



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 express my deepest gratitude and appreciation to my supervisor Dr. khaled_Eskaf for his guidance, continuous encouragement, and support during our project life cycle. Our appreciation is also extended to committee members. Finally, I would like to extend a special thank you all the staff of the Science in Artificial Intelligence Department and all persons who had helped us in all fields during our study..

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	vii
List of Figures.....	vii
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].

Nowadays object detection algorithms have two categories: two-stage detectors and one-stage detectors. Two-stage detectors, for instance Region-Based Convolutional Neural Network (R-CNN) family, it first generates region proposals and then classify objects within those regions. These methods provide high detection accuracy, they require longer processing time compared with one-stage detectors [4].

On the other hand, one-stage detectors perform object localization and classification in a single step, making them suitable for real-time applications.

2.3 Related Work

A vision-based approach for monitoring personal protective equipment (PPE) in a nuclear power station was proposed in a study conducted by Chen and Demachi [7]. The experiment was conducted on a 3808-image annotated dataset that was gathered from the actual world using a webcam and the

web using a web crawler application. During the testing phase, an alternative dataset was analyzed, which was also collected manually. Nevertheless, the detection model was designed to identify and classify two specific objects: a full-face mask and a hard hat. Furthermore, the distance and posture of the laborers were factors that facilitated the accomplishment of the study's objective. In two phases, the combined dataset was used to train the one-stage detection model YOLOv3. The first stage involved locking the final convolutional layer in Darknet-53, while the second stage involved unfreezing all convolutional layers to complete the fine-tuning process. The authors plan to enhance the dataset's accuracy and deploy the model using the on-site surveillance system in the future. Delhi et al. [8] conducted a study that utilized computer vision, a subset of deep learning capabilities, to detect personal protective equipment (PPE) in real time, thereby guaranteeing the highest level of safety at construction sites. Nevertheless, the researchers painstakingly collected the dataset, in addition to conducting web crawling, for the purpose of conducting the study. The model was trained using approximately 2500 images, which were categorized into four target classes: NOT SAFE, SAFE,

NoHardHat, and NoJacket. Additionally, YOLOv3 was trained on the data samples following the augmentation phase, based on the data. This allows the model to be generalized and robust by performing left and right rotations of 30 degrees and inverting. The dataset that was granted was partitioned using a train-validation-test method, with random dividing of 90%, 8%, and 2% for training, validation, and testing, respectively. Based on the test data, the model obtained a mAP, recall, and F1-score of 97%.

The authors of Wang et al.'s study [9] utilized deep learning neural networks that were adjusted for real-time object detection to ensure that workers adhered to safety protocols. This approach was taken to address the issue of workers' safety. Consequently, they suggested three deep-learning detectors for six classes, namely YOLOv3, YOLOv4, and YOLOv5, which are based on YOLO architectures.

The data that was employed consisted of a high-quality dataset, CHV, which was divided into six classes: headgear with four colors, person, and waistcoat. The dataset contained 1330 images [10]. According to the experiment results, YOLOv5x obtained the highest mAP of 86.55%, while YOLOv5s achieved the quickest GPU performance of 52 FPS. In a study conducted by Hayat and Morgado-Dias [11], a deep learning technique is introduced to detect the heads and headgear of construction site laborers in real-time. In this paper, the popular deep learning architecture YOLO was examined in various iterations, including YOLOv3, YOLOv4, and YOLOv5x.

The authors employed the public dataset by MakeML [12] to implement the model. A total of 3000 instances were utilized for training, while 1000 instances were employed for testing. Additionally, this investigation exclusively employed the "Head" and "Helmet" classes. The images were subjected to the Power-law transformation [13] for image pre-processing in order to address the illumination and contrast issues in the data. YOLOv5x was the most effective model, achieving 92% accuracy, 92.4% precision, 89.2% recall, and 90.8% F1-score, respectively.

The portable YOLOv4 model was employed by Ma et al. [14] to develop a combination detection algorithm for PPE. The dataset utilized consisted of approximately 25,000 samples that were extracted from security footage of a

building construction site. The data were randomly divided into six classes and subsequently separated into a training set and a test set. Two algorithms, YOLOv4 and YOLOv4-Tiny, were implemented. The finetuning and optimization strategy to enhance accuracy involved pruning the model with the original dataset. The optimal result was achieved.

