, 31–45. [[CrossRef](#)]
4. Nordlöf, H.; Wiitavaara, B.; Winblad, U.; Wijk, K.; Westerling, R. Safety culture and reasons for risk-taking at a large steel-manufacturing company: Investigating the worker perspective. *Saf. Sci.* **2015**, *73*, 126–135. [[CrossRef](#)]
5. Zhou, D.; Xu, K.; Lv, Z.; Yang, J.; Li, M.; He, F.; Xu, G. Intelligent Manufacturing Technology in the Steel Industry of China: A Review. *Sensors* **2022**, *22*, 8194. [[CrossRef](#)] [[PubMed](#)]
6. Chai, J.; Zeng, H.; Li, A.; Ngai, E.W.T. Deep learning in computer vision: A critical review of emerging techniques and application scenarios. *Mach. Learn. Appl.* **2021**, *6*, 100134. [[CrossRef](#)]

7. Feng, X.; Jiang, Y.; Yang, X.; Du, M.; Li, X. Computer vision algorithms and hardware implementations: A survey. *Integration* **2019**, *69*, 309–320. [CrossRef]
8. Lan, R.; Awolusi, I.; Cai, J. Digital computer vision for safety management in steel manufacturing. In Proceedings of the Iron & Steel Technology Conference, Detroit, MI, USA, 8–11 May 2023; pp. 31–42.
9. Marks, E.D.; Teizer, J. Method for testing proximity detection and alert technology for safe construction equipment operation. *Constr. Manag. Econ.* **2013**, *31*, 636–646. [CrossRef]
10. Awolusi, I.; Song, S.; Marks, E. Forklift safety: Sensing the dangers with technology. *Prof. Saf.* **2017**, *62*, 36–39.
11. Kursunoglu, N.; Onder, S.; Onder, M. The Evaluation of Personal Protective Equipment Usage Habit of Mining Employees Using Structural Equation Modeling. *Saf. Health Work.* **2022**, *13*, 180–186. [CrossRef]
12. Zhang, M.; Cao, Z.; Yang, Z.; Zhao, X. Utilizing Computer Vision and Fuzzy Inference to Evaluate Level of Collision Safety for Workers and Equipment in a Dynamic Environment. *J. Constr. Eng. Manag.* **2020**, *146*, 04020051. [CrossRef]
13. Ghosh, A.; Chatterjee, A. *Ironmaking and Steelmaking: Theory and Practice*; PHI Learn. Priv. Ltd.: Delhi, India, 2008.
14. Xu, Z.J.; Zheng, Z.; Gao, X.Q. Operation optimization of the steel manufacturing process: A brief review. *Int. J. Miner. Metall. Mater.* **2021**, *28*, 1274–1287. [CrossRef]
15. Bae, J.; Li, Y.; Ståhl, N.; Mathiason, G.; Kojola, N. Using Machine Learning for Robust Target Prediction in a Basic Oxygen Furnace System. *Metall. Mater. Trans. B* **2020**, *51*, 1632–1645. [CrossRef]
16. Nutting, J.; Edward, F.; Wondris, E. Steel. Encyclopedia Britannica. Available online: <https://www.britannica.com/technology/steel> (accessed on 4 October 2022).
17. Burchart-Korol, D. Life cycle assessment of steel production in Poland: A case study. *J. Clean. Prod.* **2013**, *54*, 235–243. [CrossRef]
18. Nair, A.T.; Mathew, A.; Archana, A.R.; Akbar, M.A. Use of hazardous electric arc furnace dust in the construction industry: A cleaner production approach. *J. Clean. Prod.* **2022**, *377*, 134282. [CrossRef]
19. World Steel Association. Safety and Health in the Steel Industry: Data Report 2023. 2023. Available online: <https://worldsteel.org/steel-topics/safety-and-health/safety-and-health-in-the-steel-industry-data-report-2023/> (accessed on 18 November 2023).
