Comparison of Precision and Accuracy of Image Processing Algorithm to monitor Safety Violations at a construction site

By

Jamal Farooqi

(Registration No: 00000328793)

Department of Construction Engineering and Management

NUST Institute of Civil Engineering

School of Civil and Environmental Engineering

National University of Sciences and Technology (NUST)

Islamabad, Pakistan

(2024)

Comparison of Precision and Accuracy of Image Processing Algorithm to monitor Safety Violations at a construction site

By

Jamal Farooqi

(Registration No: 00000328793)

A thesis submitted to the National University of Sciences and Technology, Islamabad,

in partial fulfillment of the requirements for the degree of

Master of Science in

Construction Engineering and Management

Thesis Supervisor: Dr. Muhammad Usman Hassan

NUST Institute of Civil Engineering

School of Civil and Environmental Engineering

National University of Sciences & Technology

(NUST) Islamabad, Pakistan

Thesis Acceptance Certificate

Certified that final copy of MS Thesis written by Mr. Jamal Farooqi (Registration No: 00000328793), of National Institute of Civil Engineering (NICE) – School of Civil & Environmental Engineering (SCEE) has been vetted by undersigned, found complete in all respects as per NUST Statutes/ Regulations/ MS Policy, is free of plagiarism, errors, and mistakes and is accepted as partial fulfillment for award of MS degree. It is further certified that necessary amendments as pointed out by GEC members and foreign/ local evaluators of the scholar have also been incorporated in the said thesis.

Signature: ____________________

Name of Supervisor Dr. Muhammad Usman Hassan

Date: 22/01/2024

Signature (HOD): HoD Construction Engineering and Management NUST Institute of Civil Engineering School of Civil & Environmental Engineering National University of Sciences and Technology

Date: 22/01/2024

Signature (Associate Dean): ____________________

Date: 22/01/2024

Dr. S. Muhammad Jamil
Associate Dean
NICE, SCEE, NUST

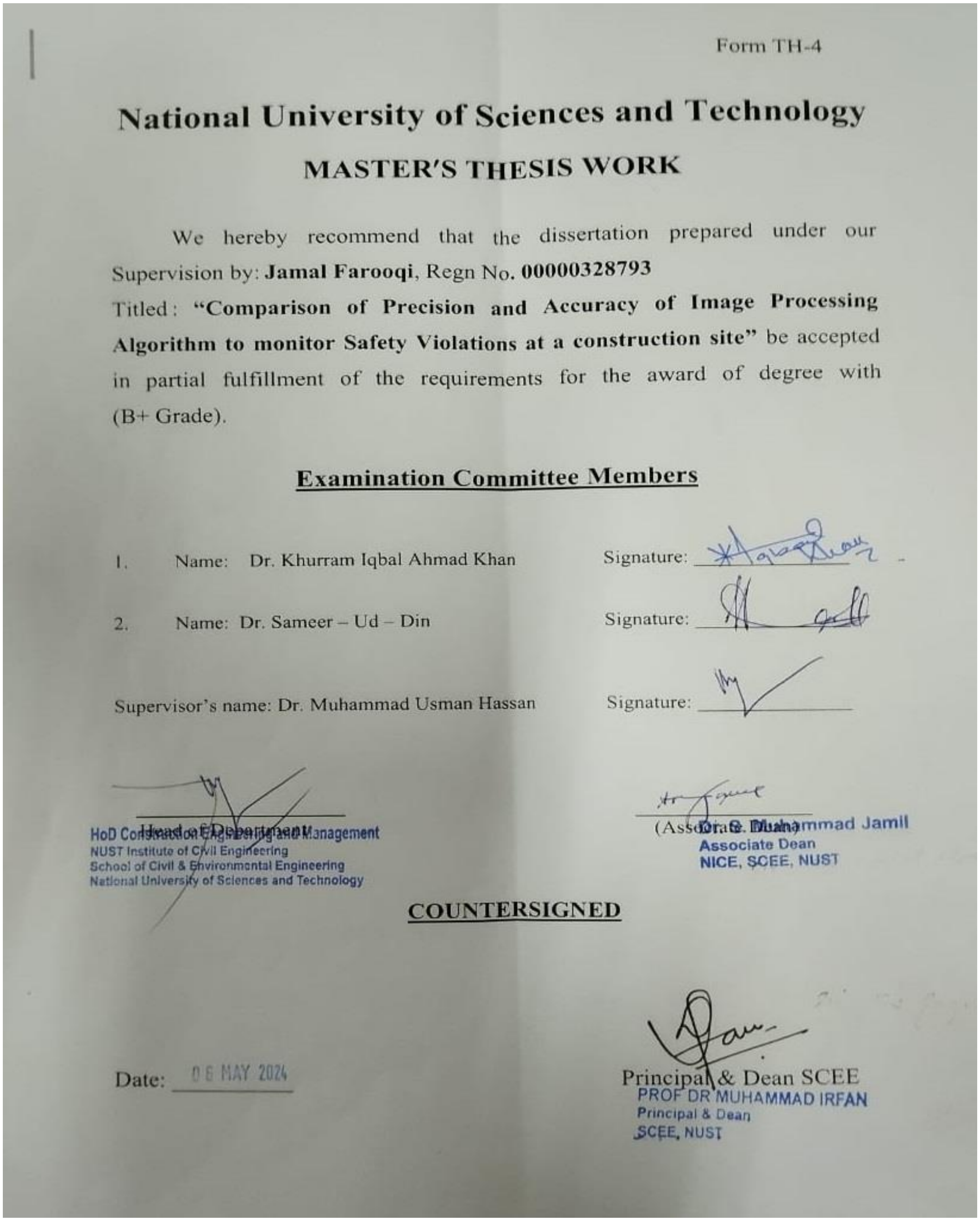
Signature (Dean/Principal): ____________________

Date: 06 MAY 2024 22/01/2024

PROF DR MUHAMMAD IRFAN
Principal & Dean
SCEE, NUST

ii

ATTACHED COPY OF TH-4 FORM



Form TH-4

National University of Sciences and Technology

MASTER'S THESIS WORK

We hereby recommend that the dissertation prepared under our Supervision by: **Jamal Farooqi**, Regn No. **00000328793**
Titled: **"Comparison of Precision and Accuracy of Image Processing Algorithm to monitor Safety Violations at a construction site"** be accepted in partial fulfillment of the requirements for the award of degree with (B+ Grade).