The portable YOLOv4 model was employed by Ma et al. [14] to develop a combination detection algorithm for PPE. The dataset utilized consisted of approximately 25,000 samples that were extracted from security footage of a building construction site. The data were randomly divided into six classes and subsequently separated into a training set and a test set. Two algorithms, YOLOv4 and YOLOv4-Tiny, were implemented. The finetuning and optimization strategy to enhance accuracy involved pruning the model with the original dataset. The CLSlim YOLOv4 model achieved the most favorable results, with a mAP loss of only 2.1%, an inference speed that was 1.8 times faster, and a compression of the model parameters by 98.2%. The primary conclusions underscored the effectiveness of the channel and layer pruning method (CLSlim) in improving the detection speed and reducing computational power consumption. It is recommended that additional research be conducted to investigate the integration of CLSlim with other lightweight strategies in order to accelerate model inference and identify more effective lightweight model techniques for mobile terminals with limited resources. Gallo et al. [15] proposed a system for detecting personal protective equipment (PPE) in hazardous industrial environments. They implemented the system by utilizing a deep neural network to analyze a video stream. Five models were trained to determine whether the laborers are donning equipment or not: YOLOv4, YOLOv4-Tiny, SSD MobileNet, CenterNet, and EfficientDet. The first dataset, which is publicly available [16], contains 7035 images. The other two datasets were collected using controlled settings and contained 215 and 236 images, respectively. Three datasets were utilized. The system was implemented with YOLOv4-tiny, which obtained a mAP of 86%, as a result of its rapid detection capabilities.

Nath, Behzadan, and Paal have suggested intelligence-based solutions to address construction fatalities that are the result of brain injuries and collisions [17]. This study suggested the development of a DL model that is based on the YOLO architecture and is intended to verify the compliance of personal protective equipment (PPE). Subsequently, the subsequent phase involves the creation of a machine learning model that will ascertain whether the worker is appropriately donning personal protective equipment (PPE). This method

involves the identification of workers, their headwear, and vests, and the subsequent use of a machine learning model (e.g., decision trees and neural networks) to determine whether each worker is appropriately donning their hat or vest. Secondly, a convolutional neural network (CNN) framework is employed to simultaneously verify PPE compliance and identify individual laborers. A third method involves initially identifying only laborers in the input image. The laborers are subsequently cropped and classified using CNN-based classifiers, such as VGG-16, ResNet-50, and Xception, based on the presence of personal protective equipment. The 23 models are trained on an in-house image dataset that is compiled through Web24 mining and crowdsourcing. 1500 annotated images and 4700 instances of 25 workers donning differing combinations of personal protective equipment are contained in the dataset Pictor-v3 [18]. The second approach, 26 generates the most favorable results in real-world environments, with a mAP of 72.3%, according to the analysis. It is capable of processing 11 frames per second on a laptop computer, which allows it to operate on lightweight mobile devices and perform real-time detection. A novel cognitive safety analysis component for a monitoring system has been proposed by Torrse et al. in another study [19]. The proposed system is implemented in this investigation to identify the appropriate utilization of personal protective equipment in real-time, as determined by CCTV camera data. In order to identify elements within the system, a deep learning algorithm was implemented. A system with a mAP of 80.19% was achieved within a real-time frame rate of 80 FPS using YOLOv4, as indicated by the findings of our study.

In a study conducted by Benedetto [20] et al., the detection of safety equipment was accomplished by utilizing computer-generated images of individuals who were wearing protective equipment, including a helmet, respirator, waistcoat, and ear protection, and those who were not.

The paper implemented transfer learning techniques, specifically YOLO and Faster-RCNN, which were pre-trained on Microsoft's COCO dataset. In addition, the pre-trained models were fine-tuned with either synthetic images only, real images only, or both synthetic and real images, after experimenting with three variations of fine-tuning. The authors utilized Rage Plugin Hook [21] to generate the synthetic data utilized in this paper. This tool functions as an API for the Grand Theft Auto V (GTA-V) computer game, enabling the generation of specific realistic images from the game. The total number of virtual images collected was 126,900 for training and 350 for assessment. The results of the experiments confirmed that the models did indeed obtain higher

scores when tested on virtual data, but they scored lower when tested on actual data.