20. Ali, M.X.M.; Arifin, K.; Abas, A.; Ahmad, M.A.; Khairil, M.; Cyio, M.B.; Samad, M.A.; Lampe, I.; Mahfudz, M.; Ali, M.N. Systematic Literature Review on Indicators Use in Safety Management Practices among Utility Industries. *Int. J. Environ. Res. Public Health* **2022**, *19*, 6198. [CrossRef] [PubMed]
21. Tang, B.; Chen, L.; Sun, W.; Lin, Z.K. Review of surface defect detection of steel products based on machine vision. *IET Image Process* **2022**, *17*, 303–322. [CrossRef]
22. International Labour Organization. Sectoral Activities Programme. In Proceedings of the Code of Practice on Safety and Health in the Iron and Steel Industry: Meeting of Experts to Develop a Revised Code of Practice on Safety and Health in the Iron and Steel Industry, Geneva, Switzerland, 1–9 February 2005.
23. Kifle, M.; Engdaw, D.; Alemu, K.; Sharma, H.R.; Amsalu, S.; Feleke, A.; Worku, W. Work related injuries and associated risk factors among iron and steel industries workers in Addis Ababa, Ethiopia. *Saf. Sci.* **2014**, *63*, 211–216. [CrossRef]
24. National Institute for Occupational Safety and Health. Hierarchy of Controls. CDC. 2015. Available online: <https://www.cdc.gov/niosh/hierarchy-of-controls/about/index.html> (accessed on 20 November 2023).
25. Houette, B.; Mueller-Hirth, N. Practices, preferences, and understandings of rewarding to improve safety in high-risk industries. *J. Saf. Res.* **2022**, *80*, 302–310. [CrossRef]
26. Berhan, E. Prevalence of occupational accident; and injuries and their associated factors in iron, steel and metal manufacturing industries in Addis Ababa. *Cogent Eng.* **2020**, *7*, 1723211. [CrossRef]
27. Awolusi, I.; Marks, E.; Hollowell, M. Wearable technology for personalized construction safety monitoring and trending: Review of applicable devices. *Autom. Constr.* **2018**, *85*, 96–106. [CrossRef]
28. Márquez-Sánchez, S.; Campero-Jurado, I.; Herrera-Santos, J.; Rodríguez, S.; Corchado, J.M. Intelligent platform based on smart ppe for safety in workplaces. *Sensors* **2021**, *21*, 4652. [CrossRef]
29. Nnaji, C.; Awolusi, I.; Park, J.W.; Albert, A. Wearable sensing devices: Towards the development of a personalized system for construction safety and health risk mitigation. *Sensors* **2021**, *21*, 682. [CrossRef]
30. Hong, X.; Lv, B. Application of Training Simulation Software and Virtual Reality Technology in Civil Engineering. In Proceedings of the 2022 IEEE International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA), Changchun, China, 25–27 February 2022; pp. 520–524.
31. Velez, D.; Zlateva, P. Virtual Reality Challenges in Education and Training. *Int. J. Learn.* **2017**, *3*, 33–37. [CrossRef]
32. Chai, X.; Lee, B.G.; Pike, M.; Wu, R.; Chieng, D.; Chung, W.Y. Pre-impact Firefighter Fall Detection Using Machine Learning on the Edge. *IEEE Sens. J.* **2023**, *23*, 14997–15009. [CrossRef]
33. Zhou, L.; Zhang, L.; Konz, N. Computer Vision Techniques in Manufacturing. *IEEE Trans. Syst. Man. Cybern. Syst.* **2022**, *53*, 105–117. [CrossRef]
34. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. *arXiv* **2015**, arXiv:1506.02640.
35. Redmon, J.; Farhadi, A. YOLO9000: Better, faster, stronger. In Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, 21–26 November 2017; pp. 6517–6525. [CrossRef]
36. Bochkovskiy, A.; Wang, C.-Y.; Liao, H.-Y.M. YOLOv4: Optimal Speed and Accuracy of Object Detection. *arXiv* **2020**, arXiv:2004.10934.