Examination Committee Members

1. Name: Dr. Khurram Iqbal Ahmad Khan Signature: ____________

2. Name: Dr. Sameer – Ud – Din Signature: ____________

Supervisor's name: Dr. Muhammad Usman Hassan Signature: ____________

Head of Department
HoD Construction Engineering and Management
NUST Institute of Civil Engineering
School of Civil & Environmental Engineering
National University of Sciences and Technology

(Associate Dean)
Dr. S. Muhammad Jamil
Associate Dean
NICE, SCEE, NUST

COUNTERSIGNED

Date: 06 MAY 2024

Principal & Dean SCEE
PROF DR MUHAMMAD IRFAN
Principal & Dean
SCEE, NUST

Certificate of Approval

This is to certify that the research work presented in this thesis, entitled "Comparison of precision and accuracy of image processing algorithms to monitor safety violations at a construction site" was conducted by Mr. Jamal Farooqi under the supervision of Dr. Muhammad Usman Hassan. No part of this thesis has been submitted anywhere else for any other degree. This thesis is submitted to the Department of Construction Engineering & Management (CE&M) in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Field of Construction Engineering & Management, Department of Construction Engineering & Management (CE&M), National University of Sciences and Technology (NUST).

Student Name: Jamal Farooqi Signature: ____________

Examination Committee:

a) GEC Member 1:
Dr. Khurram Iqbal Ahmad Khan
Assistant Professor
Signature: ____________

b) GEC Member 2:
Dr. Sameer – Ud – Din
Assistant Professor
Signature: ____________

Supervisor Name: Dr. Muhammad Usman Hassan Signature: ____________

Name of HoD: Dr. Muhammad Usman Hassan Signature: ____________

Name of Associate Dean: Dr. S. Muhammad Jamil Signature: ____________

Name of Dean/Principal: Dr. Muhammad Irfan Signature: ____________

Dr. S. Muhammad Jamil
Associate Dean
NICE, SCEE, NUST

PROF DR MUHAMMAD IRFAN
Principal & Dean
SCEE, NUST

Author’s Declaration

I, Jamal Farooqi hereby state that my MS thesis titled “Comparison of precision and accuracy of image processing algorithms to monitor safety violations at a construction site” is my own work and has not been submitted previously by me for taking any degree from this University National University of Sciences and Technology (NUST) or anywhere else in the country/ world.

At any time if my statement is found to be incorrect even after I graduate, the university has the right to withdraw my MS degree.

Name of Student: Jamal Farooqi

Date: 22/01/2024

PLAGIARISM UNDERTAKING

I solemnly declare that research work presented in the thesis titled “Comparison of Precision and Accuracy of Image processing algorithms to monitor safety violations at a construction project” is solely my research work with no significant contribution from any other person. Small contribution/ help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero-tolerance policy of the HEC and National University of Sciences and Technology (NUST) towards plagiarism. Therefore, I as an author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly referred/cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of MS degree, the University reserves the rights to withdraw/revoke my MS degree and that HEC and the University has the right to publish my name on the HEC/University website on which names of students are placed who submitted plagiarized thesis.

Student/Author Signature: ____________

Name: Jamal Farooqi

DEDICATION TO

My grandfather (Professor Saeedullah Farooqi)

ACKNOWLEDGMENTS

I am thankful to Almighty Allah, who gave me strength to complete my research. I would like to pay debt of gratitude to my advisor Dr. Muhammad Usman Hassan, for his fathomless guidance, valuable time, and encouragement, to complete my research work. I also owe my special thanks to Professor Saeedullah Farooqi and other fellows for their support. I pay my earnest gratitude with a sincere sense of respect to my parents, colleagues and to sincere family members for their unending encouragement, prayers, and patience. I would like to thank my dear friends and colleagues for their continuous support and patience that helped me in completing this study.

ABSTRACT

Ensuring the reduction of fatalities and injuries is paramount in construction site management. Recent advancements in computer vision have proven instrumental in monitoring complex site conditions and progress through image processing surveillance. Although multiple imaging algorithms are currently employed for visualizing site progress and identifying safety violations, there exists a critical need to ascertain the most robust algorithm for this purpose.This study focuses on the efficacy of different computer vision models, namely YOLO Series, Detectron 2 and GroundNino in the detection of safety violations and the improvement of safety protocols within the construction sector. The results emphasize YOLOv8 as the preferred option as compared to its predecessors and detectron 2 owing to its remarkable efficiency, precision, adaptability, and developer-centric attributes. The real-time processing capabilities of YOLOv8, in conjunction with its high precision, render it a well-suited solution for the prompt monitoring of safety in the ever-changing settings of construction sites. This study highlights the potential of computer vision technology in enhancing safety protocols within the construction industry, hence facilitating more effective safety monitoring and accident avoidance.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

The construction industry stands as a crucial driver of economic growth, infrastructure development, and urbanization. The constant occurrence of safety violations, however, casts a dark shadow under the large framework and busy work zones. These violations, which can include a wide variety of unsafe situations and activities, put construction workers in danger, cause projects to run behind schedule, and negatively impact the industry's reputation as a whole [1]. It is imperative that these infractions be corrected and safety measures enhanced. In order to determine if an image-based algorithm for detecting safety infractions on construction sites is feasible, this study goes into the fields of image processing and computer vision. In this chapter, we'll investigate how widespread safety infractions are, how they affect operations, and how cutting-edge technology can help.

1.1 Safety Violations in Construction

Multifaceted safety violations plague the building and construction business [2]. These violations encompass a diverse range of unsafe practices and hazardous conditions that jeopardize the well-being of workers and the projects they are engaged in constructing. Inadequate protective measures, poor training, noncompliance with established safety regulations, and unsatisfactory jobsite circumstances are just a few examples [3] of the many ways in which these rules are broken.

Workers often fail to take necessary precautions, for as by not wearing required PPE like helmets, safety harnesses, or steel-toed boots [4]. Inconsistent or insufficient safety training can also reduce workers' awareness of hazards and the need for preventative measures [5].

It is also disturbing that safety procedures are not always followed. This can manifest as a failure to perform necessary safety inspections, an avoidance of established safety procedures, or a delay in responding to possible threats. Poor worksite conditions further exacerbate the problem, including issues such as inadequate lighting, cluttered work areas, or unsafe scaffolding structures [6].

The Occupational Safety and Health Administration's statistical data highlights the problems. According to OSHA's documented records, there is a significant occurrence of accidents and injuries on an annual basis within the construction industry [7].

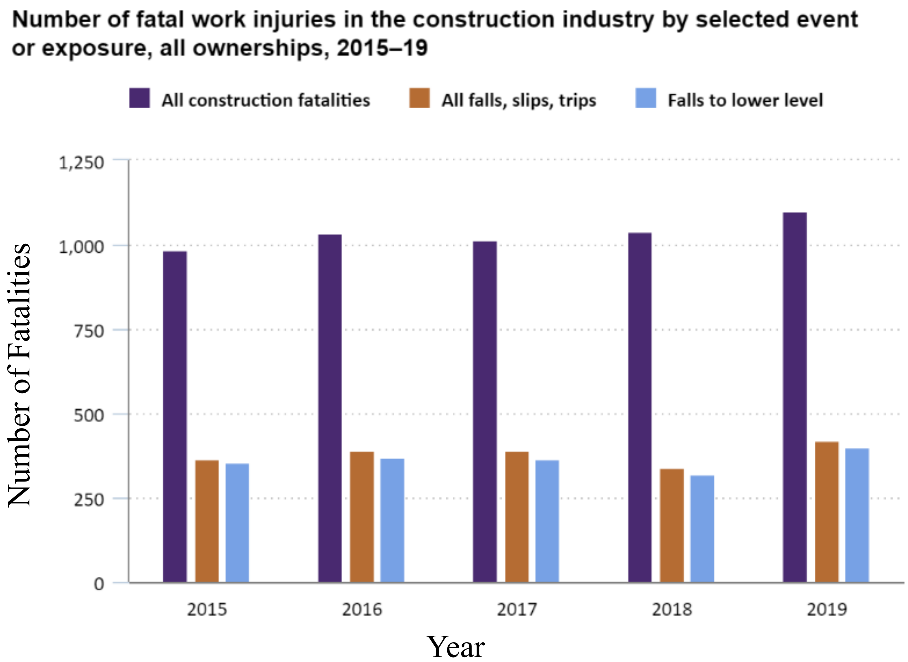


Figure 1 number of work injuries in construction industry (Source: [8])