The Faster R-CNN model, which was fine-tuned on both synthetic and real images, achieved the highest performance on real data. It achieved an overall mAP of 77.1%, with the "High-Visibility Vest" and "Person" classes having mAPs of 81.5% and 86.6%, respectively. Isailovic et al. [22] investigated the detection of head-mounted protection devices by utilizing a variety of pre-trained computer vision architectures, including YOLOv5, MobileNetv2 SSD, and Faster R-CNN, to derive the Region of Interest (ROI) surrounding the head. The dataset in question comprises 12,682 images that fall into 12 categories of personal protective equipment. The authors obtained it from the Faculty of Medicine at the University of Belgrade in Serbia and public sources, including Rob flow [15] and Pictor PPE [17]. The YOLOv5 model obtained the highest scores of all the models tested, with a precision of 92% and a recall of 61.1%. Additionally, the "Hardhats" category was the best in the YOLOv5 model, with a precision of 100% and a recall of 96%.

Cengiz et al. [23] concentrated on the detection of metal headwear by utilizing a one-stage object detector algorithm that was enhanced by the authors. The model's fundamental architecture was YOLOv5, with the primary improvement occurring during the feature extraction process. The architecture's backbone, ShuffleNetv2 and MobileNetv3, were evaluated to determine which model enhanced the network's efficiency and speed. The study classifies three objects in the image: "Helmet," "Head," and "Person," which are part of Rob flow's "Hard Hat" dataset [16]. The authors are of the opinion that the development and examination of exhaustive datasets with a greater number of classifications will be advantageous for future research. Kisa Zehra et al. [24] addressed the safety of construction sites by employing a deep neural network to determine whether or not a worker is donning a cap. The dataset was collected by Northeastern University in China and is housed in the Harvard Dataverse [25]. It comprises 7063 images of construction laborers in various poses and locations, both with and without headgear. The YOLOv5 extra-large network (YOLOv5x) demonstrated a significant improvement in the scores, as evidenced by its mAP50 of 95.8%, precision of 93.9%, recall of 91.2%, and F1-score of 92.5%, as demonstrated by the experimental results.

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 effectiveness of deep learning–based object detection models is strongly influenced by the quality and diversity of the dataset used during training.

In this study, the Construction Site Safety Image Dataset obtained from Kaggle and prepared using the Roboflow annotation framework was employed to train and evaluate the proposed PPE detection model.

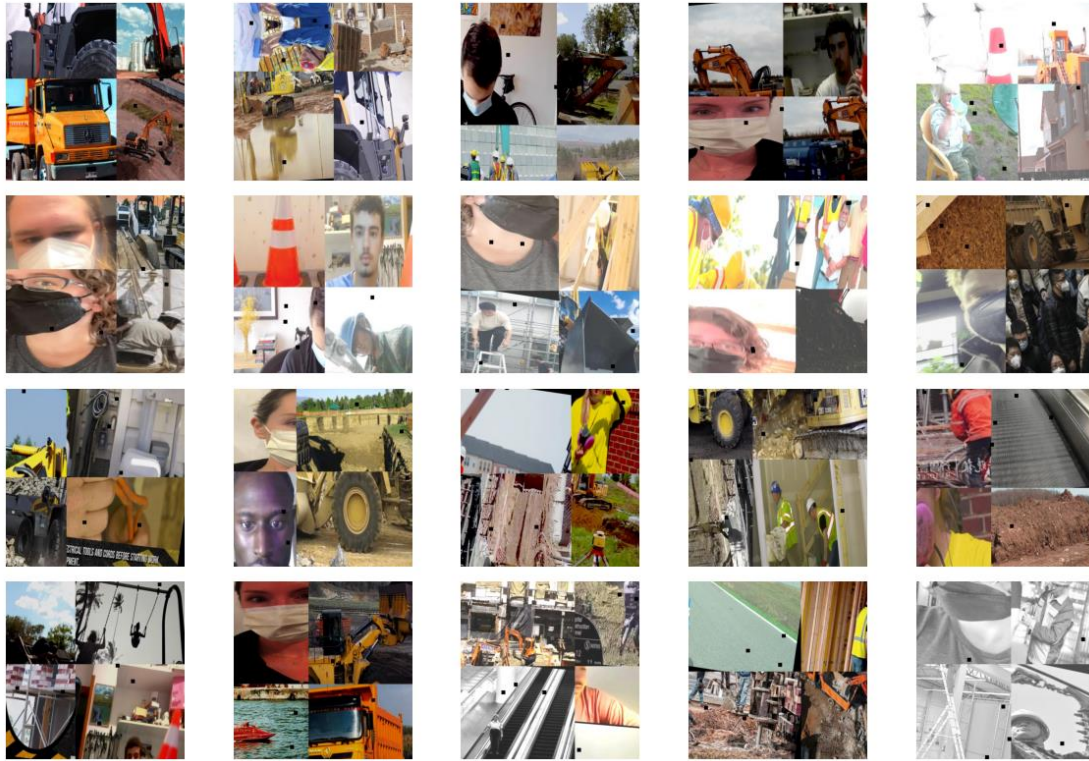
The dataset contains annotated images captured from construction site environments representing multiple safety-related object categories.