37. Terven, J.; Córdova-Esparza, D.M.; Romero-González, J.A. A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS. *Mach. Learn. Knowl. Extr.* **2023**, *5*, 1680–1716. [[CrossRef](#)]
38. Jiang, P.; Ergu, D.; Liu, F.; Cai, Y.; Ma, B. A Review of Yolo Algorithm Developments. *Procedia Comput. Sci.* **2021**, *199*, 1066–1073. [[CrossRef](#)]
39. Li, Y.; Wang, H.; Dang, L.M.; Nguyen, T.N.; Han, D.; Lee, A.; Jang, I.; Moon, H. A deep learning-based hybrid framework for object detection and recognition in autonomous driving. *IEEE Access* **2020**, *8*, 194228–194239. [[CrossRef](#)]
40. Narejo, S.; Pandey, B.; Vargas, D.E.; Rodriguez, C.; Anjum, M.R. Weapon Detection Using YOLO V3 for Smart Surveillance System. *Math. Probl. Eng.* **2021**, *2021*, 9975700. [[CrossRef](#)]
41. Ragab, M.G.; Abdulkader, S.J.; Muneer, A.; Alqushaibi, A.; Sumiea, E.H.; Qureshi, R.; Al-Selwi, S.M.; Alhussian, H. A Comprehensive Systematic Review of YOLO for Medical Object Detection (2018 to 2023). *IEEE Access* **2024**, *12*, 57815–57836. [[CrossRef](#)]
42. Lippi, M.; Bonucci, N.; Carpio, R.F.; Contarini, M.; Speranza, S.; Gasparri, A. A YOLO-based pest detection system for precision agriculture. In Proceedings of the 2021 29th Mediterranean Conference on Control and Automation, MED 2021, Puglia, Italy, 22–25 June 2021; pp. 342–347. [[CrossRef](#)]
43. Kim, K.; Kim, K.; Jeong, S. Application of YOLO v5 and v8 for Recognition of Safety Risk Factors at Construction Sites. *Sustainability* **2023**, *15*, 15179. [[CrossRef](#)]
44. Qiu, Q.; Lau, D. Real-time detection of cracks in tiled sidewalks using YOLO-based method applied to unmanned aerial vehicle (UAV) images. *Autom. Constr.* **2023**, *147*, 104745. [[CrossRef](#)]
45. Gai, R.; Chen, N.; Yuan, H. A detection algorithm for cherry fruits based on the improved YOLO-v4 model. *Neural Comput. Appl.* **2023**, *35*, 13895–13906. [[CrossRef](#)]
46. Jung, H.; Rhee, J. Application of YOLO and ResNet in Heat Staking Process Inspection. *Sustainability* **2022**, *14*, 15892. [[CrossRef](#)]
47. Mushtaq, F.; Ramesh, K.; Deshmukh, S.; Ray, T.; Parimi, C.; Tandon, P.; Jha, P.K. YOLO-v5 and image processing based component identification system. *Eng. Appl. Artif. Intell.* **2023**, *118*, 105665. [[CrossRef](#)]
48. Kou, X.; Liu, S.; Cheng, K.; Qian, Y. Development of a YOLO-V3-based model for detecting defects on steel strip surface. *Measurement* **2021**, *182*, 109454. [[CrossRef](#)]
49. Pitts, H. Warehouse Robot Detection for Human Safety Using YOLOv8. In Proceedings of the SoutheastCon 2024, Atlanta, GA, USA, 15–24 March 2024; pp. 1184–1188. [[CrossRef](#)]
50. Hao, Z.; Wang, Z.; Bai, D.; Tao, B.; Tong, X.; Chen, B. Intelligent Detection of Steel Defects Based on Improved Split Attention Networks. *Front. Bioeng. Biotechnol.* **2022**, *9*, 810876.
51. Martins, L.A.O.; Pádua, F.L.C.; Almeida, P.E.M. Automatic detection of surface defects on rolled steel using Computer Vision and Artificial Neural Networks. In Proceedings of the IECON 2010-36th Annual Conference on IEEE Industrial Electronics Society, Glendale, AZ, USA, 7–10 November 2010; pp. 1081–1086.
52. Szyzakin, R.; Voronin, V.; Gapon, N.; Zelensky, A.; Pižurica, A. Automatic detection of welding defects using the convolutional neural network. In *Automated Visual Inspection and Machine Vision III*; SPIE: Philadelphia, PA, USA, 2019.
