Enhancing Safety in Construction: A Computer Vision Approach for Personal Protective Equipment Detection.

MSc Research Project
Data Analytics

Dhrumil Nanji Sidapara
Student ID: X21241210

School of Computing
National College of Ireland

Supervisor: Sasirekha Palaniswamy

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Dhrumil Nanji Sidapara
Student ID:	X21241210
Programme:	Data Analytics
Year:	2023
Module:	MSc Research Project
Supervisor:	Sasirekha Palaniswamy
Submission Due Date:	14/12/2023
Project Title:	Enhancing Safety in Construction: A Computer Vision Approach for Personal Protective Equipment Detection.
Word Count:	5698
Page Count:	17

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Dhrumil Nanji Sidapara
Date:	14th December 2023

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	✓
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	✓
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	✓

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Enhancing Safety in Construction: A Computer Vision Approach for Personal Protective Equipment Detection.

Dhrumil Nanji Sidapara
x21241210
Programme
National College of Ireland

Abstract

Identifying and reducing potential dangers and consequences in the workplace is essential to maintaining employee security throughout production processes. The use of personal protective equipment (PPE) greatly aids in the decrease of sickness and harm connected to the workplace. Modern Deep-learning techniques are used in this research project to recognize and track PPE usage in real time. The suggested approach seeks to get beyond the drawbacks of current models by addressing issues like dim lighting, unfavorable weather on building sites, and complex image backgrounds. Through enhancing present simulations, the study endeavors to simultaneously ascertain all essential PPE components, such as helmets, gloves, safety glasses or masks, protective clothes, and shoes. In order to maximize the number of images analyzed in the second phase, the study technique will make use of the YOLOv5 single-stage recognition of objects framework. The objective of this tactical decision is to optimize effectiveness and agility while recognizing personal protection equipment (PPE) in situations that occur in the moment. The intended outcomes of this strategy include the development and implementation of a flexible and regulated technique for PPE detection. This expected result, with an emphasis on the quick and precise identification of critical safety equipment, is well-positioned to make a substantial contribution to successfully integrating and adoption of cutting-edge technology in guaranteeing risk at work.

Keywords— Employee safety, Personal Protective Equipment (PPE), Computer Vision Technology, Deep Learning, Object Detection, YOLOv5.

1 Introduction

When it comes to the construction industry, worker safety is of utmost importance since the processes of expansion and advancement are inherently risky. Given the critical role that personal protective equipment (PPE) plays in reducing possible dangers, creative ways to increase its efficacy are required. This endeavor seeks to address this significant void by employing a cutting-edge computer vision technique expressly engineered for the purpose of discerning personal protective equipment. Through the application of cutting-edge technology, particularly in the field of computer vision, this research seeks to transform safety at construction procedures. The study journey's title, sums up its focus on utilizing cutting-edge approaches for the accurate and immediate recognition of critical safety gear. As we go deeper into this investigation, the relationship between tech as well as security in the building sector comes into focus, offering a new perspective in terms of ensuring the safety of people at the vanguard of industrial advancement.

1.1 Research Background

Taking advantage of the confluence of contemporary machine vision and deep learning methodologies, this study sets out on a transformational journey in recognition of the urgent need to enhance the efficacy of PPE to be deployed. The main objective is to implement a novel method that goes above the constraints of current models for actual time identification and tracking of PPE usage. According to estimates from the (ILO), in excess of three million people pass away a year from diseases or deaths related to their jobs.Neale (2013)Building sites have an injury rate that is higher than seventy-one percent, compared to other professions.Waehrer et al. (2007) The research, which stands out for its dedication to innovation, builds on existing models to identify every crucial PPE part in detail. The simultaneous targeting of helmets, gloves, protective clothes, masks or glasses, and shoes improves the overall comprehension of safety procedures. A key component of this study approach is the deliberate use of the Yolov5 one-stage system for recognizing items, which was selected to maximize the effectiveness of visualization in immediate circumstances. Adopting YOLOv5 tactically represents a dedication to processing as many photos as possible in a second, guaranteeing quick and accurate verification of vital safety equipment. The study aims to promote the creation and application of a versatile and regulated method for PPE verification, with results that go beyond simple recognition. This expected outcome, which is supported by an emphasis on quick and precise security gear identification, has the potential to significantly alter the way that modern equipment is woven into industrial risk administration.

1.2 Research Problem Identified

Deep learning and extensive methods of learning have been applied in previous research to identify and track personal protective equipment (PPE). However, the current study is limited in its ability to concurrently identify various forms of safety equipment and take precise images of workers in a variety of weather situations.The aforementioned challenges underscore the necessity for an improved method of tackling these matters.

1.3 Research Question

In what ways the performance of computer vision methods can detect and improve images while tracking workers working in extreme weather conditions and unfavorable lighting situations?

1.4 Objective

The objective of the current investigation is to employ sophisticated machine learning techniques and cutting-edge computer vision technologies in order to enhance the efficacy of personal protective equipment (PPE) in the domain of occupational safety. In particular, the research seeks to address the shortcomings of existing models by creating a revolutionary method that can concurrently detect and track many types of personal protective equipment (PPE) in the moment, comprising protective clothing, footwear, covers or eye protection, mittens, protective gear, and garments. YOLOv5 single-stage object identification framework will be strategically used in order to raise the motion each minute that passes improve efficiency in protective gear recognition, especially in difficult circumstances. Through improving upon current models and offering a consistent and flexible approach to PPE detection, the study aims to enhance employee security in general and may lower incident and compensation rates by improving PPE identification reliability and efficacy.

1.5 Structure of the proposal

- This is a summary of the document's layout.
- Section 2 of this paper summarizes every study that has already been done on the subject.
- In Section 3, I'll talk about visual recognition technology and deep learning techniques.
- The required installation and tools are covered in the same sub-subsection.
- Section 4 of this research covers the Conclusion and Future Work
- Section 5 Acknowledgement

2 Literature Review

The following section explores several studies that clarify how one gains domain competence by understanding and using various research approaches. The following subsections include: 1. Neural Networks for Image Detection; 2. Deep Neural Network Model Comparison; and 3. A Careful Review of the Literature Already Published.

2.1 Detecting Images with Neural Networks

Neural network-based image detection uses sophisticated machine learning methods to locate and categorize items in picture data. Convolutional neural networks (CNNs) in particular, which develop layers of features to accurately recognize and localize objects inside pictures, are extremely effective at image identification tasks. This method is extensively employed in numerous diverse domains, including the examination of medical images, autonomous vehicles, and the recognition of facial attributes.