These annotations include bounding box labels identifying workers and various types of personal protective equipment such as safety helmets and reflective vests. The presence of multiple object instances within a single image enables the model to learn spatial relationships between workers and protective equipment under realistic working conditions.

In addition to class diversity, the dataset includes variations in lighting conditions, viewing perspectives, background complexity, and object scale. Such variations are essential for improving the robustness of deep learning detection

systems and enhancing their generalization capability when deployed in real-world industrial environments.

An example of dataset annotation used during the training process is presented in Figure 1.



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

3.3 Dataset Preprocessing

Prior to model training, several preprocessing procedures were applied to prepare the dataset for comity with the YOLOv8 training channel. These procedures guaranteed that the dataset was structured meetly and suitable for effective point literacy.

First, reflection lines were vindicated to ensure their thickness with the YOLO object discovery format, where each image is associated with a corresponding reflection train containing regularized bounding box equals describing object locales relative to image confines. This format enables effective integration with the YOLO training frame. Alternate, the dataset was divided into training,

confirmation, and testing subsets in order to support both model literacy and objective performance evaluation.

The training subset was used to optimize model parameters, while the confirmation subset supported hyperparameter tuning and helped overfitting during training.

The testing subset was reserved for estimating the conception capability of the trained discovery model using preliminarily unseen data. likewise, image resizing and normalization operations were automatically handled within the YOLOv8 training terrain. Input images were resized to match the network input resolution while conserving essential spatial features wanted for object localization tasks.

In addition, erected- in addition ways integrated within the YOLO training frame were employed to enhance dataset diversity. These addition ways contributed to amending the robustness of the discovery model against variations in illumination, object exposure, and scene complexity

.3.4 Proposed Detection Model Architecture

In this study, the YOLOv8 object detection architecture was used as the core detection framework because of its strong performance in real-time detection applications and its improved accuracy compared with older versions of the YOLO family.

YOLOv8 belongs to the type of single-stage object detectors, where object localization and classification are done simultaneously within a unified neural network architecture. This design helps us with faster inference speed while maintaining competitive detection accuracy, making the model particularly suitable for safety monitoring applications in dynamic construction environments.

The architecture of YOLOv8 consists of three main components: the backbone network, the feature aggregation neck, and the detection head. The backbone is responsible for extracting hierarchical spatial features from input images through convolutional operations. These extracted features are then refined within the neck module, which enhances multi-scale feature representation by combining semantic information from different network depths. Finally, the detection head

predicts bounding box coordinates and object class probabilities corresponding to detected PPE categories.

3.5 Transfer Learning Strategy

Training deep neural networks from scratch typically requires large-scale annotated datasets and extensive computational resources. To address this limitation, a transfer learning strategy was adopted in this study to improve training efficiency and detection accuracy.

Transfer learning enables the reuse of knowledge learned from large-scale datasets by initializing the detection model with pretrained weights before fine-tuning it on a domain-specific dataset. In this research, a pretrained YOLOv8s model was selected as the base detection architecture and subsequently fine-tuned using the construction site safety dataset.

This approach significantly reduced the required training time while improving feature extraction capability, since the pretrained model had already learned general visual representations applicable to object detection tasks. Fine-tuning allowed the network to adapt these representations specifically for PPE detection in construction environments.

