

Bashayer Shaheen

PROJECT2-4.docx

 Midocean University

Document Details

Submission ID

trn:oid::2945:380838489

Submission Date

Apr 28, 2026, 11:15 PM GMT+3

Download Date

May 5, 2026, 11:40 AM GMT+3

File Name

PROJECT2-4.docx

File Size

2.5 MB

33 Pages

6,414 Words

40,839 Characters

63% detected as AI

The percentage indicates the combined amount of likely AI-generated text as well as likely AI-generated text that was also likely AI-paraphrased.

Caution: Review required.

It is essential to understand the limitations of AI detection before making decisions about a student's work. We encourage you to learn more about Turnitin's AI detection capabilities before using the tool.

Detection Groups

-  **30 AI-generated only 63%**
Likely AI-generated text from a large-language model.
-  **0 AI-generated text that was AI-paraphrased 0%**
Likely AI-generated text that was likely revised using an AI-paraphrase tool or word spinner.

Disclaimer

Our AI writing assessment is designed to help educators identify text that might be prepared by a generative AI tool. Our AI writing assessment may not always be accurate (i.e., our AI models may produce either false positive results or false negative results), so it should not be used as the sole basis for adverse actions against a student. It takes further scrutiny and human judgment in conjunction with an organization's application of its specific academic policies to determine whether any academic misconduct has occurred.

Frequently Asked Questions

How should I interpret Turnitin's AI writing percentage and false positives?

The percentage shown in the AI writing report is the amount of qualifying text within the submission that Turnitin's AI writing detection model determines was either likely AI-generated text from a large-language model or likely AI-generated text that was likely revised using an AI paraphrase tool or word spinner.

False positives (incorrectly flagging human-written text as AI-generated) are a possibility in AI models.

AI detection scores under 20%, which we do not surface in new reports, have a higher likelihood of false positives. To reduce the likelihood of misinterpretation, no score or highlights are attributed and are indicated with an asterisk in the report (*%).

The AI writing percentage should not be the sole basis to determine whether misconduct has occurred. The reviewer/instructor should use the percentage as a means to start a formative conversation with their student and/or use it to examine the submitted assignment in accordance with their school's policies.

What does 'qualifying text' mean?

Our model only processes qualifying text in the form of long-form writing. Long-form writing means individual sentences contained in paragraphs that make up a longer piece of written work, such as an essay, a dissertation, or an article, etc. Qualifying text that has been determined to be likely AI-generated will be highlighted in cyan in the submission, and likely AI-generated and then likely AI-paraphrased will be highlighted purple.

Non-qualifying text, such as bullet points, annotated bibliographies, etc., will not be processed and can create disparity between the submission highlights and the percentage shown.





Project Proposal

Intelligent detection of safety equipment violations using artificial intelligence

Prepared By

Name: Kholod Alharbi

ID No.: 252001457

Name: Bashayer Shaheen

ID No.: 252001365

Submitted in fulfillment of the requirements for the degree Master
of Science in Artificial Intelligence

At

Collage of Informatics

Midocean University

UAE

Supervisor Name: **Dr. khaled_Eskaf**

Collage of Informatics, Midocean University, Comoros

Email:.....

Acknowledgements

I would like to sincerely thank my supervisor, Dr. Khaled Eskaf, for his continuous support, valuable guidance, and constructive feedback throughout this project. His insights played an important role in shaping the direction of this work. I also extend my appreciation to the committee members for their time and evaluation. Special thanks go to the staff of the Artificial Intelligence Department, as well as to everyone who supported us during our academic journey.

Abstract

Personal protective equipment (PPE) had a productive role in providing safety to the workers and labors by deducting the chances and severity of injuries related to the accident that took place along construction environments and other unsafe workplaces. The use of these equipment's such as helmets and vests contributes up to a large sum to lessen site risks and maintaining safer and a better working environment. Even after the availability of PPE at most Development Site, workforce may not constantly follow with safety rules due to inattention, environmental conditions, or improper supervision.

Traditional ways that dependent on the old way manual inspections are often more time taking and always has a space for human error, especially in big construction environments and projects where many different workers are working at the same time. So, there is the dire need for advance monitoring systems that can automatically find out PPE compliance in real time has become a necessity. Such systems can help safety supervisors by providing continuous monitoring and notify on time of the unsafe situations.

In this part, use of advance techniques based on through learning were investigate to develop a solution to detect PPE in construction environments. The Construction Site Safety Dataset collected on daily basis from the available sources on Kaggle and Roboflow was use to train and judge the proposed detection model. The dataset has images that show workers and the daily used protective equipment captured under various environmental conditions and working scenarios.

To desired performance, the YOLOv8 object detection was adopted and tuned by using transfer learning techniques. The model was made to find out multiple PPE-related classes, including safety helmets, safety vests, and workers working in construction scenes. Experimental results shows that the given model was able to achieve quite remarkable detection accuracy and quick inference performance, making it the most desireable and suitable for real-time safety monitoring applications.

This approach gives us a practical contribution toward improving occupational safety by the use of automated monitoring of PPE compliance and to help early identification of unsafe behaviors that can result to workplace accidents and injuries.

Keywords: Safety Helmet Detection, PPE, YOLOv8

Table of Contents

List of Abbreviations	viii
List of Figures.....	viii
List of Tables	Error! Bookmark not defined.

Chapter 1: Introduction.....9

1.1 Problem Definition	10
1.2 Thesis Objectives	10
1.3 Thesis Contribution	11
1.4 Thesis Organization.....	11

Chapter 2: Background and Related Work.....12

2.1 Introduction.....	12
2.2 Background.....	12
2.3 Related Work.....	13

Chapter 3: Proposed Approach.....18

3.1 Introduction.....	18
3.2 Dataset Description.....	18
3.3 Dataset Preprocessing.....	19
3.4 Proposed Detection Model Architecture.....	20
3.5 Transfer Learning Strategy.....	21
3.6 Model Training Configuration.....	21
3.8 Summary.....	22

Chapter 4: Experimental Results and Discussion.....24

4.1 Introduction.....	24
4.2 Detection Performance Results.....	24
4.3 Training Behavior and Model Convergence.....	26
4.4 Qualitative Analysis of Detection Results.....	27

4.5 Discussion of Findings.....	28
4.6 Limitations of the Proposed Work.....	29
4.7 Summary of the Chapter.....	29
Chapter 5: Conclusion and Future Work	30
7.1 Important Features	30
7.3 Future Work	30
References	31

List of Figures

3.2 Example annotated image from the Construction Site Safety Dataset	19
3.6 Training and validation performance curves	22
4.2 Precision–Recall curve of the trained YOLOv8 model	24

4.3 Training and validation loss curves	27
4.4: Sample detection results produced by the proposed model	28

Chapter 1 :Introduction

The diverse nature of these workplaces and the presence of many different risk factors are still a big challenge in fulfil the safety of workers in working environments. The most important and the most critical safety measures used to lessen the risk of injuries and prevent severe accidents is personal protective equipment (PPE). However, laborers would usually fail to follow with the PPE requirements, either intentionally or by mistake, despite the presence of safety regulations and protective equipment.