2.1.1 Safety helmet detection method based on YOLO v4

With an emphasis on reducing health dangers for employees lacking safety helmets, technological advancements have brought up creative ways to improve safety in workplaces. To solve the persistent problem of safety incidents caused by failing to comply with helmet-wearing norms,Benyang et al. (2020) this study introduces a unique helmet identification approach based on the enhanced YOLO v4 (You Only Look Once version 4) algorithm. The suggested approach achieves a more accurate a priori frame dimensional center by putting together a self-constructed dataset made up of movies from actual building sites and using the K-means algorithm for clustering. This allows for the extraction of targeted edge information. 7 The system's training procedure next employs a broad conditioning technique to strengthen the algorithm's flexibility over a range of sensing levels. The test findings demonstrate the effectiveness of the framework with a (mAP) of 92.89 percent of jobs involving the recognition of helmet wearers and a rate of recognition of 15 frames per second. This represents a substantial boost over the baseline YOLO v4, indicating that the current needs for helmet identification have been practically fulfilled. The study also emphasizes how important deep learning is to improving pinpointing targets skills, particularly with regard to YOLO v4. The investigation places its findings within the larger framework of employee security, highlighting how crucial it is to wear helmets in order to avoid disasters. The high cost of current approaches, such as human testing and camera recording, is criticized, highlighting the importance of actual time computerized identification in supporting medical control. The study ends with an overview that restates the benefits of the suggested approach in terms of satisfying immediate detection needs and guaranteeing accurate detection in the intricate and fluid manufacturing setting. In summary, the existing study highlights the importance of the suggested helmet detecting technique for employee protection and sophisticated target identification approaches.

2.1.2 Detection of Personal Protective Equipment (PPE) Compliance on Construction Site Using Computer Vision Based Deep Learning Techniques

The development industry is faced with ongoing difficulties in maintaining employee security in the face of the potential hazards associated with building sites. Workers are still vulnerable to safety threats even after thorough danger assessments and management mechanisms are put in place, which makes wearing personal protective equipment (PPE) essential. Delhi et al. (2020) This paper presents a system based on computer vision and deep learning algorithms to identify failure to comply with the use of PPE in an organized and actual-time manner. In especially the simplest form of the YOLOv3 algorithm is trained with Convolutional Neural Networks (CNN) and machine learning is applied to create an accurate framework for actual time detection of project employees' willingness to wear PPE, especially helmets as well as protective jackets. 7 The algorithm used is trained on a set of 2,509 pictures taken from footage captured at different building locations. On the actual dataset, the algorithm achieves an F1 score of 0.96, with a mean recall and precision rate of 96 percent. The suggested paradigm separates No Helmets and No Jacket situations in addition to classifying conformance as Secure or NOT Secure. The algorithm creates a recorded summary and sounds an alert when it detects failure to comply, which can be integrated in actual time into processes for building web security. The study emphasizes how computer vision-based methods, particularly those utilizing deep learning, may be used to automate hazardous requirements in the building trade. 7 The study highlights a paradigm shift away from human-centric monitoring methods and toward automated, real-time safety monitoring on construction sites. In order to provide a thorough and automatic reaction to safety threats, the conclusion highlights the possible wider uses of the created framework and envisions its integration into a bigger safety monitoring system that may make use of big data and the Internet of Things (IoT).

2.1.3 Real-time Personal Protective Equipment (PPE) Detection Using YOLOv4 and TensorFlow

The COVID-19 pandemic continues to be a worldwide concern that requires creative solutions to slow down the propagation of the virus. In order to deal with interaction between people, a crucial component of the spread of viruses,Protik et al. (2021) this research focuses on the virus's entrance sites via accessible regions such the mouth, nose, and eyes. The researchers stress the need of personal protective equipment (PPE), such as gloves, face shields, and masks, as essential components in stopping the spread of viruses. But they do admit that some people are reluctant to use personal protective equipment (PPE), which emphasizes the necessity of an on-demand surveillance system. The paper presents a detector that was created with the use of the well-known YOLOv4 computer vision model, which excels in actual time object recognition. The writers painstakingly assembled a composite dataset made up of both taken and gathered photos labeled in the YOLO style. With a (mAP) of 79 percent, the study shows the viability of the created sensor. The following categories are specifically examined: face mask, no face mask, face shield, and hand gloves. The detector's excellent performance on a variety of components, such as internet pictures, recorded pictures, and live webcam feeds, demonstrates how its use can be adapted to practical situations. The study offers a thorough assessment of the detector's efficiency, going over parameters for each class that include mAP, average loss, and average accuracy. This discovery is important because it has the possibility of helping improve the community's health by promoting the use of personal protective equipment (PPE) and facilitating close surveillance, which could ultimately slow the global spread of COVID-19.

2.1.4 Computer Vision for Industrial Safety and Productivity

The researchShetye et al. (2023) emphasizes the value of personal protective equipment (PPE) in lowering incidents at work and stresses the necessity of a dependable, ongoing surveillance system. To close this gap, the suggested system uses the YOLOv3 algorithm to identify and classify safety gear, including jackets, helmets, and safety goggles. This 3500 picture collection from factories is specifically designed to highlight the types of safety equipment that are frequently seen in production plants. Because of its effectiveness, speed, and accuracy, the YOLOv3 algorithm is a good option for the actual time identification of items, enabling the system to determine if employees are using the appropriate safety gear. The suggested technique offers a thorough way of improving job security and includes data preparation, algorithm application, and subsequent alarm messages. The subject matter of the research is made clearer by concentrating on popular protection equipment elements and restricting the dataset to photographs. The characteristics of the suggested algorithm are described, namely its capacity to recognize safety equipment, quicker processing using YOLO, and actual time alarm alerts. The processes are depicted in a process graph, starting with sample classification and ending with the training and assessment of the YOLOv3 model its determination of the existence or missing safety gear, and the issuing of alarms. The convolutional neural network design, speed, and efficacy of the YOLOv3 deep learning model are highlighted in its introduction. To sum up, the YOLOv3-based computer vision system that has been suggested offers a reliable way to keep an eye on PPE compliance and improve safety in the industrial industry. Mean Average Precision (mAP) and other performance measures, together with the outcomes, attest to the system's success.

2.1.5 Design and Implementation of Safety Helmet Detection System Based on YOLOv5

The research Guan et al. (2021)that is currently available emphasizes how common deep learning techniques are for solving problems related to target identification and acknowledgment, such as headgear detection. By presenting an improved YOLOv5 model and implementing changes to the input size, anchor box parameters, and gain function, this study adds to the existing body of research. The improved YOLOv5 model meets the specifications for actual time helmet identification with notable improvements in precision, attaining a noteworthy 90 percent, and a phenomenal recognition speed of 37.8fps. This demonstrates how well deep learning models—YOLOv5, in particular—can handle safety issues in construction-related settings. Several important metrics, including accuracy, recall rate, standard deviation of mean, and frequency mean, are evaluated in the YOLOv5 model assessment process. The model is stable and reliable; its accuracy and recall are higher than 90 percent. The experimental set's (mAP) of 90.1 percent highlights the effectiveness of the YOLOv5 algorithm in helmet use detection. The research offers significant perspectives on the pragmatic implementation of the model, demonstrating 37 frames per second identification on a GTX 2070 graphics card, satisfying the requirements of instantaneous identification.