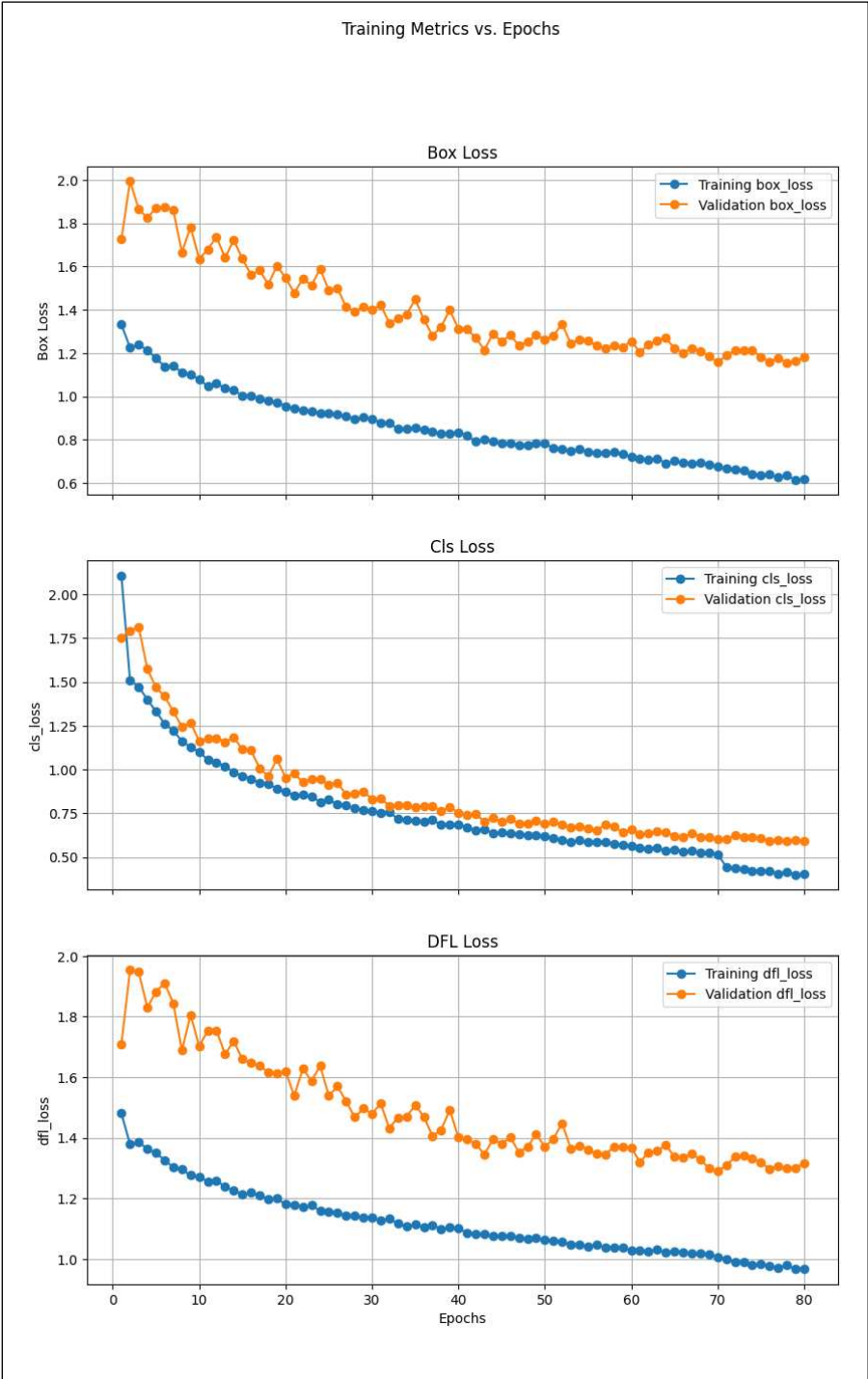
3.6 Model Training Configuration

The training process was conducted using the Ultralytics YOLOv8 framework within the Kaggle GPU environment. The selection of training parameters was guided by the objective of achieving an effective balance between detection accuracy and computational efficiency.

During training, the pretrained YOLOv8s weights were initialized and updated iteratively through multiple training epochs using the annotated construction site safety dataset. Input images were resized to a fixed resolution suitable for object detection tasks, allowing the network to learn both global contextual information and fine-grained spatial features necessary for accurate PPE localization.