The above figure shows that There were 1,102 fatal injuries in the construction industry in 2019 in private industry and government. These deaths represented 20.7 percent of total workplace fatalities in the United States (5,333). Falls, slips, and trips were the most frequent type of fatal event in the construction industry, representing 37.9 percent of all fatalities (418 of 1,102). This was a 22.9-percent increase in fatal falls, slips, and trips over 2018. Most fatal falls, slips, and trips are from falls to a lower level [8].

1.3 Number of incidents on construction site

Safety violations on construction sites result in a significant number of incidents, causing harm to workers, project delays, and financial burdens. These incidents underline the urgent need for effective safety monitoring and enforcement measures within the construction industry.

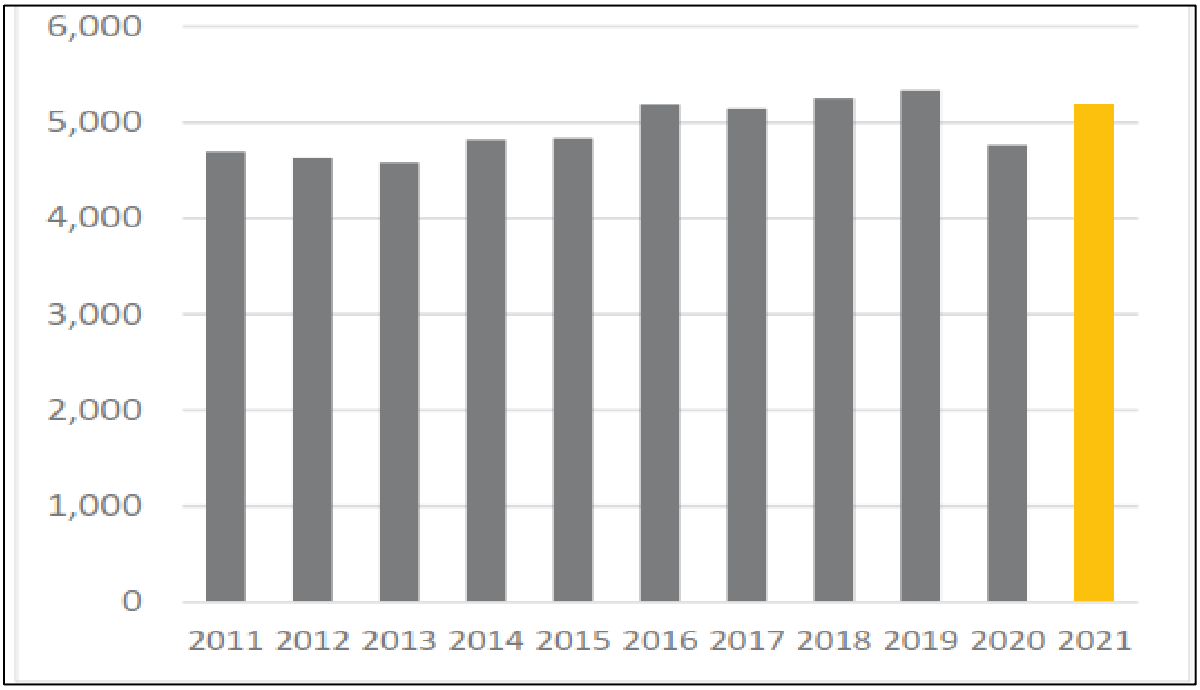


Figure 2 number of fatal work injuries (Source: [9a])

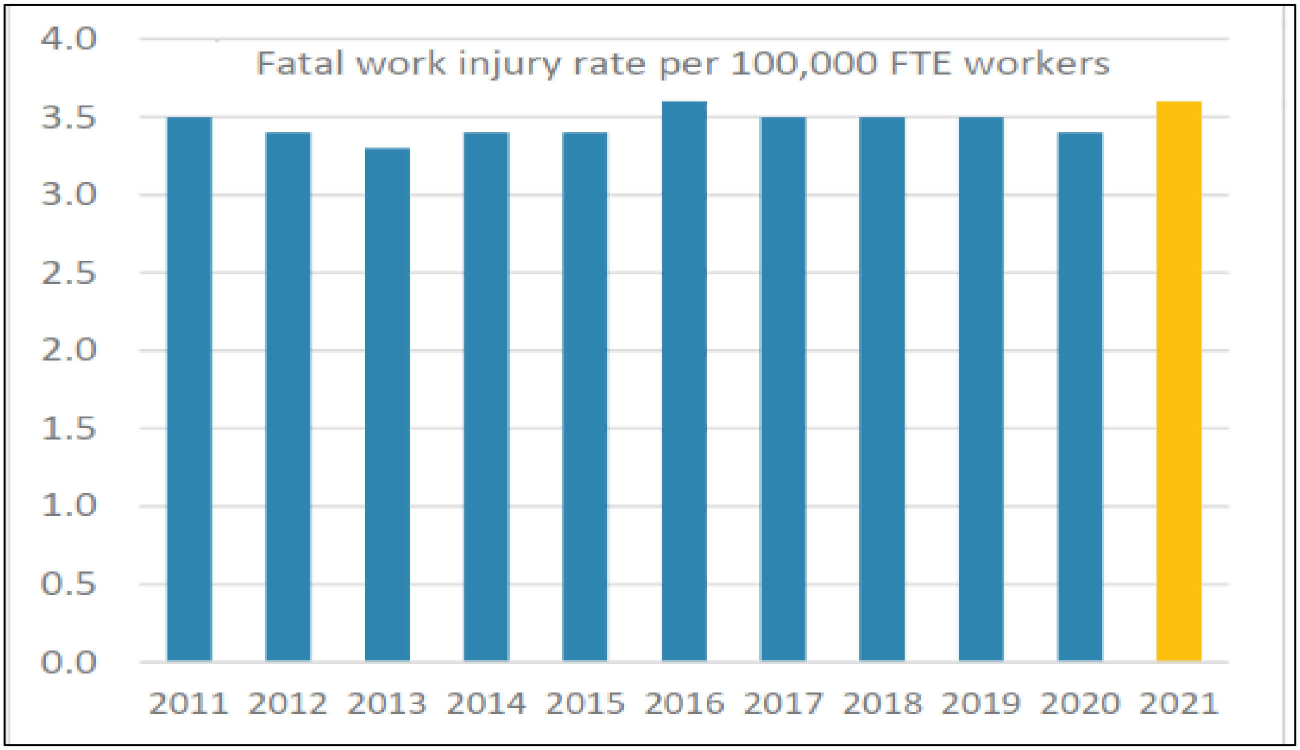


Figure 3 fatal work injury rate (Source: [9b])]