2.2 Deep neural network comparability of models

The title "Deep Neural Network Comparability of Models" implies that the contrast between various deep neural network models would be the main emphasis. This entails evaluating their performance, appropriateness for a certain purpose, and interoperability. The analysis is expected to cover topics including model design, training approaches, and general efficacy in managing challenging tasks, offering insights into the advantages and disadvantages of different deep learning models.

2.2.1 A Smart System for Personal Protective Equipment Detection in Industrial Environments Based on Deep Learning

Actual object identification systems using video streaming analysis are highly valued in today's industrial environment. This approach is widely used in a variety of fields, including safety management and surveillance of security. To guarantee employees' safety in manufacturing facilities, exact compliance to Personal Protective Equipment (PPE) guidelines is essential. But workers frequently forget to wear essential PPE, including helmets, especially when they're working indoors. Gallo et al. (2021) Gallo and colleagues (2021) have introduced a developed intelligent system designed to identify personal protective equipment (PPE) in real-time through the analysis of live video, utilizing sophisticated machine learning models. The primary focus of their investigation lies in the detection of helmets, aiming to address this issue effectively and mitigate the potential hazards associated with accidents. The authors use various iterations of the YOLOv4 network that have been refined with a public PPE dataset in order to assess the precision, latency, and overall PPE detection performance. The design that has been suggested fits within the Industry 4.0 framework and supports apps that require immediate and local manufacturing, which is in line with the expanding trend of cloud computing. The study highlights how this customized strategy can lead to lower reliance on services hosted in the cloud and more dependability. The dataset that was used to fine-tune the model includes a variety of industrial contexts, which enhances item detection resilience. The authors carry out an extensive test campaign, taking into account factors like identification speed, classification ability, and training time. They pay particular attention to the YOLOv4-Tiny version, proving that it can be used for real-time PPE detection in computer systems.

2.2.2 Safety Helmet Detection Using Deep Learning: Implementation and Comparative Study Using YOLOv5, YOLOv6, and YOLOv7

The use of deep learning (DL) models for safety helmet detection has attracted a lot of interest, which is indicative of the increased significance placed on protecting worker safety, especially in high-risk settings like building sites.Yung et al. (2022) Advances in computational imaging techniques, particularly those that use deep learning (DL) for robust object recognition, have contributed to this increase of interest. Deep learning (DL) and its variant deep cognitive networks are especially well-suited for jobs involving object recognition since they have demonstrated efficacy in producing complex characteristics from unprocessed input. In this regard, the paper examines how well different DL algorithms that are used for headgear identification work. In particular, the study explores the YOLOv5s (small) model, YOLOv6s (small) model, and YOLOv7 model, assessing as well as teaching the models to determine their effectiveness. The challenge regarding helmet detection is framed in the context of artificial intelligence and its related fields, with a focus on the advancement of health surveillance systems through the use of deep learning (DL), machine learning, neural networks, and computer vision. Since failure to follow personal protective equipment (PPE) regulations is a common cause of injuries on construction sites, this work employs safety helmet identification as a sample instance in the larger area of object detection. A thorough analysis of the methods and results related to the use of DL models for safety helmet detection is made possible by the literature study, which also provides insights into the development of these advancements and their effects for maintaining worker safety.

2.2.3 Generic compliance of industrial PPE by using deep learning techniques

Industries have serious concerns about the execution of workplace safety rules that are effective, especially when it comes to using Personal Protective Equipment (PPE) appropriately. Many injuries and large sums of money may arise from safety managers' failure to quickly detect and address PPE abuse. Vukicevic et al. (2022) This study suggests a unique method utilizing artificial intelligence (AI) for the binary categorization of conformity to PPE in response to the increasing need for cutting-edge solutions to support workplace safety. The method provides a versatile solution that can be tailored to meet the demands of different industries by accommodating a wide variety of PPE kinds across multiple body areas. The research thoroughly assesses eighteen different PPE categories that protect five biological body regions and/or activities. The work presents outstanding results using six various image categorization architectures, namely MobileNetV2, Dense-Net, and ResNet, and the HigherHRNet pose estimator for PPE area identification. Because it strikes the best possible equilibrium of processing effectiveness and fastness, MobileNetV2 is the preferred option. The suggested method stands out for being flexible, making it simple to customize for various body parts and PPE kinds. Unlike earlier research, this technique is resilient and adaptable, allowing it to be used in a variety of sectors. The study addresses the difficulties caused by the size and variety of employees in business environments and suggests its deployment at self-checkpoints and hazardous industries. This study makes a substantial contribution to the digitization of PPE compliance and offers a workable way to improve workplace safety in the face of changing industry regulations.

2.2.4 Real-Time Personal Protective Equipment Compliance Detection Based on Deep Learning Algorithm

The building business is widely recognized as one of the world's most hazardous industries. Its workers are constantly at danger of accidents, injuries, and fatalities. Appropriate control over the use of personal protective equipment (PPE) is necessary to mitigate these dangers. In order to check the building workers are adhering to safety requirements and lower the risk of workplace tragedies caused by inappropriate PPE use.Lo et al. (2023) this study uses deep learning in a PPE inspection model. The approach uses the YOLOv3, YOLOv4, and YOLOv7 algorithms to recognize employee helmets and high-visibility jackets in actual time from pictures or videos. This work presents a new PPE dataset with 11,000 photos and 88,725 labels, which will be an important tool for future study. Throughout evaluation, the model performs at 25 frames per second (FPS) and achieves a noteworthy 97 percent mean average accuracy (mAP). The results demonstrate the effectiveness of the suggested method's identification and calculating capacities and point to its potential for real-time PPE monitoring at building sites. In light of the construction industry's ongoing elevated incidences of injuries at work, integrating deep neural networks and artificial intelligence technology presents a viable way to improve safety procedures by automating, accurately, and promptly checking PPE compliance. The work advances the conversation on the critical nexus of cutting-edge technology and safety in the building by creating a comprehensive PPE dataset in addition to developing an improved YOLOv7-based detection algorithm.

2.2.5 Safety Helmet Detection Based on YOLOv5

The development of recognition of items methods has brought in a new age of precision and rapidity in a variety of image identification projects, especially in the YOLO (You Only Look Once) series of methods. This paperZhou et al. (2021) explores the topic of security at work, emphasizing the critical function that security helmets have in safeguarding employees. Notwithstanding its importance, workers frequently fail to put on helmets due to failures in safety knowledge. The paper suggests utilizing the YOLOv5 method to create a computerized helmet-based surveillance system, capitalizing on the progress made in detecting objects. Training and testing were conducted using datasets with annotations including 6045 photos, with an emphasis on assessing four YOLOv5 models (s, m, l, x) with varying parameters. According to research findings, the YOLOv5s can identify objects at roughly a speed of 110 FPS, which satisfies real-time detection requirements. The following study is expected to use a wider set of data that includes actual-life incidents in order to improve the model's capacity for extension. All things considered, the safety helmet detection approach based on YOLOv5 shows promise as an actual time and reliable work environment security surveillance system.