Supervision is time and again labor-intensive, inefficient, and prone to human error in conventional safety monitoring methods. Continuous monitoring becomes very difficult to maintain through normal methods as construction sites grow in size and the nature of the work is complex. So, the development of such technology-based automated monitoring systems has come up as a research area of dire importance for increasing workplace safety.

Technology has progressed to find visual information and to identify objects as a result of progress in Artificial Intelligence, particularly in the fields of Deep Learning and Computer Vision. YOLO (You Only Look Once) is a family of advance object detection algorithms that has gained alot attention for its capacity to do fast and accurate real-time detections. As Compared to older versions, the most updated version, YOLOv8, gives enhanced performance and adaptability, making it the best for industrial safety monitoring applications. The primary objective of this thesis is to create an adaptable PPE detection system that uses image datasets to automatically find safety conformance on a construction site environment by the use of the YOLOv8 architecture.

1.1 Problem Definition

Construction and industrial workplaces are among one of the most dangerous working environments due to the availability of heavy equipment, heightened structures, and complex operational activities. Even though Personal Protective Equipment is known as an essential safety requirement, making sure continuous compliance remains a difficult yet achievable task when keeping an eye and it depends in the last on human supervision.

Older inspection methods have many limitations, including irregularity in observation, limited coverage of large work areas, and late response to safety violations. This increases the chances of accidents that result due to the unavailability of important protective equipment such as helmets and safety vests.

With the passage of time and advancement of computer vision technologies, there is a dire need to invent automated monitoring systems that have the ability to detect PPE usage in real time and without any error. Therefore, the main problem mentioned in this thesis is the design and implementation process of an object detection system that has the ability of automatically identifying PPE compliance in construction site environments by the help of deep learning techniques.

1.2 Thesis Objectives

The aim of this thesis is to provide a automated PPE detection system using a learning-based object detection to enhance workplace safety monitoring efficiency.

The specific objectives of this study include:

- find the importance of automated PPE monitoring systems in construction site environments
- look after the capabilities of learning techniques for object detection tasks
- use a PPE detection model using the YOLOv8 architecture
- use YOLOv8 model using a construction safety dataset
- evaluate the performance of the trained model using standard evaluation metrics such as Precision, Recall, and mean Average Precision (mAP)
- find out the effectiveness of the proposed model for real-use for workplace safety monitoring applications

1.3 Thesis Contribution

This thesis presents the use of a smart PPE discovery system by using one of the stylish object discovery algorithms grounded on the YOLOv8 armature. This work helps in amending occupational safety monitoring by the help of an automated approach which can detect safety compliance in construction surroundings.

The main contributions of this thesis are as following:

- Development of a deep learning-based PPE detection framework using YOLOv8
- Application of transfer learning techniques to improve detection performance
- Utilization of a construction site safety dataset containing annotated PPE classes
- Performance evaluation of the trained detection model using standard object detection metrics
- Demonstration of the feasibility of deploying real-time PPE monitoring systems in practical workplace environments

1.4 Thesis Organization

This thesis is organized into six chapters to present the research work.

Chapter 1 tells the research background, tells us about problem definition, its objectives, contributions, and overall structure.

Chapter 2 provides a short review of related work in the field of PPE detection and object detection using advance techniques.

Chapter 3 tells us the dataset which is used in the study and elaborates the steps involved before training the model.

Chapter 4 gives information about the methodology, including the architecture of the YOLOv8 model and the training configuration used for tuning.

Chapter 5 discusses the experimental results that we get from the trained model and judges its performance using standard metrics.

Chapter 6 finishes the thesis by the help of summarizing the findings of the study and provide possible options for future research.

Chapter 2 : Background and Related Work

2.1 Introduction

Making sure worker safety in such environments is a concern because of the increasing complexity of worksites and the potential high risk that comes with the dangerous operations. PPE plays a important role in lessening workplace injuries and minimizing serious accidents. To make sure safety rules are being followed is still a tough task when monitoring is dependent on human supervision.

Recent advancement in AI, especially in the field of Computer Vision and Deep Learning, had enabled the development of automated monitoring systems which can detect safety violations in real time. Object detection algorithms have resulted in performance in locating and identifying workers and verifying if they follow required protective equipment such as helmets and gear.

This chapter tells us about the background which is essential to understand the technologies that are used in this study. It also gets help from recent research efforts which are related to PPE detection with the help of deep learning–based object detection approach, by the help of a focus on the YOLO family of algorithms.

2.2 Background

Computer Vision is an essential part of AI that helps machines to understand visual information obtained from the images and videos. It plays an essential role in real time applications such as the surveillance systems, medical imaging, autonomous driving tech, and industrial safety monitoring.

The important task in Computer Vision is to find object, which involves identifying it in an image and finding its locations using bounding boxes. Old object detection techniques depend on handcrafted feature extraction methods, which are limited in accuracy and scalability. Whereas, the development of Deep Learning techniques greatly improves object detection performance by the use of automatic feature extraction using Convolutional Neural Networks (CNNs).

Deep Learning is machine learning that has multi-layer neural networks that are responsible to model complex patterns in data. CNNs, in particular, have become the dominant architecture to image recognition tasks because of their unique characteristic to learn spatial hierarchies of features directly from given images [3].

Object detection approaches are commonly grouped into two main categories: two-stage and one-stage detectors. In two-stage methods, such as the R-CNN family, the process begins by generating candidate regions, which are then classified into object categories. Although these methods are known for their high accuracy, they often require more processing time. In contrast, one-stage detectors perform localization and classification simultaneously within a single pass, making them more suitable for real-time applications where speed is critical.

2.3 Related Work

Recent research on personal protective equipment (PPE) detection has shown a clear shift toward deep learning and computer vision methods, particularly in construction and industrial safety applications. A large portion of this work has focused on object detection models because of their ability to identify workers and safety equipment directly from images or video frames with relatively high efficiency.

Chen and Demachi proposed a vision-based system for monitoring PPE usage in a nuclear power plant environment. Their study was based on 3,808 annotated images collected from both real-world scenes and online sources. The dataset was divided into four categories representing safe and unsafe conditions. To improve generalization, YOLOv3 was trained after applying augmentation techniques such as rotation and image flipping. The data were split into training, validation, and testing subsets, and the model achieved strong results in terms of mAP, recall, and F1-score [7]. In an additional testing stage, another manually collected dataset was used to evaluate the model further. The detection task was limited to two main objects: a full-face mask and a hard hat. The authors also noted that worker posture and distance had an effect on performance. Training was performed in two phases, where the final convolutional layer in Darknet-53 was first frozen and then all layers were unfrozen for fine-tuning. The authors suggested that future work should focus on improving dataset quality and integrating the model with an on-site surveillance system.

Similarly, Delhi et al. [8] investigated the use of computer vision for real-time PPE detection in construction sites. Their dataset was collected through manual gathering and web crawling, and it contained approximately 2,500 images classified into four categories: NOT SAFE, SAFE, NoHardHat, and NoJacket. YOLOv3 was trained after augmentation, including rotations and inversion, to increase robustness. The dataset was split into training, validation, and testing sets in a 90%, 8%, and 2% ratio, respectively. Their results showed promising performance, with high values for mAP, recall, and F1-score on the test set.