The above figure shows that There were 5,190 fatal work injuries recorded in the United States in 2021, an 8.9-percent increase from 4,764 in 2020, the U.S. Bureau of Labor Statistics reported today [9a]. The fatal work injury rate was 3.6 fatalities per 100,000 full-time equivalent (FTE) workers, up from 3.4 per 100,000 FTE in 2020 and up from the 2019 pre-pandemic rate of 3.5. (See chart 2.) These data are from the Census of Fatal Occupational Injuries (CFOI) [9b].

1.4 Impact of Safety Violations

The repercussions of construction site safety violations, extends far beyond immediate safety concerns, affecting both human lives and the outcomes of construction projects [10]. Safety incidents and accidents within construction sites can have dire consequences, with wide-ranging ramifications that reverberate through various facets of the industry [11].

Foremost among these consequences is the financial toll exacted by safety violations. Construction accidents frequently lead to substantial increases in project costs [12]. These costs encompass a spectrum of expenditures, ranging from immediate medical expenses for injured workers to long-term legal liabilities and compensation claims filed by affected parties. The financial burden arising from accidents can be substantial, significantly eroding profit margins and diverting resources away from project completion.

1.5 Controlling the Impact of Violations

The building industry and regulatory organizations have devised a number of steps to mitigate the effects of safety violations. Safety training, inspections, and severe fines for noncompliance are all required by law under occupational health and safety rules [13]. Companies also spend money on safety management systems and tools to guarantee they're following the rules [14].

1.6 Modern Techniques for Monitoring Violations

New methods for detecting and reporting construction site safety violations have been the subject of extensive study in recent years. While manual inspections and checklist-based systems have been the standard up until now, they have their limitations when it comes to effectively detecting infractions. As a result of their extraordinary ability to automate the detection of safety infractions through the analysis of visual data, cutting-edge technologies like image processing and computer vision [15] provide intriguing answers.

1.7 Traditional Limitations:

Construction safety enforcement has traditionally relied on manual inspections, typically conducted by safety inspectors or supervisors [16]. However, human variables such as the possibility of oversight or subjectivity in determining safety compliance are inherent limitations of such inspections. In addition, they may not be able to cover all regions of a construction site at once, and they can be time-consuming and resource-intensive [17]. Although checklists are systematic, they are not always effective at spotting safety lapses due to changes in circumstances.

1.8 Image Processing and Computer Vision:

When used to construction site monitoring, image processing and computer vision techniques have resulted in an evolutionary change [18]. Real-time analysis of photos and videos collected at building sites is made possible by these technologies [19] thanks to the use of high-powered digital cameras, sensors, and sophisticated algorithms. This allows for the immediate detection of a wide variety of safety violations, such as employees working without proper PPE, unapproved entry into potentially dangerous areas, and risky actions like standing on uneven ground.

Fast and precise processing of large amounts of visual data is the key to the success of image processing and computer vision [20]. These tools provide a preventative method of enforcing safety standards by identifying infractions that could otherwise go undetected during manual inspections. In addition, accidents can be avoided and the occurrence of safety violations can be reduced thanks to real-time monitoring that permits fast responses to potential safety threats.

1.8.1 Effects and Benefits:

Image processing and computer vision for construction safety monitoring have been beneficial. First, it may drastically minimize safety violations and accidents, making workplaces safer and reducing injuries. Safety monitoring automation frees up human resources for other vital construction site management activities.

1.9 Image Processing and Computer Vision

Computers can now analyze and interpret visual data from images and videos thanks to developments in image processing and computer vision [21]. These safety monitoring devices have revolutionized construction safety monitoring. Image processing and computer vision technologies can accurately and efficiently identify a wide range of safety infractions in building site photos and videos.

1.9.1 The Power of Computer Vision:

However, computer vision goes beyond traditional image processing by giving computers the ability to grasp and make sense of visual data. It requires sophisticated algorithms that can spot abnormalities, patterns, and objects in visual data. The construction sector makes extensive use of computer vision systems because of their ability to analyze the content of photos and videos and spot potential safety violations, such as employees not wearing helmets or entering prohibited areas.

1.10 Enhancing Safety Monitoring:

There are many benefits to using image processing and computer vision technologies into construction safety monitoring systems. Continuously scanning enormous amounts of visual input is an impossible effort for humans, but not for these systems. They function in real time, enabling the detection of security breaches as they happen.

In addition, these technologies may be relied on to do inspections without bias or fatigue affecting the results, unlike human inspectors. This ensures that all construction sites are held to the same standards of safety enforcement, which is essential for preventing accidents and injuries.

1.11 Problem Statement:

Accidents, injuries, and delayed projects are all too common results of safety infractions in the construction sector [22]. Inefficient and lacking in real-time information, current monitoring approaches frequently rely on intrusive procedures or human inspections. The overarching question this study seeks to answer is, then, whether or not an image-based algorithm with the potential for precise and efficient identification of safety infractions happening on construction sites is viable to design. The question arises where we find the efficiency and precision of multiple image processing algorithms to monitor which one works best for the safety violations to be monitored.

1.12 Previous Work and Research Gap

In order to keep everyone on the building site safe at all times, a new and innovative safety model has been developed that uses the Internet of Things (IoT) [23]. Computer vision is a sophisticated and automated method for extracting and processing video and image information to monitor construction workers' health and safety.

A scholarly article detailed the implementation of a system that facilitates the transmission of safety-related data from various construction projects to a centralized database. This database generates real-time safety indicators [24]. The system offers safety indicators pertaining to specific projects as well as industry-wide statistics.

Real-time control based on intelligent video monitoring was employed to prevent worker violations on site. This article merges BIM and WSN into a unique solution that allows the building site to visually monitor safety via a spatial, colored interface and automatically remove hazardous gas. Researchers have successfully incorporated Bluetooth low-energy (BLE)-based location detection technology, building information model (BIM)-based hazard identification, and a cloud-based communication platform [25]. A new framework is presented to quantify safety infractions by feeding data from real-time location sensors to a real-time data visualization platform.

One researcher proposed a wearable technology which helped in getting the Realtime results so that they were combined to provide better understanding of the violations that were occurring author proposed the use of novel kernel filter to guess the position of the construction worker or vehicles at site by continuously judging their previous movements [26].