2.3 A Comprehensive Overview of the Literature

The examined research offers a thorough summary of the most recent developments in protective helmet identification techniques, highlighting the crucial role those advances plays in improving worker safety. Numerous investigations have looked into the use of deep learning algorithms—specifically, those from the YOLO series—for the detection and tracking of safety helmets in industrial environments. Combining machine vision with advanced learning methods has shown promise in reducing safety concerns related to non-adherence to personal protective equipment (PPE) regulations.

3 Proposed Methodology & Specifications

Knowledge Discovery in Databases (KDD) approach is an effective instrument for analyzing gathered data. Its efficacy originates from its emphasis on real-world application as opposed to complex system design. When it comes to analyzing information, this method is quite helpful, especially when it comes to finding and predicting. The study that is being suggested makes use of neural networks and computer vision methods to create a dynamic system that is designed to continuously identify and monitor personal safety gear (PPE). The primary goal is to improve the recognition and monitoring of various personal protective equipment (PPE) components, such as helmets, gloves, safety glasses or masks, protective clothing, and shoes. The study uses the YOLOv5 recognition of objects methodology in order to do this.

3.1 Proposed Methodology

1 - Data Selection: The study has assembled a large picture collection of employees using a range of Personal Protective Equipment (PPE) in a variety of work environments. The "Personal Protective Equipment Dataset" that can be found on universe.roboflow.com Equipment (2022)will be used to manage this dataset, which includes annotations for subclasses such as goggles, helmets, masks, suits, shoes, no-shoes, no-helmet, no-mask, gloves, and no-glove. This set of pictures, which comprises 11,978 images, offers a wealth of information for the PPE identification technique's testing, validation, and training. Workers may be seen in the photos donning a variety of PPE, such as protective suits, helmets, goggles, masks, gloves, and safety shoes. To enable detailed examinations, the collection also includes photos of workers sans particular PPE components. The cornerstone resource for building, training, and assessing the models in the research endeavor is this painstakingly collected data.

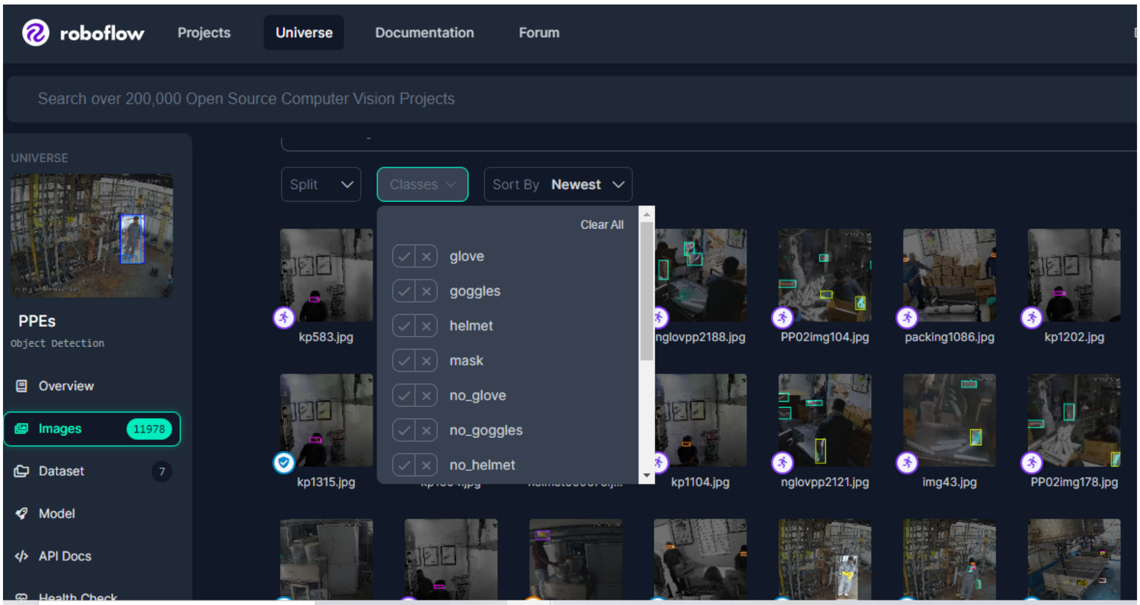

Figure 1: Dataset

2 - Initial assessment Utilizing Deep Learning Methodology: Google Colaboratory is the software system of choice used in this research to run several mathematical models using deep learning. The 16GB memory chip of the Tesla T4 graphics processing unit (GPU) is set up for virtual connection and will be used continuously during every test. 12GB of Bandwidth is provided by Google Colaboratory, enabling faster neural network development. This research uses the TensorFlow software package, which comes pre-installed in Google Colaboratory, to generate graphical representations and confusion matrices. The average mean accuracy is the main statistic used to evaluate the effectiveness of the deep learning models (mAP). This measurement instrument, which is often used in PPE detection investigations, shows patterns in recall and accuracy across several simulated training scenariosYung et al. (2022). Criteria for assessment that may be expressed quantitatively are used, including precision rate (Pr), recall rate (Re), and F1-score (F1). The evaluation of True Positive (TP), False Positive (FP), and False Negative (FN) samples is required for these indicators to be useful. When an object is correctly predicted to be an effective sample, it is indicated as a True Positive (TP). False Positive (FP) refers to the detection of

$$Pr = \frac{TP}{TP+FP},$$

$$Re = \frac{TP}{TP+FN},$$

$$F1 = \frac{2 \times Pr \times Re}{Pr + Re}.$$

Figure 2: Formula

an erroneous item in an accurate sample. Conversely, False Negative (FN) describes situations in which a legitimate dispute is inadvertently classified as an unwanted sample.

The following are the techniques for calculating mean average precision (mAP) and average precision (AP):

$$AP = \int_0^1 Pr(Re)\ dRe,$$

Figure 3: Average precision

$$mAP = \frac{1}{n}\sum_{k=1}^{k=n} AP_k,$$

Figure 4: Mean Average Precision

3.2 Design Specification

This carefully chosen dataset provides the basis for building, training, and evaluating models. The first evaluation makes use of TensorFlow for model construction, Google Colaboratory, and a Tesla T4 GPU with 16GB of RAM. The main metrics used to evaluate a model are Mean Average Precision (mAP), F1-score (F1), recall rate (Re), and precision rate (Pr). The suggested architecture focuses on smooth integration and model adaptability. It entails cloning the YOLOv5 repository, using the ROBOFLOW API KEY for expedited dataset downloads, and employing YOLO for efficient data processing. An important phase in YOLOv5's iterative development is being marked by the continuous discussion about its functionality and possible enhancements.

The process of my suggested design is shown in Figure 5 below.