Wang et al. [9] further explored real-time object detection for worker safety by comparing several YOLO-based detectors. In their study, they evaluated YOLOv3, YOLOv4, and YOLOv5 on the CHV dataset, which contains 1,330 images and six classes, including workers, waistcoats, and helmets of different colors [10]. The results showed that YOLOv5x achieved the highest mAP, while YOLOv5s offered the fastest inference speed. This comparison highlighted the trade-off between accuracy and processing speed, which remains an important issue in practical PPE detection systems.

Hayat and Morgado-Dias [11] also studied real-time detection of workers' heads and helmets using several YOLO versions, including YOLOv3, YOLOv4, and YOLOv5x. Their experiments were carried out using the public MakeML dataset [12], which included 3,000 training instances and 1,000 testing instances. The study focused only on the "Head" and "Helmet" classes. To address issues related to illumination and contrast, the images were preprocessed using power-law transformation [13]. Among the tested models, YOLOv5x performed the best, achieving 92% accuracy, 92.4% precision, 89.2% recall, and 90.8% F1-score.

In another line of work, Ma et al. introduced a PPE detection method based on YOLOv4 using around 25,000 images collected from construction-site surveillance footage. Their dataset was divided into multiple classes and split into training and testing subsets. Both YOLOv4 and YOLOv4-Tiny were evaluated, and model pruning was used to improve efficiency. The optimized CLSlim YOLOv4 model showed only a small loss in mAP while being significantly faster and much lighter in terms of parameters. This study demonstrated that lightweight optimization techniques can be highly effective when deploying PPE detection models on resource-constrained systems [14].

Gallo et al. [15][16] developed a PPE detection system for hazardous industrial environments using deep learning models applied to video streams. They

compared multiple architectures, including YOLOv4, YOLOv4-Tiny, SSD MobileNet, CenterNet, and EfficientDet. The experiments were conducted on three datasets of different sizes, and YOLOv4-Tiny produced the best balance between speed and accuracy. Their findings support the use of lightweight detectors in real-time monitoring scenarios where processing speed is essential.

Nath, Behzadan, and Paal [17] proposed intelligent approaches to reduce construction fatalities by verifying PPE compliance through a series of deep learning and machine learning techniques. Their work first identified workers and safety items such as headwear and vests, and then applied additional classifiers to determine whether PPE was worn correctly. They also explored a CNN-based framework for simultaneous worker and PPE verification. A third strategy involved detecting workers first and then classifying cropped regions using CNN-based models such as VGG-16, ResNet-50, and Xception. Their models were trained on the Pictor-v3 dataset [18], which contains annotated images and worker instances with different PPE combinations. Among the proposed methods, the second approach produced the strongest results in real-world conditions, reaching a mAP of 72.3% and processing 11 frames per second on a laptop, which made it suitable for lightweight deployment.

Torres et al. [19] presented a cognitive safety analysis component for a monitoring system based on CCTV footage. Their system used deep learning to identify correct PPE usage in real time. The reported results showed a mAP of 80.19% and a processing speed of 80 FPS using YOLOv4, which indicates that such systems can support efficient safety monitoring in industrial settings.

Benedetto et al. [20] explored PPE detection using both synthetic and real images of people wearing or not wearing protective equipment such as helmets, respirators, waistcoats, and ear protection. Their study employed transfer learning with YOLO and Faster R-CNN models pretrained on the COCO dataset. The models were fine-tuned using different combinations of synthetic-only, real-only, and mixed datasets. Synthetic images were generated using Rage Plugin Hook [21], an API for GTA V that allowed the creation of realistic training scenes. The dataset included 126,900 virtual images for training and 350 for evaluation. The results showed that models trained on synthetic data performed better on similar synthetic inputs, but performance dropped when tested on real images. The best performance on real data was obtained by Faster R-CNN fine-tuned on both synthetic and real images, reaching an overall mAP of 77.1%.

Isailovic et al. [22] studied the detection of head-mounted protective devices using pre-trained architectures such as YOLOv5, MobileNetv2 SSD, and Faster R-CNN to locate the region of interest around the head. Their dataset consisted of 12,682 images spanning 12 PPE categories and was collected from the Faculty of Medicine at the University of Belgrade, along with public sources such as Roboflow [15] and Pictor PPE [17]. Among the tested models, YOLOv5 achieved the best overall performance, with 92% precision and 61.1% recall. The “Hardhats” category performed especially well, with a precision of 100% and a recall of 96%.

Cengis et al. [23] focused on the detection of metal headwear using an improved one-stage detector based on YOLOv5. Their main contribution was in the feature extraction stage, where alternative backbones such as ShuffleNetv2 and MobileNetv3 were examined to improve speed and efficiency. The study classified three classes: Helmet, Head, and Person, using the Hard Hat dataset from Roboflow [16]. The authors suggested that future studies would benefit from larger datasets with more detailed class coverage.

Finally, Kisa Zehra et al. [24] addressed PPE compliance in construction sites by developing a deep neural network to determine whether workers were wearing helmets. Their dataset, sourced from the Harvard Dataverse [25], contained 7,063 images of construction laborers in different poses and locations, both with and without headgear. The experimental results showed that YOLOv5x achieved the highest performance among the tested models, with a mAP50 of 95.8%, precision of 93.9%, recall of 91.2%, and F1-score of 92.5%.

Overall, the reviewed studies show that YOLO-based detectors have become the preferred choice in PPE detection because they provide a strong balance between accuracy and inference speed. At the same time, the literature also shows that performance is still influenced by dataset quality, class diversity, occlusion, and deployment conditions. These findings motivate the present study to focus on a fine-tuned YOLOv8 model trained on a construction-site safety dataset, with the aim of achieving reliable PPE detection under realistic conditions.

Chapter 3 Proposed Approach (Methodology)

3.1 Introduction

This chapter presents the methodological framework adopted to develop the proposed PPE detection system using a deep learning-based object detection approach.

The methodology is designed to enable accurate identification of safety compliance in construction environments through the utilization of a fine-tuned YOLOv8 detection model trained on a construction site safety dataset.

The implementation process followed a structured pipeline consisting of dataset preparation, preprocessing, model selection, transfer learning-based training, and performance evaluation. These stages were carefully designed to ensure that the developed detection system achieves reliable performance while maintaining suitability for real-time workplace safety monitoring applications.