Prior research has examined safety violations in the construction industry. However, there exists a noticeable gap in the scholarly literature when it comes to conducting a thorough evaluation of image processing algorithms for the purpose of monitoring such infractions. Most existing research has primarily focused on traditional methods or lacked a thorough evaluation of modern technologies. This research aims to fill this gap by conducting a systematic analysis of various image processing algorithms' effectiveness in improving construction site safety.

1.13 Research Aim:

This research aims to make sure that we monitor safety violations of a construction site using computer vision algorithms. While monitoring the violations, we will be doing the comparisons of those image processing algorithms and hence finding which one works best for the violations monitoring. Using methods from image processing and computer vision, this study aims to develop a real-time monitoring system capable of automatically detecting security breaches. By reaching this goal, the research hopes to greatly increase safety on construction sites, lower the number of safety infractions, and help raise industry-wide safety norms. This study is to develop a technologically superior and effectively implemented strategy for ensuring the safety of construction workers and the completion of building projects.

1.14 Research Question

Can we find out which algorithm is the most efficient and precise to monitor safety violations at a construction site?

1.15 Research Objectives:

The research objectives given below aim to guide the investigation into developing a non-intrusive image-based algorithm for enhanced safety monitoring at construction sites, ultimately improving safety outcomes and operational efficiency.

1) To Identify the violations that are occurring all over the construction industry.
2) To monitor Safety violations using different Image processing Algorithms
3) To determine which algorithm is accurate and precise for monitoring construction violations

CHAPTER 2: LITERATURE REVIEW

2.1 Violations by construction workers

Cognitive biases, also known as CBs, are psychological factors that have a significant impact on the decision-making process of individuals regarding unsafe behaviors.

Based on the established definition of cognitive biases (CBs), previous research has provided evidence that CBs have the potential to result in a distorted comprehension of both the external circumstances and one's own self, subsequently giving rise to inappropriate behavioral responses [27]. Individuals rely on easily recalled information when interpreting the context, as evidenced by the available details. The availability of easily retrievable information is inherently constrained and can lead to flawed and shallow decision-making. Robson et al., (2020) have observed that construction workers frequently refrain from utilizing fall prevention equipment [28].

The utilization of fall arrest equipment was infrequent. The prevalence of job insecurity within the dynamic work environment significantly impacted the conduct of apprentices, despite their awareness of proper safety protocols. This issue was particularly pronounced among workers with limited experience.

2.2 Violations by Construction Co workers

Safety violations in this context bear resemblance to the concept of "workarounds," which has been extensively employed in healthcare field [29]. "workarounds" terminology points towards the practice observed in intricate socio-technical systems, wherein workers frequently devise alternative methods to overcome obstacles or impediments in the workflow in order to successfully complete their tasks.

When engaging in routine violations, workers tend to consistently opt for the path that requires the least amount of effort or engage in what is commonly referred to as "cutting corners" [30]. In contrast, situational violations are typically influenced by situational constraints within the workflow, thereby presenting challenges or rendering adherence to established rules difficult or unattainable.

2.3 Computer vision through Image processing

Object recognition is an integral part of image processing. The researchers have ensured the utilization of the bounding box and class probability within an image as a unified regression problem. The utilization of the YOLO technique facilitated the determination of both the spatial coordinates and the categorical classification of the object under investigation. In this study Kim et al., (2021) use the geometric transformation in order to map the captured image with its position on the drawings [32]. The techniques used were Scale, Distortion, Cut, Replace and Stretching.

The improvement of image quality for its processing to make sure that the safety risk assessment is done. According to Nie et al., (2020), The segmentation of images can be achieved through the application of grey scale techniques [33], which employ the weighted average method to obtain the most appropriate grey image. The author employs a weighted average methodology to calculate the components.

Pour Rahimian. et al., (2020), use the gaming engine such as unity for the interoperability and integration of BIM and on-site photographs to give a realistic on-site model update [34a]. The author strengthens the use of image processing as a tool because according to him it is still the easiest and most reliable method of getting the site information [34b].

Bartol et al., (2021) Proposed a framework in which RGB image color acquis ion was done in order to update the 4D model in place [35]. Rahimian et al., (2020), used the construction site images and superimposed those images onto the BIM model to estimate the delay that has been occurring at the site [36]. Some of the factors that must be catered while getting a good quality image is the quality of the camera. The cameras that are used must be placed at a height so that there are no such interruptions to the line of site during the construction site.

In their proposal, Wu and H suggested that computer vision has the potential to utilize images obtained from construction sites in order to provide valuable information. They further proposed that the formal representation of safety regulatory knowledge can be achieved through the utilization of ontology and SWRL (semantic web rule language) rules [37].

2.4 Image Classification through Image processing

Fusion was used by Zhung et al. (2021) to combine RGB and LiDAR data for 3D semantic segmentation [38]. The ability to map an RGB image onto a polar grid is the main innovation of this study. This representation is subsequently utilized in the development of mid-level fusion architectures.

Zhuang et al., (2021), employed the deep convolutional neural network methodology known as deeplabv3+ to identify and measure subtle interlayers. To validate this approach, the author employed a comprehensive database comprising numerous images [39]. The edge detection technique is used to detect the deformations in the structures [40]. In the detection of cracks

Jacintha et al., (2022), used the techniques of fuzzy C-means clustering and chose the image with the lowest brightest value in the cluster as the one with the crack [41]

2.5 Computer Vision and Machine Learning

Several algorithms have been created for the purpose of monitoring construction safety violations using computer vision. One set, known as shallow learning methods, includes algorithms like support vector machines and histograms of directed gradients, while the other, known as deep learning methods, includes algorithms like convolutional neural networks (CNNs) and recurrent neural networks (RNNs) [42, 43]. The accuracy of detection is hampered by shallow learning approaches since they rely so heavily on features that must be developed by hand.

Computer vision has been employed to analyze job sites by utilizing several cameras, which enable the collection of data for visual data analysis. Integration of BIM framework is done with the use of PMGA for the placement of cameras at a construction site [44].

In computer vision technology people have tried to extract information from short and long sequence of videos. Hassanin and M. Khan (2021) used different types of models such as a discriminative action classifier based on support vector Machines(SVM) and a hidden Markov Model for disintegrating the images from RGBD into sequence of atomic actions [45]. Aggarwal, (2018) use the deep neural algorithms [46], to find out if any non-certified workers were working on site so that any unprofessional work could be stopped from carried around.

Xu and Wang, (2020) use the computer vision along with safety prewarning system in which the risk assessment model is integrated with the surveillance cameras and the extracted images are used to let the model find the results [47]. The application of computer vision was done in order to extract some real time information from the images that were captured through the cameras. Once the information is taken out, it is used with the safety assessment model.

Fang, W. et al. (2020) use the integration of computer vision technology with ontology enables the creation of a knowledge graph that facilitates the automated identification of hazards [48a]. The model proposed in this study employs a hybrid approach, integrating Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) techniques, in order to autonomously detect and classify hazardous behaviors demonstrated by workers [48b].