Throughout the training process, several evaluation metrics were monitored continuously to assess model convergence and detection performance. These metrics included training loss, validation loss, precision, recall, and meanAverage Precision. Monitoring these indicators enabled early detection of overfitting and ensured that the training process progressed toward optimal model 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 presented the methodological framework adopted for developing the proposed PPE detection system using the YOLOv8 deep learning architecture. The dataset preparation process, preprocessing procedures, transfer learning strategy, and training configuration were described in detail to clarify the implementation pipeline of the proposed detection approach.

The next chapter presents the experimental results obtained from training the detection model and discusses its performance using quantitative evaluation metrics derived from the testing dataset.

Chapter 4 :Experimental Results and Discussion

4.1 Introduction

This chapter presents a comprehensive evaluation of the proposed PPE detection system based on the fine-tuned YOLOv8 model. The objective of this evaluation is to assess the effectiveness of the proposed approach in accurately detecting PPE components in construction site environments under realistic conditions.

The experimental analysis is conducted using standard object detection evaluation metrics, including Precision, Recall, and mean Average Precision (mAP). In addition to quantitative evaluation, qualitative analysis is also provided through visual inspection of detection outputs. The results are further interpreted in relation to the characteristics of the dataset and the capabilities of the YOLOv8 architecture.

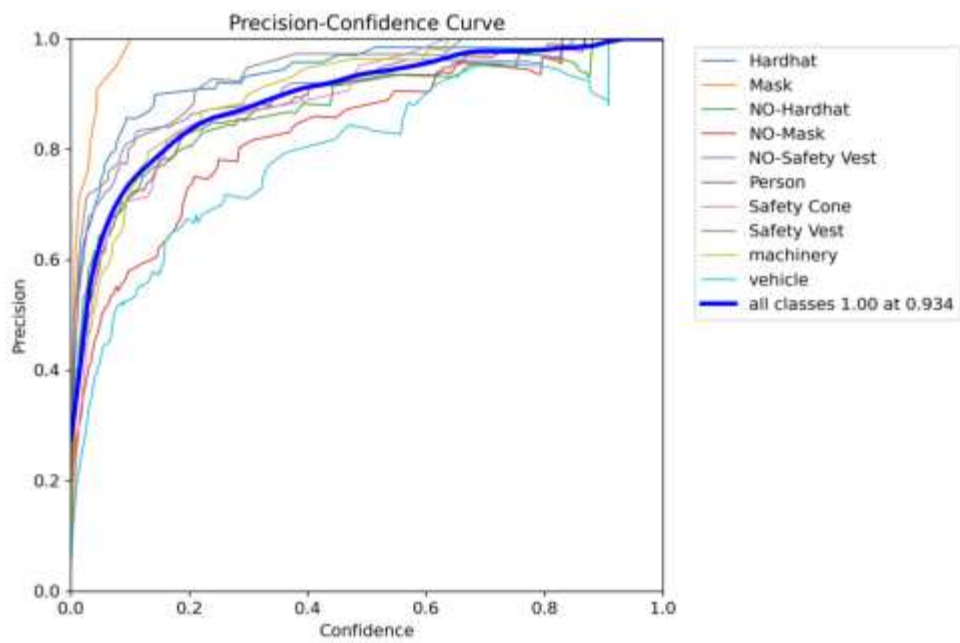
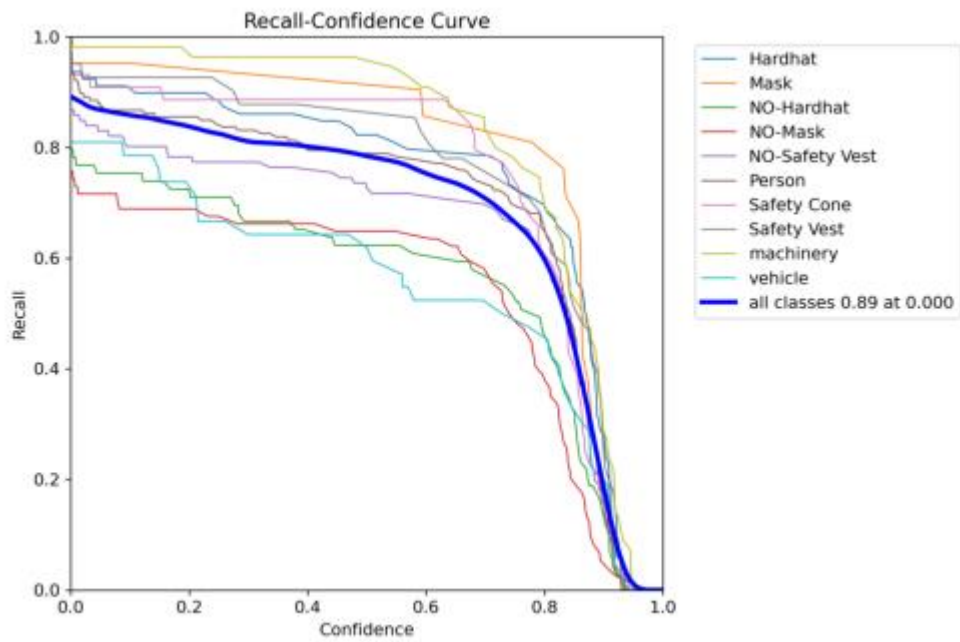
4.2 Detection Performance Results