3.2 Dataset Description

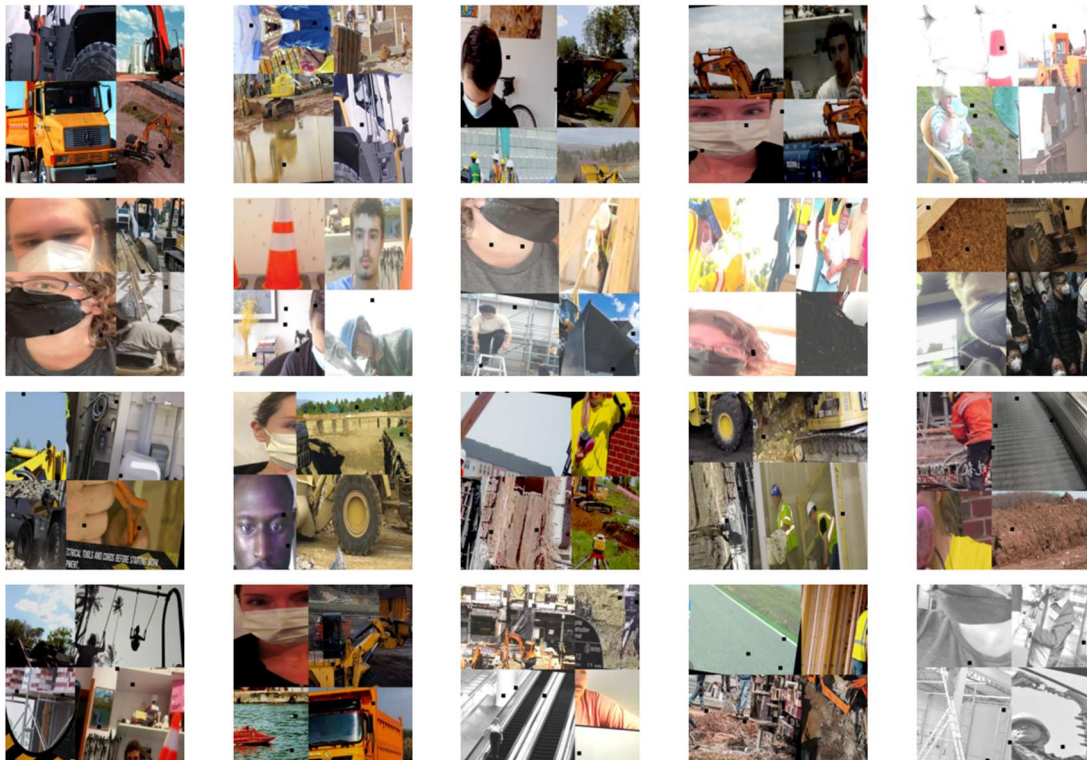
The performance of deep learning-based detection models depends heavily on the characteristics of the dataset used during training. For this reason, particular attention was given to selecting a dataset that reflects realistic construction-site conditions where personal protective equipment compliance must be monitored.

In this study, the **Construction Site Safety Image Dataset**, obtained from Kaggle and organized using the Roboflow annotation platform, was adopted for training and evaluation purposes. The dataset contains images captured from actual construction environments and includes multiple annotated object categories related to worker safety. These annotations provide bounding-box information for workers as well as several PPE elements such as safety helmets and reflective safety vests.

An important advantage of this dataset is that many images contain more than one object instance within the same scene. This characteristic allows the detection model to learn spatial relationships between workers and safety equipment under conditions that closely resemble real operational environments rather than simplified laboratory scenarios.

Furthermore, the dataset includes natural variations in illumination levels, viewing angles, background clutter, and object sizes. These variations are particularly valuable during training because they help the model learn more general visual representations and improve its ability to operate reliably when applied to unseen construction-site images.

An example of the annotation format used during training is illustrated in Figure 1.



(Figure 1: Example annotated image from the Construction Site Safety Dataset)

3.3 Dataset Preprocessing

Before initiating the training stage, several preprocessing steps were applied to prepare the dataset for compatibility with the YOLOv8 training pipeline and to ensure stable learning behaviour during fine-tuning.

First, all annotation files were verified to match the YOLO object detection format, where each image is associated with a corresponding label file containing normalized bounding-box coordinates. This step ensured that object location information could be interpreted correctly by the detection framework during training.

After validating the annotation structure, the dataset was divided into three subsets: training, validation, and testing. The training subset was used to update the network parameters, while the validation subset supported hyperparameter adjustment and helped monitor the learning progress during training. The testing subset remained unused during training and was reserved exclusively for evaluating the generalization capability of the trained model.

In addition to dataset partitioning, image resizing operations were handled automatically within the YOLOv8 training environment in order to match the required network input resolution. This resizing process preserved essential spatial features necessary for accurate object localization.

To further improve robustness, built-in augmentation techniques provided by the YOLOv8 framework were enabled during training. These augmentation operations introduced controlled variations in object appearance and scene conditions, which helped the model become more resilient to illumination changes, viewpoint differences, and background complexity typically encountered in real construction environments.

.3.4 Proposed Detection Model Architecture

The detection component developed in this work is based on the YOLOv8 architecture, which was selected due to its ability to achieve a practical balance between inference speed and detection accuracy in real-time monitoring scenarios.

Unlike traditional two-stage detectors, YOLOv8 follows a single-stage detection strategy in which object localization and classification are performed simultaneously within a unified neural network structure. This design reduces computational overhead and makes the model particularly suitable for safety-monitoring applications where timely detection is essential.

Structurally, the YOLOv8 network consists of three main modules: the backbone, the neck, and the detection head. The backbone extracts hierarchical spatial features from the input images through convolutional layers operating at multiple resolution levels. These features are then refined inside the neck module, which combines semantic information from different feature scales in order to strengthen the representation of objects appearing at various sizes within the image. Finally, the detection head produces bounding-box coordinates together

with class probability estimates corresponding to the PPE categories considered in this study.

This architecture enables the detection model to identify safety equipment efficiently even when objects appear at different distances or under varying scene conditions.

3.5 Transfer Learning Strategy

Training deep convolutional detection models from scratch typically requires very large annotated datasets and extensive computational resources. Since the available dataset in this study represents a domain-specific construction safety scenario, a transfer learning strategy was adopted to improve both training efficiency and detection performance.

Specifically, the YOLOv8s pretrained model was used as the starting point for the training process. The pretrained weights provided an initial representation learned from large-scale object detection datasets, allowing the network to reuse general visual features such as edges, textures, and object structures.

After initialization, the model was fine-tuned using the construction-site safety dataset so that these general representations could be adapted to PPE detection tasks. This approach significantly reduced the required training time while improving convergence stability during optimization.

As a result, the model became capable of focusing more effectively on safety-related object categories relevant to construction environments.