A CNN model was developed by Fang et al., (2022), used to ensure if the workers were wearing the harness belt or not. What this model did was that it developed algorithms which detected the workers presence and also if they were wearing the harness or not [49].

The essential use of machine learning comes in places where the computer is programmed in such a way to learn from past experiences or trained by a programmer to model and predict using statistical programs. The machine learning might be supervised or unsupervised. The supervised machine learning is based on how the computer program has learnt from the labelled datasets. It is further classified into labelled and unlabeled datasets. Unsupervised learning is concerned with making datasets that are unlabeled is categorized into clustering and dimension reduction techniques. Knowledge based system is concerned with machine decision making based on existing knowledge. Robotics have come to the help in the construction industry due to the shortages in the skilled labor. One application of machine learning was highlighted in the safety analysis of scaffolding where the safety predictions were done based on the strain datasets of the failure of the scaffolding columns.

2.6 Use of BIM for Safety Monitoring

According to Lu et al (2020), BIM is a "digital representation of the building process" that allows for the interoperability and sharing of data between different software programs [50]. In the AEC industries, BIM represents the cutting edge of technology. However, BIM implementation is slowest in some regions.

The application of BIM can be seen in multiple factions of the construction industry. One such application can be the monitoring of progress and its analysis using BIM and Information Integrated technology. Rising interest in the use of BIM to facility management has led to the development of a prototype that focuses on the delivery of an asset information model to a shared data environment. Safety hazard can be identified by the virtual integration of BIM. These hazards are usually identified during the planning face and hence allows their mitigation to become easy.

Xu and Wang, (2020) proposed a BIM model to do the real time risk analysis by incorporating the computer vision into it [52]. This analysis is done through a risk model which is categorized into different risk levels and suggestions are given to the corresponding risk bearers. The study aimed at creating a dense photo cloud which is then fused with the IFC based BIM model to show the automated physical progress the ongoing project site.

Boje et al., (2020) and others colleague identified the shortcomings in BIM due to the semantic understanding of a construction site [53]. The essential use of machine learning comes in places where the computer is programmed in such a way to learn from past experiences or trained by a programmer to model and predict using statistical programs.

2.7 Different method used to monitor the violation outside of construction:

While monitoring safety violations within construction sites is critical, it is equally important to extend surveillance beyond the boundaries of construction areas to ensure a safe environment for the public and neighboring communities. The computer vision field has witnessed remarkable advancements in monitoring violations occurring outside of construction sites. The section below explores the various methods employed to achieve this goal.

2.7.1 Geospatial Analysis

Geospatial analysis, as exemplified by the research conducted by Afzal et al. (2021), utilizes geographic information systems (GIS) and satellite imagery to observe construction-related activities and identify instances of safety violations in the adjacent regions [54]. This methodology has the capability to monitor alterations in the physical environment, patterns of land utilization, and incursions into areas designated for safety purposes.

2.7.2 Object Detection and Recognition

Object detection and recognition is a key approach within the field of computer vision. Xiao and Kang (2021) conducted a study whereby they investigated the use of deep learning methodologies for the purpose of identifying and classifying objects or actions that could potentially present safety hazards in areas next to building sites [55]. This encompasses the identification of vehicles without authorization, pedestrians trespassing in restricted zones, and the recognition of possible risks in adjacent streets.

2.7.3 Environmental Monitoring

Environmental factors play a significant role in safety. Authors like Salamone et al., (2021), have explored computer vision methods for monitoring environmental parameters, such as air quality and noise levels, in areas surrounding construction sites [56]. This information can help detect violations related to environmental regulations.

2.7.4. Traffic Surveillance

The maintenance of road user safety in the vicinity of building sites is of paramount importance. Researchers such as Outay et al. (2020) have conducted investigations on computer vision systems that are equipped with cameras for the purpose of monitoring traffic conditions (57). These systems has the capability to identify instances of traffic violations, congestion, or accidents in close proximity to construction sites, so guaranteeing the safety of both construction personnel and the general public.

2.7.5 Community Engagement and Reporting Apps

Community involvement is vital in identifying safety violations outside construction sites. Researchers like Bubalo et al., (2019), have developed smartphone apps that allow residents to report safety concerns with geotagged images [58]. These apps can streamline communication between the public and construction site management.

2.8.YOLO Versions (YOLOv8,v7,v6,v5)

Significant progress has been made in the field of image and video object detection thanks to the You Only Look Once (YOLO) algorithm, which plays a critical role in identifying safety violations. The most recent updates to this system, YOLOv8 and YOLOv7, have significantly improved its usefulness for monitoring worker safety on construction sites. Several improvements have been made to the YOLO algorithm since its conception by researchers including Terven & Cordova (2023), increasing its accuracy and efficiency [59]. YOLO's fundamental strengths lie in its detection and localisation capabilities, making it ideal for the detection of security breaches in real time. Fast responsiveness and pinpoint accuracy make YOLO a useful tool for job site safety.

2.8.1 A Detailed Breakdown of YOLO's Functionality:

Bounding Box Prediction:

The capacity of YOLO for forecasting is regarded as one of its primary advantages. During the analysis of each individual grid cell, the process involves the creation of hypothetical boxes that fully enclose any detected items or areas of interest. The bounding boxes exhibit a high level of precision, enabling accurate localization of objects within the image. The attainment of high accuracy is necessary in order to effectively monitor safety infractions, since it enables the algorithm to accurately pinpoint the precise site of the breach.

Real-time Object Detection:

In contrast to alternative approaches that depend on predefined regions of interest (ROIs) or sliding windows, YOLO has the capability to efficiently analyze photos and videos in real-time. This characteristic renders YOLO highly advantageous for the ongoing surveillance of construction sites. The YOLO algorithm partitions the image into a grid and methodically evaluates each grid cell to detect the presence of objects or violations.

Class Probability Assignment:

Not only can YOLO forecast bounding boxes for objects, but it also assigns class probabilities to those items. Differentiating between different things or transgressions is greatly aided by this categorization stage. YOLO's flexibility in categorizing objects makes it a useful tool for spotting security breaches. It can tell the difference, for instance, between a worker in a building site who is wearing protective gear and one who is not.

2.8.3 Versatile Applications of YOLO in Construction Site Safety:

The versatility of YOLO extends to its applications in addressing a wide range of safety infractions on construction sites:

Unauthorized Access:

The algorithm works well to detect trespassers and other unauthorized visitors to the building site. YOLO improves site security and safety by identifying things and people that don't follow specified safety rules.

Safety Equipment Compliance:

When construction workers aren't putting on the appropriate helmets, bright vests, and safety glasses, YOLO can pick up on it. To do this, it analyzes photos in real time and uses object classification to determine whether or not certain safety devices are present.

Equipment Misuse:

whether properly educated, YOLO can detect whether construction tools are being used improperly or are being misused. For instance, it can detect instances of dangerous machinery operation, hence lowering the probability of mishaps resulting from equipment infractions.