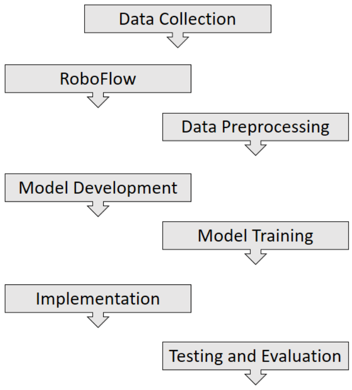

Figure 5: Design of Detecting PPE Kits using Images.

3.3 Proposed Implementation

The first step toward creating the YOLOv5 architecture and training setup from GitHub is to carefully copy the YOLOv5 repository Jocher (2020)along with all of its necessary preconditions. Setting the ROBOFLOW API KEY is a crucial parameter that simplifies the dataset download procedure. This key makes it easier to access the PPEs-8 dataset using the Roboflow API. The collected data is organized under the "dataset" folder. The responsibility for handling the complexities of gathering information lies with YOLO. Following a rigorous development process, an algorithm is used to training and validate sets of pictures. Following that, the outcomes are carefully produced and thoroughly discussed.

The YOLOv5 repository acts as a dynamic platform for the YOLO model's progress in addition to being a fundamental resource. This repository uses GitHub's collaborative capabilities to guarantee continuous updates and enhancements. The interaction between dataset gathering, the ROBOFLOW API KEY, and YOLO's skilled data processing highlights the need of smooth integration in the YOLOv5 configuration. The model's versatility is highlighted by its use on various picture sets, with an emphasis on effectiveness. The important conclusions drawn from extensive training and validation add to the continuing discussion about the functionality and possible improvements of YOLOv5, signifying a pivotal point in the iterative development and improvement of this state-of-the-art deep learning model.

3.4 Setup and Configuration

The system's preliminary setup listed below will be used to construct the model:

Device: Acer Aspire E15

RAM: 8:00 GB

Operating System: Windows 10 64-bit

CPU: Intel(R) Core(TM) i5-7200U CPU @ 2.50GHz 2.70 GHz

GPU: NVIDIA GEFORCE 940MX

3.5 Evaluation and Results Analysis

3.5.1 Evaluation

The sequential method of fine-tuning hyperparameters aims to improve deep learning model performance measures such as accuracy, precision, and recall. These hyperparameters in the context of Ultralytics YOLO include things like the learning rate and architectural details, such as the quantity of neurons or the kinds of activation functions used.

Learning Rate lr0: Chooses the step size for every cycle as it approaches the loss function's minimum.

Batch Size batch: Number of images processed simultaneously in a forward pass.

Number of Epochs epochs: An epoch is one complete forward and backward pass of all the training examples. In my I have kept for 5 epochs.

Here I have used the model.tune() method to utilize the Tuner class for hyperparameter tuning of YOLOv5n on PPE-8 for 5 epochs with an AdamW optimizer and skipping plotting, checkpointing and validation other than on final epoch for faster Tuning.

```
lr0: 0.01001
lrf: 0.00919
momentum: 0.9209
weight_decay: 0.0005
warmup_epochs: 2.72434
warmup_momentum: 0.79397
box: 7.37398
cls: 0.50774
dfl: 1.51006
hsv_h: 0.01526
hsv_s: 0.69304
hsv_v: 0.38952
degrees: 0.0
translate: 0.09797
scale: 0.48178
shear: 0.0
perspective: 0.0
flipud: 0.0
fliplr: 0.50675
mosaic: 0.9757
mixup: 0.0
copy_paste: 0.0
```

Figure 6: Hyperparameters and Optimizer Selection

The optimizer employs a stochastic gradient descent (SGD) optimizer with momentum and weight decay coefficients set at 0.9209 and 0.0005, respectively.

3.5.2 Results Analysis

1. Model Evaluation Indicators- I used the PPE-8 dataset from Roboflow to simulate YOLOv5s, and I found that it took about 5.282 hours to train 50 epochs. This research measures the percentage of overlap between the predicted and real bounding boxes using an Intersection over Union (IoU) criterion with a score of 0.5. If the score is more than 0.5, the projection is judged good and yields a true output; if not, it is regarded as false. As seen in Figure 7, the YOLOv5s algorithm reached its maximum mean Average Precision (mAP) of 31.1 percent at epoch 20 during testing. Interestingly, the mAP did not change at epochs 37 or 49, suggesting stability and obviating the no need for more epochs. The mAP values for epochs 37 and 49 are shown in Figures 8 and 9. This result implies that the model's performance reaches a plateau after a specific amount of time and that further training epochs result in very little progress. The model's effectiveness in bounding box predictions is enhanced by the IoU threshold of 0.5, which functions as a crucial criterion for differentiating meaningful results.

```
Epoch    GPU_mem   box_loss   obj_loss   cls_loss  Instances       Size
20/49      1.83G    0.03665    0.01343   0.001827         23       416:
           Class     Images  Instances          P          R      mAP50
             all       3570       7710      0.665      0.458      0.312
```

Figure 7: mAP of YOLOv5s at epoch number 20.

The research's main modelling assessment specifications are harmonic mean(F1-Score), accuracy, recall rate, and average precision mean. Better recall rates and accuracy are positively correlated with improved identification of Personal Protective Equipment (PPE), however not positively. The means of standard deviation and harmonics are quantitative measures that include recall and accuracy. The effectiveness of PPE wear recognition increases with the values of these indicators. These metrics provide a thorough evaluation of the model's performance by taking precision and recall into account. This highlights the significance of striking a balance between accuracy and recall for the best PPE identification efficacy.

Epoch	GPU_mem	box_loss	obj_loss	cls_loss	Instances	Size
37/49	1.83G	0.033	0.0124	0.001091	43	416:
	Class	Images	Instances	P	R	mAP50
	all	3570	7710	0.68	0.435	0.304

Figure 8: mAP of YOLOv5s at epoch number 37.

Epoch	GPU_mem	box_loss	obj_loss	cls_loss	Instances	Size
49/49	1.83G	0.02991	0.01172	0.0006721	36	416:
	Class	Images	Instances	P	R	mAP50
	all	3570	7710	0.677	0.43	0.304

Figure 9: mAP of YOLOv5s at epoch number 49.

2. Evaluation of the Experimental Results- After 50 training epochs, the YOLOv5 model is evaluated for PPE wear detection; the results are shown in Figure 10.