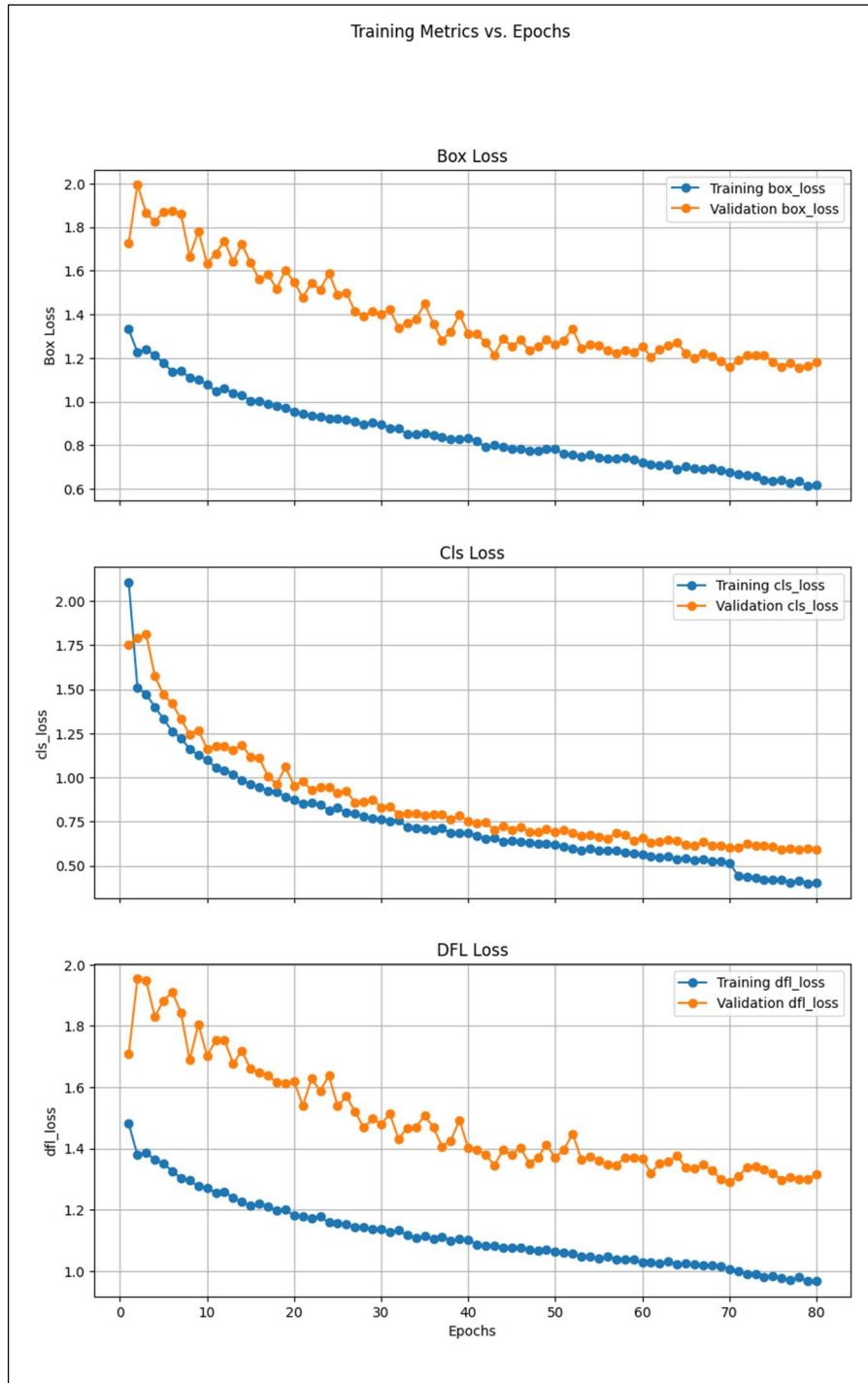
3.6 Model Training Configuration

The training experiments were conducted using the Ultralytics YOLOv8 implementation within the Kaggle GPU environment. This setup provided sufficient computational capability for performing iterative fine-tuning while maintaining a practical experimental workflow.

During training, the pretrained YOLOv8s weights were updated progressively across multiple epochs using the annotated dataset described earlier. Input images were resized to a fixed resolution compatible with the detection network so that both global contextual information and fine spatial details could be preserved during feature extraction.

Throughout the training process, several evaluation indicators were monitored continuously in order to assess convergence behaviour and detection quality. These indicators included training loss, validation loss, precision, recall, and mean average precision (mAP). Observing these metrics during training made it possible to detect early signs of overfitting and helped guide the optimization process toward achieving stable detection performance.

Training performance curves obtained during the optimization process are illustrated in Figure 2.



(Figure 2: Training and validation performance curves)

3.8 Summary

This chapter described the methodological steps followed to develop the proposed PPE detection system based on the YOLOv8 object detection framework. The discussion covered the dataset selection process, the preparation and preprocessing procedures applied before training, and the transfer learning strategy adopted to adapt the pretrained detection model to construction-site safety scenarios. In addition, the configuration of the training environment and the selected optimization settings were presented to clarify how the experimental pipeline was implemented in practice.

The next chapter presents the experimental evaluation of the trained detection model and analyzes its performance using quantitative indicators derived from the validation and testing datasets together with visual inspection of representative detection outputs.

.

Chapter 4 :Experimental Results and Discussion

4.1 Introduction

This chapter presents the experimental evaluation of the PPE detection system developed in this study using the fine-tuned YOLOv8 model. The purpose of this evaluation is to examine how effectively the trained detection model can identify workers and safety equipment in construction-site images that reflect realistic environmental conditions.

To provide a comprehensive assessment, the trained model was analyzed using commonly adopted object detection performance indicators such as Precision, Recall, and mean Average Precision (mAP). In addition to these numerical measures, visual inspection of selected detection outputs was also performed in order to better understand how the model behaves when applied to images containing multiple objects, scale variations, and complex backgrounds. The obtained results are interpreted with reference to both the dataset characteristics and the structure of the YOLOv8 detection architecture.

4.2 Detection Performance Results

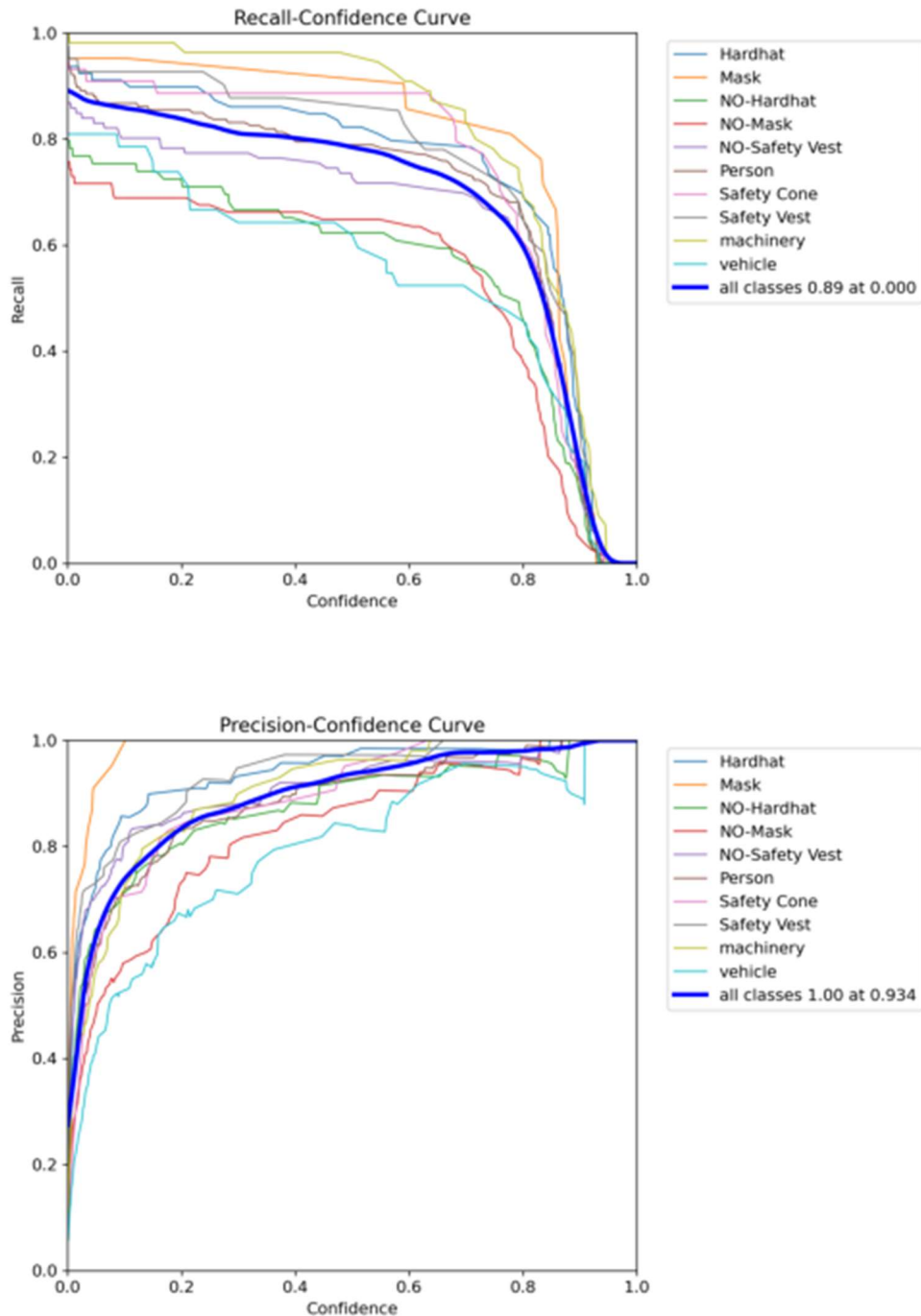
The trained YOLOv8 model demonstrated consistent performance in detecting PPE-related objects across both validation and testing datasets. The model was able to identify workers, safety helmets, and reflective safety vests in images captured under different scene conditions, including variations in background complexity and viewing perspectives.