2.9 Detectron 2

Facebook's AI Research team created Detectron 2, a free software framework for object detection and image segmentation, which has since become an essential tool for safety monitoring on building sites [60]. Detectron 2's variety of pre-trained models has made it a popular tool for researchers and industry practitioners engaged in monitoring construction site safety in order to improve safety standards and reduce breaches. Wu et al. (2019) highlight Detectron 2's exceptional versatility as a defining characteristic [61]. Detectron 2's modular design and architecture makes it easy to incorporate additional data sources and tailor detection jobs to specific needs. By allowing for customization and expansion, the platform may be made optimal for use in construction safety monitoring.

By incorporating data from actual building sites into current models, Detectron 2 is used to pinpoint safety failures. Thanks to the platform's data-driven ability to detect and flag safety infractions in real time, incidents can be avoided before they even happen.

Below are some key applications of Detectron 2:

- **Identification of Workers Without PPE**:

Detectron 2 can be set up to monitor construction sites for the absence of required personal protective equipment (PPE) such hard helmets, gloves, and safety vests. Because of this, you may rest assured that all necessary safety measures are taken, thus cutting down on the number of possible injuries.

- **Detection of Unattended Materials**:

Additionally, the system has the capability to be configured in a manner that enables the monitoring of construction sites with regards to unattended equipment and gear. The capability of Detectron 2 to identify and annotate such occurrences becomes advantageous in preventing accidents and ensuring the efficient operation of systems.

- **Unauthorized Visitors**:

The degree of harm resulting from individuals trespassing onto work places is indeed substantial. In order to ensure the presence of solely authorized individuals at a given location, Detectron 2 has the capability to be programmed for the purpose of identifying and notifying security personnel regarding the presence of unauthorized visitors.

Detectron 2's flexibility, wide variety of pre-trained models, and ability to integrate real-world data make it an excellent choice for creating construction-specific security monitoring programs. Detectron 2 is an indispensable tool for guaranteeing conformity with safety requirements and averting potential problems in the construction business, where safety ranks among the highest priorities.

2.10 Darknet Keras

Another well-liked open-source neural network framework, Darknet [62] is especially effective and quick at object detection. The Darknet Keras variation combines Darknet's security with Keras's ease of use. This framework is often used to create specialized models for detecting safety violations on building sites.

Researchers that are interested in developing custom solutions will find Darknet Keras to be especially helpful. Scientists have trained models to spot certain types of safety infractions, like employees working in prohibited areas or misusing safety gear. With such granular tailoring, it's possible to reliably identify illegal acts.

2.11 Application of Algorithms by Modern Researchers

Researchers today have been actively employing computer vision and image processing methods in order to augment the monitoring of safety on building sites. The researchers employ deep learning frameworks such as YOLOv8, Detectron 2, and Darknet Keras to construct models capable of real-time detection of violators. Analysts have implemented a methodology centered on data analysis, wherein they have gathered and annotated visual media depicting construction sites. Subsequently, the aforementioned datasets are employed for the purpose of training and refining the current models. Munir and Siddiqui (2023) utilized the YOLOv8 framework to develop a bespoke model aimed at identifying safety breaches within building sites [63]. This strategy showcases the versatility and efficacy of these algorithms.

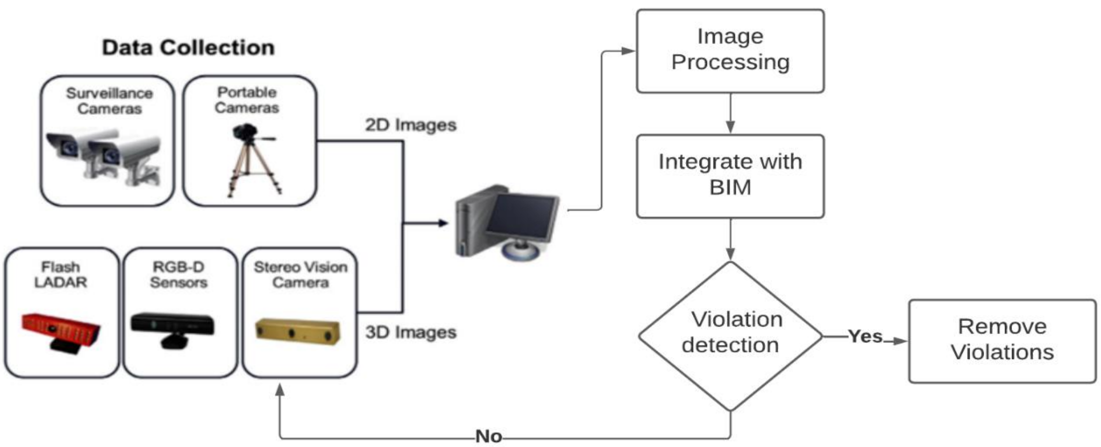
2.12 Image Classification in Safety Violation Detection

The task of classifying images plays a crucial role in the detection of safety violations. Researcher like Chen et al., (2023), utilize convolutional neural networks (CNNs) for the purpose of categorizing objects and activities depicted in photographs of building sites [64]. Convolutional neural networks (CNNs) possess the ability to discern between conditions that are deemed safe and those that are considered harmful by relying on visual signals.

Image classification models are trained using datasets that consist of tagged images together with accompanying labels that indicate whether safety violations are present or absent. For example, Yang et al. (2019) employed a convolutional neural network (CNN) methodology to categorize photos into "safe" or "unsafe" categories by discerning the existence of safety equipment on construction workers [65]. This methodology enables the automatic identification of non-compliance instances pertaining to personal protective equipment.

CHAPTER 3: RESEARCH METHODOLOGY

3.1 Research Framework



To facilitate a comprehensive comparison of YOLOv8, Detectron 2, and RCNN for safety violation detection at construction sites, a meticulously crafted dataset of images and videos captured within construction environments was collected. This dataset is essential for algorithm training, validation, and evaluation.

3.1.1 Image Collection

It is vital to ensure the diversity and representativeness of the dataset [66], which involves collecting photographs with extreme care. These images should capture a wide variety of construction circumstances and scenarios. The following are the stages of an image collection process:

- **Scenario Diversity:** A systematic approach was taken to capture a wide spectrum of construction scenarios. That include images showcasing construction workers engaged in various tasks, heavy equipment in operation, materials in transit, and the presence of safety gear [67].

- **High-Resolution Imaging:** High-resolution cameras was employed to capture detailed images [68]. This ensures that the dataset contains images with the clarity and precision required for object detection and safety violation identification.

- **Safety Violation Scenarios:** Special attention was given to capturing images that depict safety violations. These may include instances of workers not wearing personal protective equipment (PPE), unauthorized access to restricted areas, and improper equipment usage [69].

3.1.2 Video Footage

Video footage offers a dynamic perspective of construction site activities, providing a more realistic and challenging dataset for algorithm evaluation [70]. The collection of video footage was adhered to the following principles:

- **Continuous Recording:** Video cameras continuously record construction site activities, ensuring that a rich and continuous stream of video data is available for evaluation. This approach replicates the dynamic nature of construction environments.