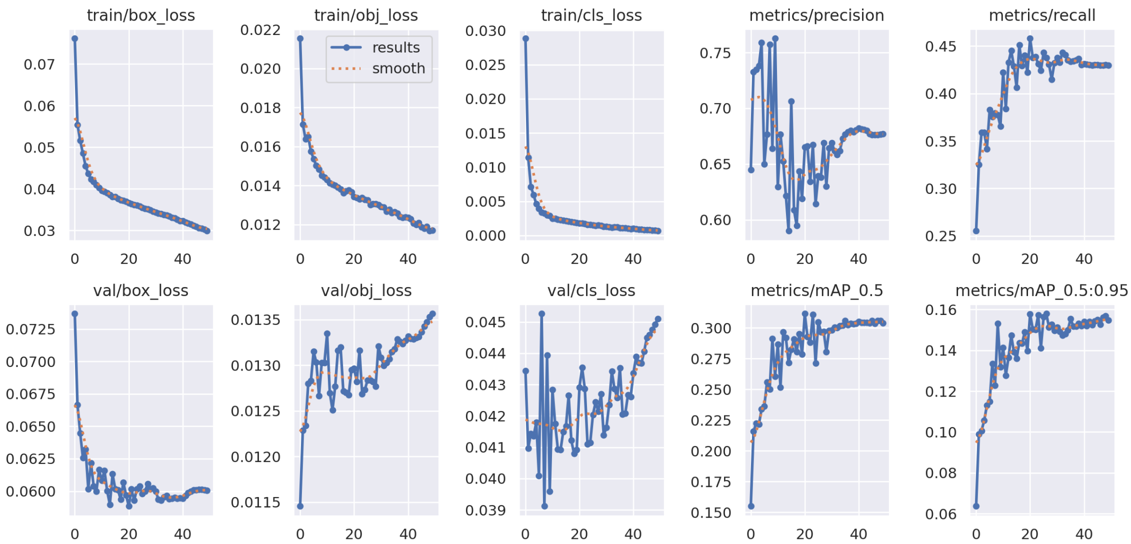

Figure 10: Findings from the model's analysis.

Figure 11 shows the evaluation outcomes of the YOLOv5s algorithm for the detection of personal protective equipment (PPE). After 50 epochs, the framework achieves an average state, showing steady gains in recall and accuracy over the course of training. After completion, the recall stays close to 43 percent and the precision stays over 31 percent. Both the harmonic(F1-Score) mean and the average accuracy means continuously stay at high levels. As seen in Figure 12, the mean Average Precision (mAP) index offers an unbiased assessment, demonstrating the model's efficacy in recognizing different PPE categories, such as goggles, helmet, mask, no-suit, no-glove, no-goggles, no-helmet, no-mask, no-shoes, shoes, suit, and overall mAP. The mean Average Precision (mAP) on the assessment set after the training phase was a remarkable 31.1 percent. This noteworthy result indicates that the YOLOv5 algorithm performs satisfactorily in the field of Personal Protective Equipment (PPE) detection.

The YOLOv5s model's confusion matrix and detection performance matrix are shown in Figures 12 and 13. The effectiveness of the model in categorizing 12 distinct object classes is summarized in the confusion matrix, and patterns of the F1-score, accuracy rate, and recall rate are shown in the object recognition performance matrix. In particular, Figure 12 shows the confusion matrix for identifying twelve different classes. Examining this matrix shows that the algorithm can recognize

Class	Images	Instances	P	R	mAP50	mAP50-95:
all	3570	7710	0.665	0.458	0.311	0.158
glove	3570	984	0.712	0.937	0.91	0.475
goggles	3570	1192	0.81	0.601	0.643	0.384
helmet	3570	283	1	0	0	0
mask	3570	253	1	0	0	0
no-suit	3570	15	0.016	0.8	0.0306	0.0261
no_glove	3570	1548	0.553	0.731	0.6	0.277
no_goggles	3570	1384	0.702	0.699	0.702	0.338
no_helmet	3570	229	1	0	0	0
no_mask	3570	640	1	0	0	0
no_shoes	3570	525	0.157	0.169	0.104	0.0168
shoes	3570	639	0.992	0.556	0.599	0.231
suit	3570	18	0.0404	1	0.149	0.144

Figure 11: Model Summary

safety kits with an overall accuracy score of 31 percent. It does exceptionally well in identifying gloves, goggles, no-suit, no-glove, no-goggles, no-shoes, shoes, and suits. With a false prediction score ranging from 0.0 to 0.24, the model performs average.

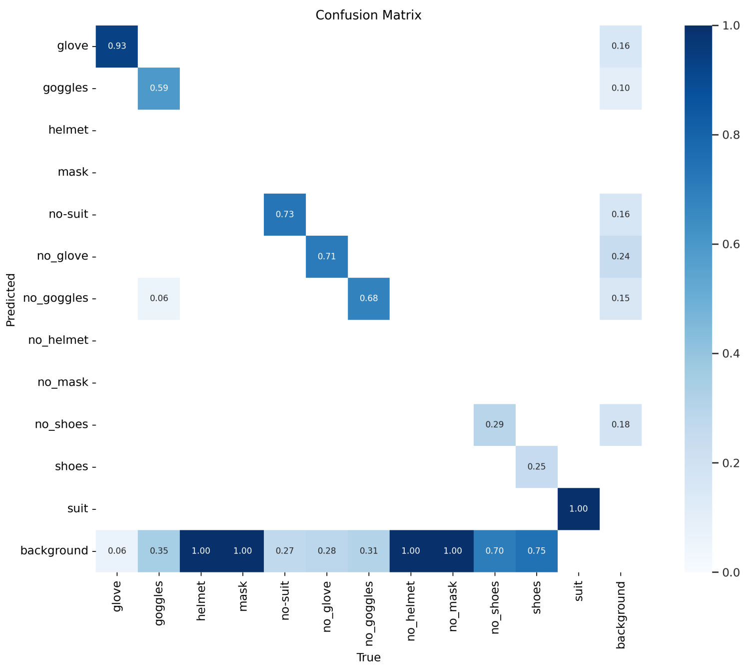

Figure 12: Confusion matrix of the YOLOv5s model

It is clear from looking at Figure 13's F1-score versus probability graph that the model that was learned achieves an F1-score of 0.32 for all object classes at a certain confidence level (blue curve). Interestingly, the accuracy rate and recall rate are balanced by the F1-score. As the learned YOLOv5s model's confidence level rises for particular classes, Figure 13's plotted graphs of precision vs recall, precision versus confidence, and recall versus confidence show an inverse connection between the accuracy rate and recall rate. This finding emphasizes the inverse relationship between recall rate and accuracy rate, providing detailed information about the model's degree of confidence in various object classes.