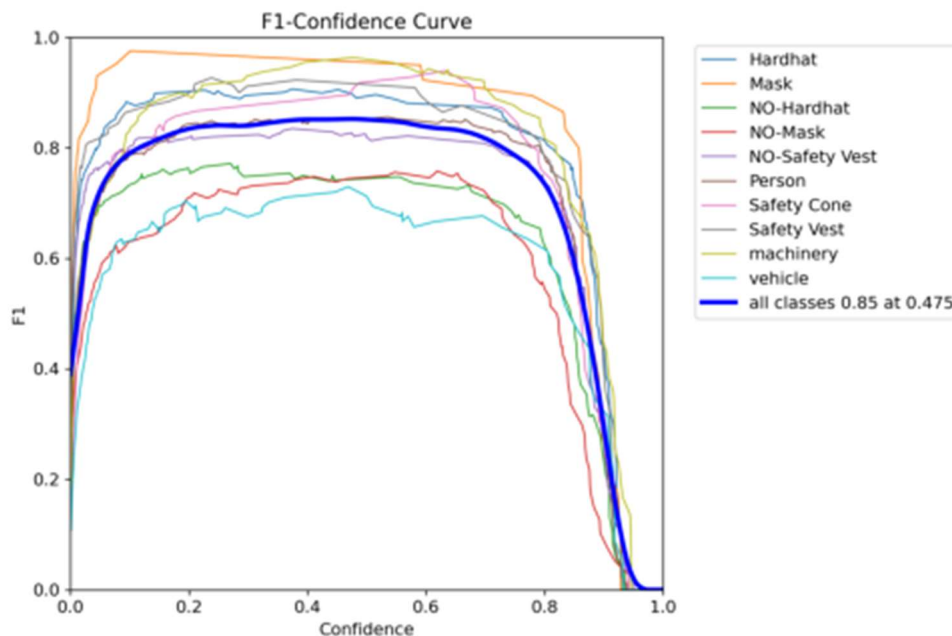
The obtained mean Average Precision at an intersection-over-union threshold of 0.5 (mAP@50) indicates that the model achieved reliable localization accuracy when detecting target objects. Furthermore, the mAP@50–95 evaluation range confirms that the detection performance remained stable even when stricter localization criteria were applied, which reflects the robustness of the trained detection pipeline.

Precision values observed during evaluation suggest that the majority of predicted detections corresponded correctly to actual object categories, indicating that the number of false positive detections remained limited. Similarly, the recorded recall values demonstrate that most relevant PPE objects present in the dataset were successfully identified by the detection model,

which reflects good coverage of safety-related object instances during inference.

A graphical representation of the Precision–Recall relationship obtained during model evaluation is illustrated in Figure 3.





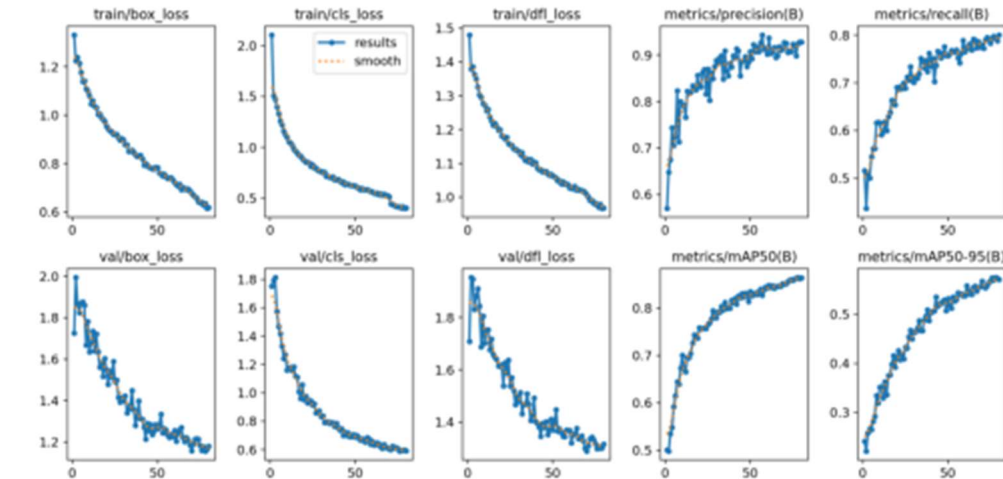
(Figure 3: Precision–Recall curve of the trained YOLOv8 model)

4.3 Training Behavior and Model Convergence

The training process showed stable convergence behaviour throughout the experimental iterations. At the early stages of training, loss values decreased rapidly, indicating that the pretrained detection network was able to adapt efficiently to the PPE detection task. As training progressed, the rate of improvement became more gradual, suggesting that the optimization process approached a stable solution.

An important observation is that no significant gap appeared between training loss and validation loss during later epochs. This behaviour indicates that the model maintained balanced learning performance without showing clear signs of severe overfitting. The use of transfer learning together with built-in augmentation strategies available in the YOLOv8 framework contributed to improving the model's ability to generalize beyond the training subset.

The evolution of training and validation loss values across epochs is illustrated in Figure 4.