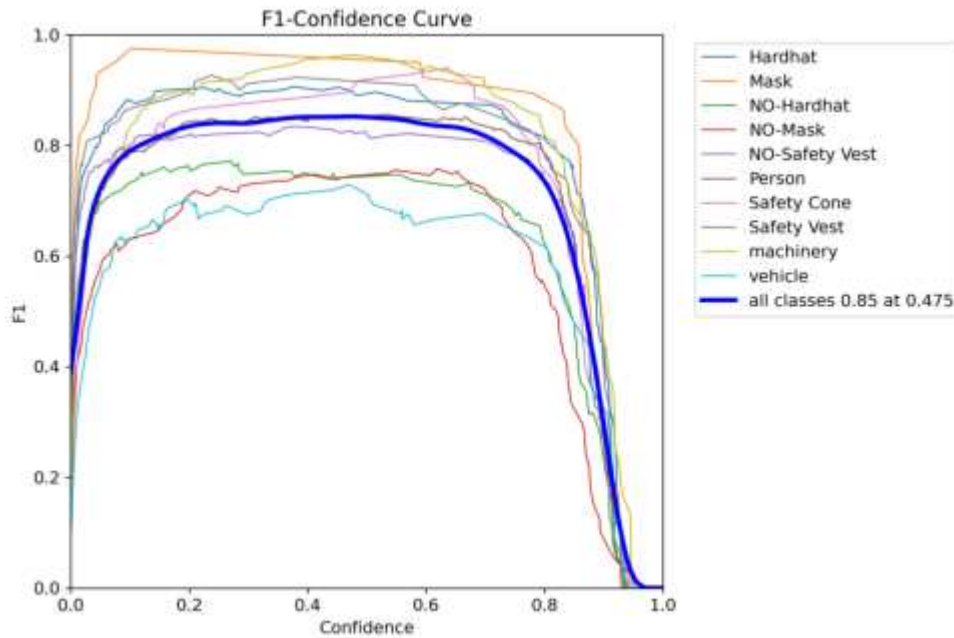
The trained YOLOv8 model showed strong results in identifying PPE-related objects across the validation and testing datasets. The results show that the model achieved a high level of accuracy in identifying multiple object classes, including workers, helmets, and safety vests, even in complex visual environments.

The model achieved a high mean Average Precision at IoU threshold 0.5 (mAP@50), reflecting its ability to correctly detect and localize objects with acceptable overlap accuracy. Additionally, the mAP@50–95 metric indicates that the model maintains consistent performance across stricter localization thresholds, which confirms the robustness of the detection system.

Precision values which are obtained during evaluation show that the maturity of prognosticated findings correspond to correct object groups, indicating a low false positive rate. also, recall values indicate that the model successfully identifies utmost applicable objects within the dataset, suggesting a low false negative rate.

A detailed visualization of the evaluation metrics, including Precision–Recall curves, is presented in Figure 3.