Examples of findings from tests are shown in Figure 14, giving an idea of the results attained after YOLOv5's models were trained. Figures 15 and 16, which provide effectiveness samples for the model, provide more light on the evaluation results. Interestingly, the YOLOv5s models performed less well than ideal in low light. The models' inability to identify safety helmet, mask, no-helmet, and no-mask

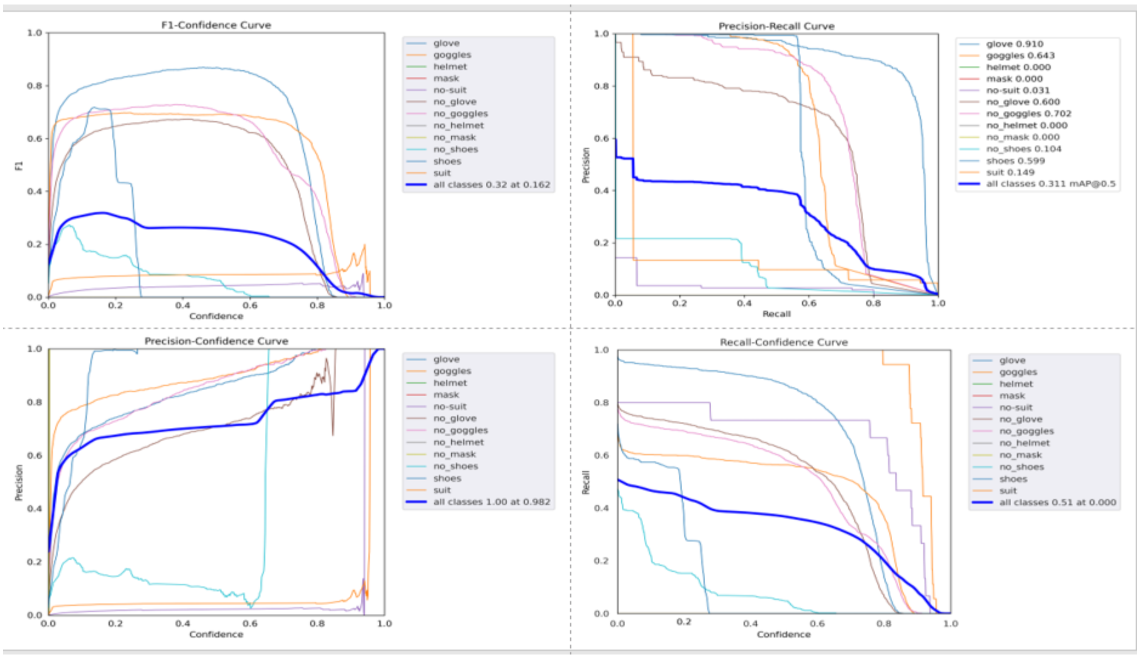

Figure 13: Detection performance metrics of the YOLOv5s model

Figure 14: Sample testing images from the dataset

situations under safety Personal Protective Equipment (PPE) is one of the noted flaws. The reasons for these detecting difficulties include things like overall image quality and backdrop quality. Although the accuracy as a whole peaked at 31 percent, looking at individual scores shows that certain areas of achievement were excellent. The subtle challenges in identifying the lack of a safety mask and helmet emphasize the influence of backdrop and clarity on the model's overall performance.

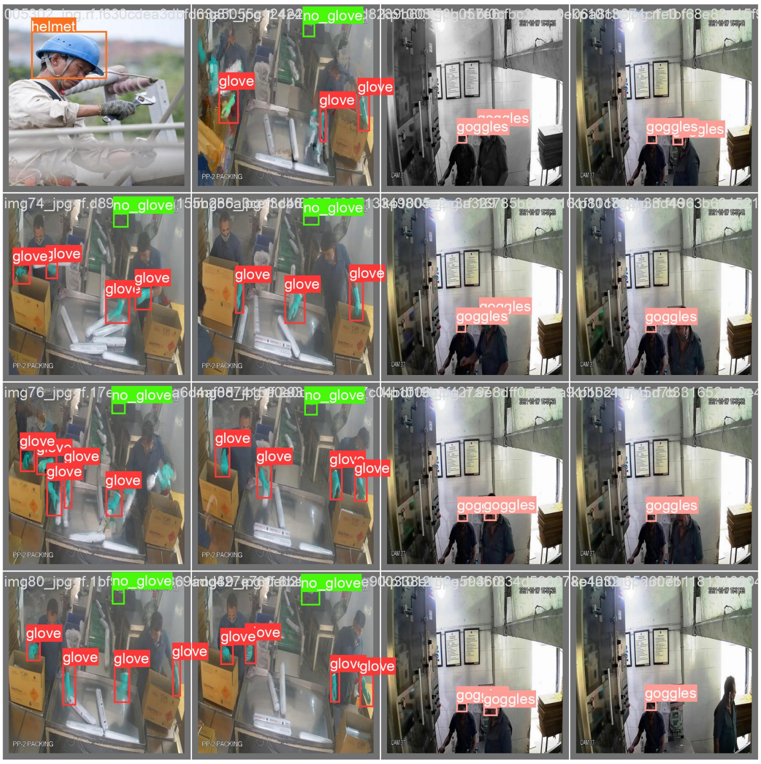

Figure 15: Testing results with labels of YOLOv5s model

4 Discussion

The explanation of the research's assumptions is covered in detail in this section. A structure with 157 layers and 70,42,489 parameters is revealed in the model summary. It makes predictions for 12 different classes: goggles, helmet, mask, no-suit, no-glove, no-goggles, no-helmet, no-mask, no-shoes, shoes, and suit. The average precision (mAP) for the entire sample is 31.1% . When looking at individual class mean average points (mAP), the glove class scores the most (91%), followed by no-goggle, goggle, no-suits, no-shoes, and shoes, which range from 10–70%.

However, the model has shortcomings, particularly in terms of recognizing scenarios in which a mask, helmet, or neither is present. The complexity of the dataset—which includes challenging backgrounds and image characteristics—is the reason for this constraint. The model is unable to analyze photographs effectively due of their complexity; Rather, human involvement is necessary for accurate assessments. In circumstances like accident reviews, relying just on the model might not be sufficient. It gets essential for independent verification to ensure a comprehensive analysis. The accuracy of the model is significantly impacted by the quantity and complexity of the dataset, underscoring the necessity in some situations for human validation.

5 Conclusion

In summary, this research marks a significant breakthrough in the implementation of modern deep learning techniques, namely the YOLOv5 single-stage recognition framework, to enhance the tracking and real-time identification of personal protective equipment (PPE) in occupational environments. The goal of this investigation is to improve upon the weaknesses of existing models by investigating problems