(Figure 4: Training and validation loss curves)

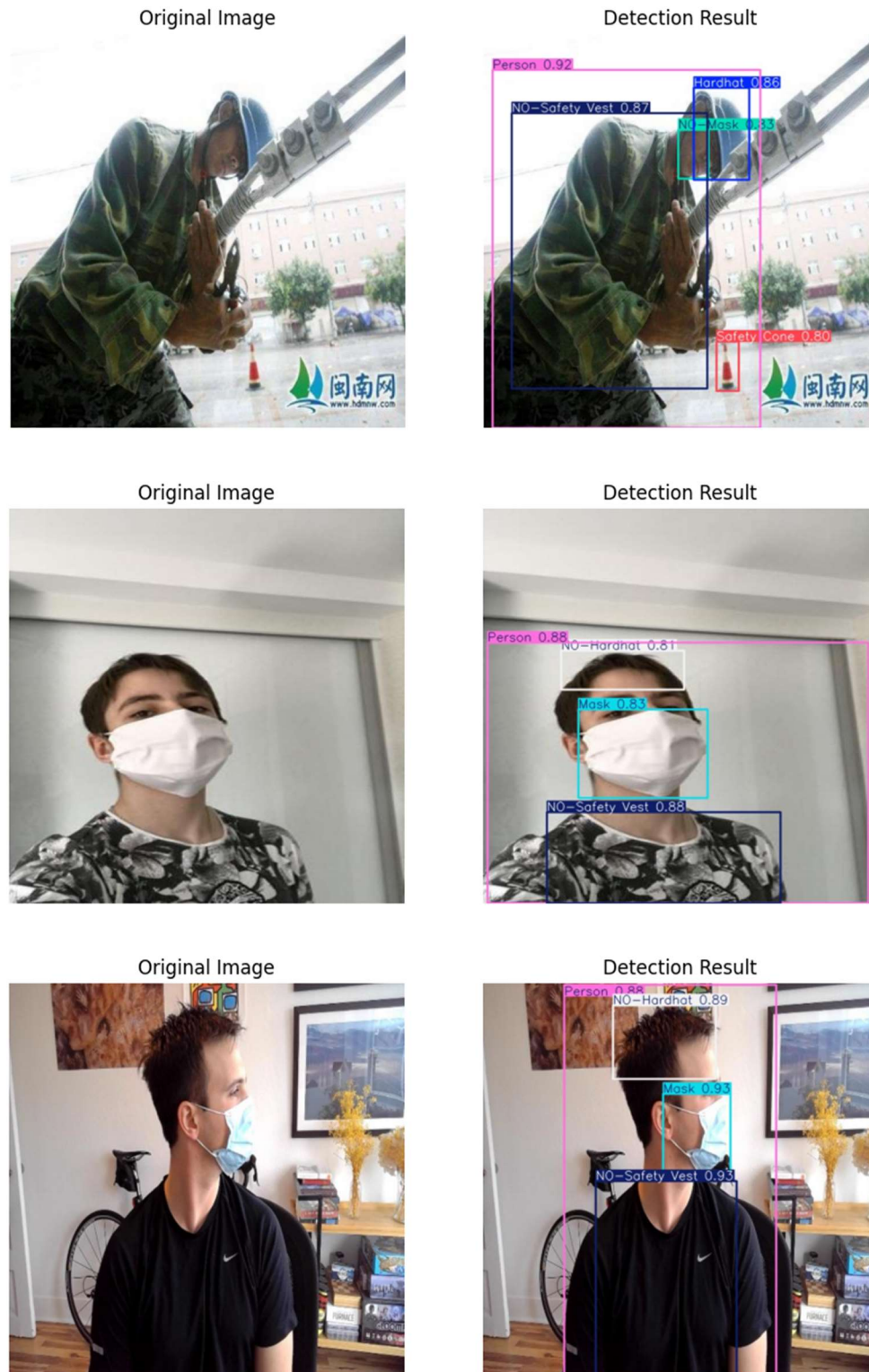
4.4 Qualitative Analysis of Detection Results

In addition to the numerical evaluation results, visual inspection of detection outputs was performed in order to better understand how the trained model behaves under practical operating conditions. Several representative test images were analyzed to observe detection accuracy when multiple objects appear simultaneously within the same scene.

The experimental results show that the trained model was capable of detecting PPE components under a variety of realistic conditions, including scenes containing partial occlusions, crowded object arrangements, and variations in illumination. The model also maintained stable localization performance when objects appeared at different scales within the same image.

However, a small number of challenging cases were observed in situations involving strong occlusion effects or very small object instances. In these scenarios, detection confidence decreased slightly compared with normal conditions. Despite these limitations, the overall qualitative results confirm that the proposed detection approach remains suitable for practical safety-monitoring applications in construction environments.

Examples of representative detection outputs are shown in Figure 5.



(Figure 5: Sample detection results produced by the proposed model)

4.5 Discussion of Findings

The results confirm that the YOLOv8-based detection model gives an effective solution for automated PPE monitoring in work related environments. The combination of precision and recall values shows that the model achieves a good balance between removing false detections and increasing object coverage.

Compared with traditional object detection approaches, the use of a one-stage detector such as YOLOv8 helps real time performance without compromising accuracy. This is essential in safety applications, where there is a requirement for quick response to non-compliance .

The use of transfer learning helped to improve performance of model, as it allowed the detection system to allow pre-learned visual features and use them to the PPE detection task. More than this, the vast variety of the dataset plays a role in increasing the validity of the model under varying environmental conditions.

Overall, the findings show that the proposed approach is quite good for deployment in safety monitoring systems aimed to reduce workplace accidents and improve compliance with safety rules regulations.

4.6 Limitations of the Proposed Work

Despite the results we got from this study, many limitations should be considered. One of the limitations is the dependency of performance on the quality and variety of the training dataset. Although the dataset used in the study got different environmental conditions, it still might not be able to fully represent all possible real scenarios encountered on worksites.

Else then this, the model may come across reduced detection accuracy in some cases such as extreme occlusion, low-resolution objects, or highly cluttered backgrounds. These are the common challenges in object detection tasks and may need further model optimization or dataset expansion.

Another limitation is related to computational requirements. While YOLOv8 provides fast and accurate detection, using the model in resource-constrained

environments may require further optimization techniques such as model pruning, quantization etc.

4.7 Summary of the Chapter

This chapter tells us a detailed explanation of the proposed PPE detection system working on the basis of the YOLOv8 architecture. The results tell us that the model has the potential to achieve strong detection performance in terms of precision, recall, and mean Average Precision.

Both the quantitative analysis and the qualitative analyses shows the effectiveness of the proposed approach to detect PPE components in realistic worksite conditions. There are certain limitations but the overall performance indicates that the system can work as a reliable solution for automated safety monitoring of the workplace.

Chapter 5 Conclusion and Future Work

This study presented the development and evaluation of a deep learning-based detection system designed to identify personal protective equipment in construction-site environments using the YOLOv8 object detection architecture. The main objective of the work was to design an automated solution capable of recognizing essential safety elements such as workers, helmets, and reflective vests in order to support monitoring activities in workplace safety scenarios.