53. Bolderston, A. Conducting a research interview. *J. Med. Imaging Radiat. Sci.* **2012**, *43*, 66–76. [[CrossRef](#)]
54. Guest, G.; Namey, E.; Taylor, J.; Eley, N.; McKenna, K. Comparing focus groups and individual interviews: Findings from a randomized study. *Int. J. Soc. Res. Methodol.* **2017**, *20*, 693–708. [[CrossRef](#)]
55. Creswell, J.W. Editorial: Mapping the field of mixed methods research. *J. Mix. Methods Res.* **2009**, *3*, 95–108. [[CrossRef](#)]
56. Çelikbilek, Y.; Tüysüz, F. An in-depth review of theory of the TOPSIS method: An experimental analysis. *J. Manag. Anal.* **2020**, *7*, 281–300. [[CrossRef](#)]
57. Kraujalienė, L. Comparative Analysis of Multicriteria Decision-Making Methods Evaluating the Efficiency of Technology Transfer. *Bus. Manag. Educ.* **2019**, *17*, 72–93. [[CrossRef](#)]
58. Olson, D.L. Comparison of weights in TOPSIS models. *Math. Comput. Model.* **2004**, *40*, 721–727. [[CrossRef](#)]
59. Chakraborty, S. TOPSIS and Modified TOPSIS: A comparative analysis. *Decis. Anal. J.* **2022**, *2*, 100021. [[CrossRef](#)]
60. Yahya, M.N.; Gökçekuş, H.; Ozsahin, D.U.; Uzun, B. Evaluation of wastewater treatment technologies using topsis. *Desalin. Water Treat.* **2020**, *177*, 416–422. [[CrossRef](#)]
61. Rombach, R.; Blattmann, A.; Lorenz, D.; Esser, P.; Ommer, B. High-Resolution Image Synthesis with Latent Diffusion Models. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022.
62. Tharwat, A. Classification assessment methods. *Appl. Comput. Inform.* **2018**, *17*, 168–192. [[CrossRef](#)]
63. Obi, J.C. A comparative study of several classification metrics and their performances on data. *World J. Adv. Eng. Technol. Sci.* **2023**, *8*, 308–314. [[CrossRef](#)]
64. Yan, X.; Zhang, H.; Li, H. Computer vision-based recognition of 3D relationship between construction entities for monitoring struck-by accidents. *Comput. Aided Civ. Infrastruct. Eng.* **2020**, *35*, 1023–1038. [[CrossRef](#)]
65. Anwar, Q.; Hanif, M.; Shimotoku, D.; Kobayashi, H.H. Driver awareness collision/proximity detection system for heavy vehicles based on deep neural network. *J. Phys. Conf. Ser.* **2022**, *2330*, 012001. [[CrossRef](#)]
66. Wu, H.; Wu, D.; Zhao, J. An intelligent fire detection approach through cameras based on computer vision methods. *Process Saf. Environ. Prot.* **2019**, *127*, 245–256. [[CrossRef](#)]
67. Krestenitis, M.; Orfanidis, G.; Ioannidis, K.; Avgerinakis, K.; Vrochidis, S.; Kompatsiaris, I. Oil spill identification from satellite images using deep neural networks. *Remote Sens.* **2019**, *11*, 1762. [[CrossRef](#)]

68. Lee, H.; Lee, G.; Lee, S.H.; Ahn, C.R. Assessing exposure to slip, trip, and fall hazards based on abnormal gait patterns predicted from confidence interval estimation. *Autom. Constr.* **2022**, *139*, 104253. [[CrossRef](#)]
69. Luo, H.; Wang, M.; Wong, P.K.Y.; Cheng, J.C.P. Full body pose estimation of construction equipment using computer vision and deep learning techniques. *Autom. Constr.* **2020**, *110*, 103016. [[CrossRef](#)]
70. Balakrishnan, B.; Richards, G.; Nanda, G.; Mao, H.; Athinarayanan, R.; Zaccaria, J. PPE compliance detection using artificial intelligence in learning factories. *Procedia Manuf.* **2020**, *45*, 277–282. [[CrossRef](#)]
71. Abd El-Rahiem, B.; Sedik, A.; El Banby, G.M.; Ibrahim, H.M.; Amin, M.; Song, O.Y.; Khalaf, A.A.M.; Abd El-Samie, F.E. An efficient deep learning model for classification of thermal face images. *J. Enterp. Inf. Manag.* **2020**, *36*, 706–717. [[CrossRef](#)]
72. Gupta, A.; Ramanath, R.; Shi, J.; Keerthi, S.S. Adam vs. SGD: Closing the generalization gap on image classification. In Proceedings of the OPT2021: 13th Annual Workshop on Optimization for Machine Learning, Virtual, 22 October 2021.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.