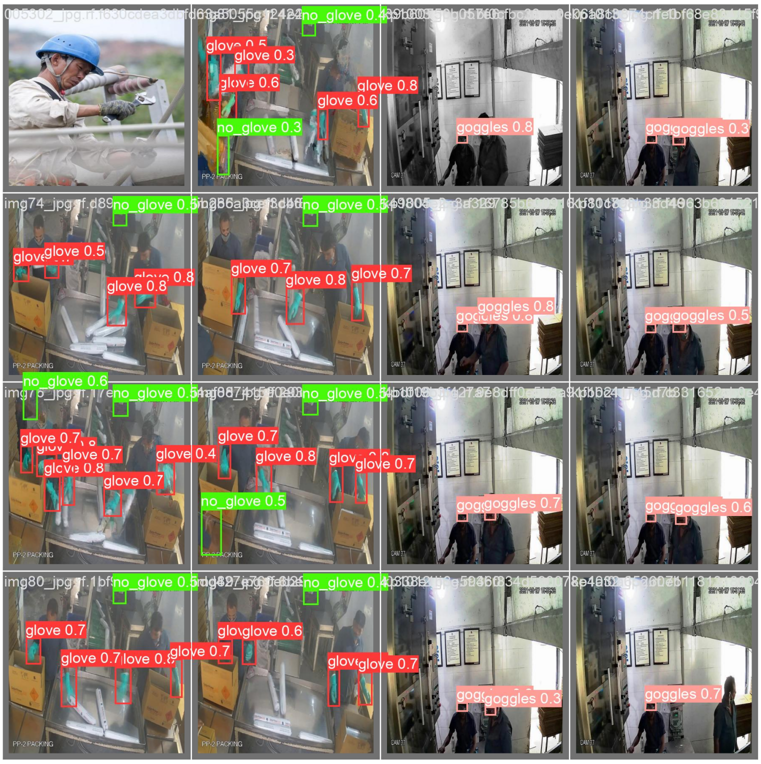

Figure 16: Testing results with prediction on YOLOv5s model

such as dim lighting, adverse circumstances on construction sites, and intricate image backgrounds. After assessment, the YOLOv5 algorithm displays 14 Figure 16: Predictive testing outcomes on the ordinal performance of the YOLOv5 framework, with a mean Average Precision (mAP) of 31.1%. This parameter is crucial for assessing the model's ability to distinguish between the expected and real limitations, which is necessary for accurate PPE detection. Despite this accomplishment, the study points out a number of flaws, namely in the identification of scenarios including helmets, masks, no-helmets, and no-masks. The dataset's complications, that involve minute details like picture quality, background intricacy, and overall overall difficulty, are blamed for these challenges. The efficiency of the model remains consistent over a particular amount of simulated epochs, highlighting how important it is to understand the model's overload limit to prevent losing computational power. The findings demonstrate how intricately safety and innovation interact in a construction industry.

The project's goal is to transform safety measures by utilizing state-of-the-art image recognition technology, providing a fresh outlook on managing risks. We talk about the iterative evolution of the YOLOv5 model and the current debate over its usefulness and possible improvements. The project intends to make a contribution to the field of PPE detection as well as to the wider use and incorporation of cutting-edge technology to ensure workplace safety.

5.1 Future Work

Further studies in this field may focus on enhancing and expanding the YOLOv5 model's capacity for better PPE detection. Improving the model's ability to manage certain challenges, such as low light levels and complex backgrounds, might significantly boost its usefulness. Using state-of-the-art computer vision methods and exploring various neural network architectures may assist to further address the stated limits. More research may be done to create a larger and more diverse dataset that encompasses a wider range of working conditions and scenarios in order to train the model. This would ensure that the computer model maintains to perform effectively in a range of everyday circumstances and deliver a more dependable and ubiquitous efficiency.

It could also be crucial to address any moral issues about data security and privacy and find ways to make the model's findings easier to understand in order to ensure widespread adoption. By focusing on these components repeatedly, more study may be done to develop more complex, reliable, and ethically sound PPE detection algorithms for improved employee security.

6 Acknowledgement

I would like to thank Dr. Sasirekha Palaniswamy for her inspiration and guidance in helping me effectively complete my research project. Dr. Palaniswamy helped me choose a well-organized study and guided me through each phase of the way to make sure I didn't make any mistakes. Dr. Palaniswamy urged tenacity and stressed the value of preserving mental health while overcoming obstacles throughout trying times. His constant encouragement and coaching have been crucial to my research project's fulfillment, and I am deeply appreciative of the direction she has given me during this course of study.

References

Benyang, D., Xiaochun, L. & Miao, Y. (2020), Safety helmet detection method based on yolo v4, *in* '2020 16th International Conference on Computational Intelligence and Security (CIS)', pp. 155–158.

Delhi, V. S., Sankarlal, R. & Thomas, A. (2020), 'Detection of personal protective equipment (ppe) compliance on construction site using computer vision based deep learning techniques', *Frontiers in Built Environment* **6**, 540603.

Equipment, P. P. (2022), 'Ppes dataset', `https://universe.roboflow.com/personal-protective-equipment/ppes-kaxsi`. visited on 2023-12-01.
URL: *https://universe.roboflow.com/personal-protective-equipment/ppes-kaxsi*

Gallo, G., Di Rienzo, F., Ducange, P., Ferrari, V., Tognetti, A. & Vallati, C. (2021), A smart system for personal protective equipment detection in industrial environments based on deep learning, *in* '2021 IEEE International Conference on Smart Computing (SMARTCOMP)', pp. 222–227.

Guan, Y., Li, W., Hu, T. & Hou, Q. (2021), Design and implementation of safety helmet detection system based on yolov5, *in* '2021 2nd Asia Conference on Computers and Communications (ACCC)', pp. 69–73.

Jocher, G. (2020), 'YOLOv5 by Ultralytics'.
URL: *https://github.com/ultralytics/yolov5*

Lo, J.-H., Lin, L.-K. & Hung, C.-C. (2023), 'Real-time personal protective equipment compliance detection based on deep learning algorithm', *Sustainability* **15**(1).
URL: *https://www.mdpi.com/2071-1050/15/1/391*

Neale, R. (2013), 'Ten factors to improve occupational safety and health in construction projects', *African Newsletter on Occupational Health and Safety* **23**(3), 52–54.

Protik, A. A., Rafi, A. H. & Siddique, S. (2021), Real-time personal protective equipment (ppe) detection using yolov4 and tensorflow, *in* '2021 IEEE Region 10 Symposium (TENSYMP)', pp. 1–6.

Shetye, S., Shetty, S., Shinde, S., Madhu, C. & Mathur, A. (2023), Computer vision for industrial safety and productivity, *in* '2023 International Conference on Communication System, Computing and IT Applications (CSCITA)', pp. 117–120.

Vukicevic, A. M., Djapan, M., Isailovic, V., Milasinovic, D., Savkovic, M. & Milosevic, P. (2022), 'Generic compliance of industrial ppe by using deep learning techniques', *Safety Science* **148**, 105646.
URL: *https://www.sciencedirect.com/science/article/pii/S0925753521004860*

Waehrer, G. M., Dong, X. S., Miller, T., Haile, E. & Men, Y. (2007), 'Costs of occupational injuries in construction in the united states', *Accident Analysis Prevention* **39**(6), 1258–1266.
URL: *https://www.sciencedirect.com/science/article/pii/S0001457507000589*

Yung, N. D. T., Wong, W. K., Juwono, F. H. & Sim, Z. A. (2022), Safety helmet detection using deep learning: Implementation and comparative study using yolov5, yolov6, and yolov7, *in* '2022 International Conference on Green Energy, Computing and Sustainable Technology (GECOST)', pp. 164–170.

Zhou, F., Zhao, H. & Nie, Z. (2021), Safety helmet detection based on yolov5, *in* '2021 IEEE International Conference on Power Electronics, Computer Applications (ICPECA)', pp. 6–11.