To achieve this objective, a construction-site image dataset collected from publicly available sources including Kaggle and Roboflow was used for training and evaluation. A transfer learning strategy was adopted to fine-tune the pretrained YOLOv8s detection model so that it could adapt effectively to PPE-related object categories. The experimental results demonstrated that the trained detection model achieved reliable performance across different scene conditions, confirming the suitability of modern object detection techniques for supporting automated safety-monitoring tasks.

The results obtained in this study indicate that intelligent PPE detection systems can support safety supervision processes by reducing reliance on manual observation and enabling faster identification of non-compliance situations inside construction environments. Such systems can assist safety personnel in monitoring multiple work areas more efficiently and may contribute to improving overall workplace safety practices.

Although the achieved performance was encouraging, several possible extensions can be considered in future work. One important direction involves deploying the detection model in real-time monitoring scenarios using live surveillance camera streams instead of still image datasets. This step would allow the system to operate more effectively in practical construction-site environments. Another potential improvement includes expanding the dataset by incorporating images collected from additional construction locations with different environmental characteristics in order to further strengthen model generalization capability.

Future studies may also investigate the detection of additional protective equipment categories such as gloves and safety glasses to extend the monitoring

coverage provided by the proposed system. Furthermore, integrating the trained detection model with IoT-based safety monitoring platforms could enable automatic alert mechanisms when violations are detected, supporting more proactive workplace safety management strategies.

Overall, the findings of this research demonstrate that deep learning-based PPE detection systems represent a practical and scalable approach for improving construction-site safety monitoring and provide a strong foundation for further development of intelligent safety-assistance solutions in industrial environments.

References

- [1] Richard Szeliski, *Computer Vision: Algorithms and Applications*, Springer, 2010.
- [2] David Lowe, “Distinctive Image Features from Scale-Invariant Keypoints,” *International Journal of Computer Vision*, 2004.
- [3] Yann LeCun, Yoshua Bengio, Geoffrey Hinton, “Deep Learning,” *Nature*, 2015.
- [4] Ross Girshick, “Fast R-CNN,” *IEEE International Conference on Computer Vision (ICCV)*, 2015.
- [5] Joseph Redmon et al., “You Only Look Once: Unified, Real-Time Object Detection,” *CVPR*, 2016.
- [6] Chien-Yao Wang et al., “YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors,” 2022.
- [7] Chen, S.; Demachi, K. A vision-based approach for ensuring proper use of personal protective equipment (PPE) in decommissioning of fukushima daiichi nuclear power station. *Appl. Sci.* **2020**, *10*, 5129.
- [8] Delhi, V.S.K.; Sankarlal, R.; Thomas, A. Detection of Personal Protective Equipment (PPE) Compliance on Construction Site Using Computer Vision Based Deep Learning Techniques. *Front. Built Environ.* **2020**, *6*, 136.
- [9] Wang, Z.; Wu, Y.; Yang, L.; Thirunavukarasu, A.; Evison, C.; Zhao, Y. Fast Personal Protective Equipment Detection for Real Construction Sites Using Deep Learning Approaches. *Sensors* **2021**, *21*, 3478. [CrossRef] [PubMed]
- [10] ZijianWang-ZW. (n.d.). Real-Time PPE Detection Based on YOLO. Open High-Quality Dataset. Available online: https://github.com/ZijianWang-ZW/PPE_detection (accessed on 20 May 2023).
- [11] Hayat, A.; Morgado-Dias, F. Deep Learning-Based Automatic Safety Helmet Detection System for Construction Safety. *Appl. Sci.* **2022**, *12*, 8268.
- [12] Hard Hat Workers Dataset|MakeML—Create Neural Network with Ease. Available online: <https://makeml.app/datasets/hardhat-workers> (accessed on 17 May 2023).
- [13] Kumar, D.; Ramakrishnan, A.G. Power-law transformation for enhanced recognition of born-digital word images. In *Proceedings of the 2012 International Conference on Signal Processing and Communications (SPCOM)*, Bangalore, India, 22–25 July 2012; pp. 1–5. [CrossRef]

- [14] Ma, L.; Li, X.; Dai, X.; Guan, Z.; Lu, Y. A Combined Detection Algorithm for Personal Protective Equipment Based on Lightweight YOLOv4 Model. *Wirel. Commun. Mob. Comput.* **2022**, *2022*, 3574588.
- [15] Gallo, G.; Di Rienzo, F.; Garzelli, F.; Ducange, P.; Vallati, C. A Smart System for Personal Protective Equipment Detection in Industrial Environments Based on Deep Learning at the Edge. *IEEE Access* **2022**, *10*, 110862–110878.
- [16] Hard Hat Workers Dataset. Available online: <https://public.roboflow.com/object-detection/hard-hat-workers> (accessed on 18 May 2023). Nath, N.D.; Behzadan, A.H.; Paal, S.G. Deep learning for site safety: Real-time detection of personal protective equipment. *Autom.Constr.* **2020**, *112*, 103085.
- [17] PPE Object Detection Dataset (V9, Pictor-V3-Revised) by PPE. Roboflow. Available online: <https://universe.roboflow.com/ppeorxtt/ppe-u7jtr/dataset/9> (accessed on 1 June 2023).
- [18] Torres, P.; Davys, A.; Silva, T.; Schirmer, L.; Kuramoto, A.; Itagyba, B.; Salgado, C.; Comandulli, S.; Ventura, P.; Fialho, L.; et al. A Robust Real-time Component for Personal Protective Equipment Detection in an Industrial Setting. In *Proceedings of the 23rd International Conference on Enterprise Information Systems (ICEIS 2021)*, Online Streaming, 26–28 April 2021; pp. 693–700.
- [19] Di Benedetto, M.; Carrara, F.; Meloni, E.; Amato, G.; Falchi, F.; Gennaro, C. Learning accurate personal protective equipment detection from virtual worlds. *Multimed. Tools Appl.* **2020**, *80*, 23241–23253. [CrossRef]
- [20] What Is RAGE Plugin Hook? RAGE Plugin Hook. Available online: <https://ragepluginhook.net/> (accessed on 16 May 2023).
- [21] Isailovic, V.; Peulic, A.; Djapan, M.; Savkovic, M.; Vukicevic, A.M. The compliance of head-mounted industrial PPE by using deep learning object detectors. *Sci. Rep.* **2022**, *12*, 16347. [CrossRef] [PubMed]
- [22] Cengil, E.; Çınar, A.; Yıldırım, M. An Efficient and Fast Lightweight-Model with ShuffleNetv2 Based on YOLOv5 for Detection of Hardhat-Wearing. *Rev. Comput. Eng. Stud.* **2022**, *9*, 116–123. [CrossRef]
- [23] Kisaiezehra; Farooq, M.U.; Bhutto, M.A.; Kazi, A.K. Real-Time Safety Helmet Detection Using Yolov5 at Construction Sites. *Intell. Autom. Soft Comput.* **2022**, *36*, 911–927. [CrossRef]
- [24] Xie, L. Hardhat-Hardhat-Wearing. 2019. Available online: <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/7CBGOS> (accessed on 19 May 2023).