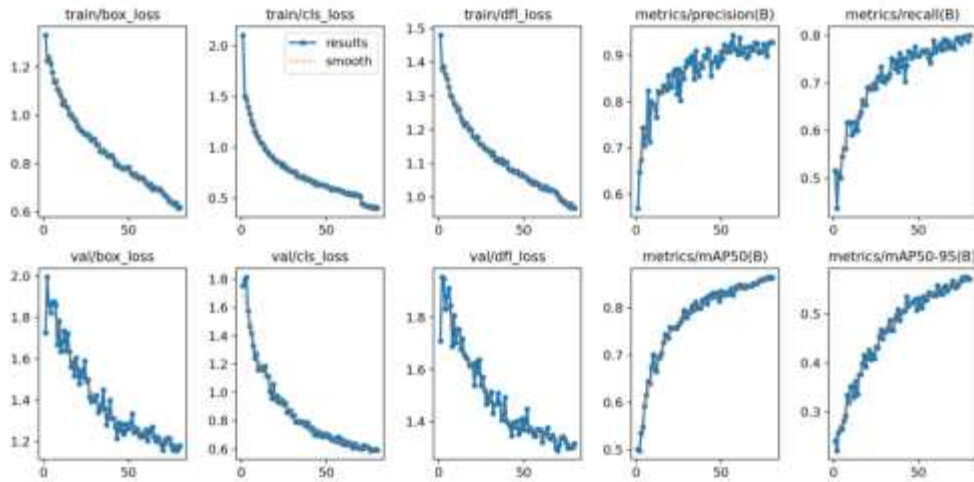
(Figure 3: Precision–Recall curve of the trained YOLOv8 model)

4.3 Training Behavior and Model Convergence

The training process reveals stable Merging behavior, as observed from the training and validation loss curves. In the start, the model showed a rapid decrease in loss values during the early epochs, indicating effective learning of basic visual features. As training progressed, the rate of improvement gradually stabilized, reflecting convergence toward an optimal solution.

The absence of significant divergence between training and validation loss suggests that the model did not suffer from severe overfitting. This can be attributed to the use of data addition ways and the transfer literacy strategy, which bettered conception capability.

The training dynamics are illustrated in Figure 4



(Figure 4: Training and validation loss curves)

4.4 Qualitative Analysis of Detection Results

In addition to numerical evaluation, the qualitative analysis was carried by looking at the detection outputs which are created by the trained model on test images.

The model delivers capability in identifying PPE components under various different conditions which includes crowded places, partial occlusion, and changing lighting conditions.

The system was accurate in locating multiple objects in the same image and hence it maintained an accurate bracket indeed though the objects appear at different scales. still, some limitations were noticed in similar cases which involved severe occlusion or veritably small object sizes, where the discovery delicacy was slightly dropped.

Examples of 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 achieves 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 a deep learning–based approach for detecting personal protective equipment (PPE) in construction site environments using the YOLOv8 object detection model. The main objective of the research was to develop an automated monitoring solution capable of identifying essential safety elements such as safety helmets, reflective vests, and workers in construction scenes in order to support workplace safety monitoring processes.

The system was trained and evaluated using the Construction Site Safety Dataset collected from publicly available sources such as Kaggle and Roboflow.

Fine-tuning techniques were applied to adapt the pretrained YOLOv8 architecture to the PPE detection task, which helped improve detection performance while reducing training complexity. The experimental results demonstrated that the model achieved reliable performance in detecting PPE components under different conditions, confirming the effectiveness of modern object detection techniques in supporting occupational safety applications.

The findings of this exploration indicate that automated PPE discovery systems can reduce reliance on homemade supervision and give a fresh subcaste of safety monitoring in construction surroundings.

similar intelligent results can help safety administrators in relating on-compliance situations more efficiently and may contribute to reducing plant accidents caused by the absence or incorrect use of defensive outfit.

Despite the desirable results that are obtained from this study, there are several options for improvements. In future work may focus on extending approach to operate in real-time environments using live surveillance cameras instead of still image datasets. This step would help the system to be applied more effectively in real life workplace scenarios. In addition, adding the size and diversity of the training dataset by including images collected from different construction surroundings may further ameliorate model robustness and conception performance. unborn exploration may also explore the discovery of fresh defensive outfit similar as gloves, safety thrills, and defensive spectacles to enhance the overall monitoring capability of the system.

Another possible direction for future development involves integrating the detection model with Internet of Things (IoT)–based monitoring platforms to enable automatic warning systems when safety violations occur. Such integration may support proactive safety management strategies and improve decision-making processes in workplace environments.

Overall, the results of this study show that deep learning–based PPE detection systems represent a solution for betterment of construction site safety monitoring and can serve as a base for developing more advanced safety management systems in future applications.

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).