- **Multi-Camera Setup:** Multiple surveillance cameras strategically positioned across the construction site was used to capture video footage from various angles and viewpoints.

This setup mirrors real-world monitoring scenarios, where construction sites are equipped with comprehensive camera networks.

- **Variation in Lighting Conditions:** To account for diverse lighting conditions, video capture was performed throughout the day and night, as well as under varying weather conditions [71]. This variation introduces complexity and challenges to the dataset, allowing for a comprehensive assessment of algorithm performance.

3.1.3 Annotation

The annotation process is a critical step to facilitate supervised learning and model training. Each collected image and video were undergone meticulous annotation to label objects, regions of interest, and safety violations [72]. Annotations are instrumental in providing ground truth data for training and evaluation. Key aspects of annotation include:

- **Safety Violation Classification:** Safety violations, such as the absence of PPE, unauthorized access, and equipment misuse, were categorized and labeled in the dataset [73]. These annotations enable the models to differentiate between safety compliance and violations accurately.

- **Object Labeling:** Objects of interest within the images and videos was labeled [74], including construction workers, safety gear, equipment, materials, and potential safety violations. These labels are essential for training the computer vision models to recognize and classify objects accurately.

- **Bounding Box Annotation:** Bounding boxes were defined around labeled objects to precisely identify their location within the images and video frames. Those bounding boxes serve as reference points for object detection algorithms.
- **Annotation Consistency:** To maintain annotation quality and consistency, annotators were adhered to predefined guidelines and undergo training to ensure that annotations across the dataset are accurate and coherent.

3.2 Data Preprocessing

The effectiveness of the computer vision models, YOLOv8, Detectron 2, and RCNN, in safety violation detection relies significantly on the quality and consistency of the dataset. Prior to inputting the data into these models, a series of preprocessing steps was meticulously executed to ensure uniformity and compatibility.

3.2.1 Data Cleaning

Data cleaning is a pivotal step in the data preprocessing phase [75]. This process involves the identification and removal of noisy, irrelevant, or corrupt data instances from the dataset. The data cleaning process adheres to the following technical guidelines:

- Quality Assurance: Images or video frames with poor quality, such as those affected by motion blur, excessive noise, or low resolution, were identified and excluded from the dataset. This quality assurance step guarantees that the data used for training and evaluation is of the highest quality possible.

- Noise Removal: Outliers, anomalies, and data instances that do not contribute to the research objectives were identified and eliminated. This ensures that the dataset maintains a high signal-to-noise ratio, which is crucial for effective model training.

- Relevance Filtering: Irrelevant or redundant data that does not align with the research objectives was filtered out. This ensures that the dataset remains focused on construction site safety violation detection.

3.2.2 Data Augmentation

The utilization of data augmentation strategies is crucial in improving the resilience and variety of the dataset [76]. By implementing controlled modifications, these methodologies effectively equip the dataset to effectively address a diverse array of real-world problems. The subsequent technical specifications delineate the procedure of data augmentation.

- Geometric Transformations: Techniques such as rotation, scaling, and flipping were applied to the images and video frames. Rotation introduces variations in object orientations, while scaling simulates objects at different distances from the camera [77]. Flipping horizontally and vertically introduces mirror images, expanding the dataset's diversity.

- The photos and frames undergone brightness and contrast modifications in order to replicate diverse lighting conditions. This step involves preprocessing the dataset to effectively manage variations in illumination levels, hence enhancing the adaptability of the models.

- Artificial noise injection involved the deliberate introduction of controlled noise, namely Gaussian noise, to specific data instances (Reference 78). This simulation replicates real-life situations that involve different degrees of noise and interference, hence improving the models' ability to tolerate and handle noise.

- Color Space Variations: Data augmentation may involve converting images to different color spaces (e.g., RGB to grayscale) [79] to simulate scenarios where color information is limited or unnecessary.

3.3 Model Selection

This research compares three state-of-the-art computer vision models: YOLOv8, Detectron 2, and RCNN. These models have demonstrated effectiveness in object detection and are suitable for safety violation identification.

3.3.1 Model Training

- **Detectron 2:** Similarly, Detectron 2 were undergo training with the annotated dataset. Customizations will be made to adapt the model to construction site safety monitoring tasks.

- **YOLO series:** YOLO series was trained using the preprocessed dataset. These models were fine-tuned to detect safety violations, including the absence of safety gear, unauthorized access, and equipment misuse.

- **Ground Nino:** Ground Nino was trained and evaluated as part of the comparison. It were fine-tuned to recognize safety violations based on object detection and classification.

3.4 Object Detection Algorithm:

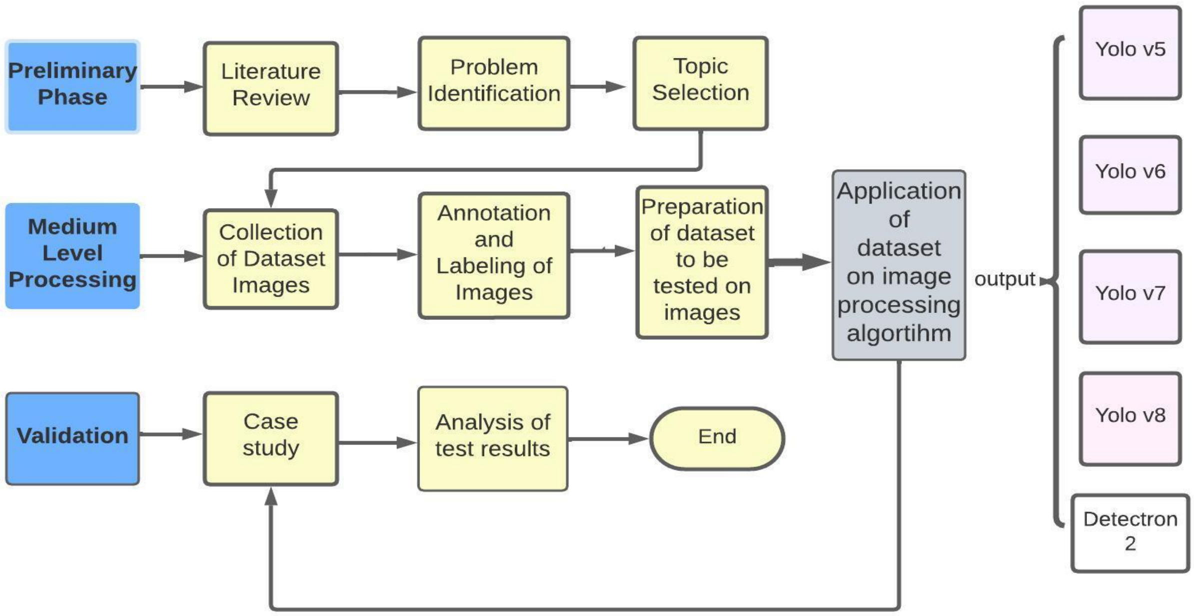


Figure 4: Algorithm

3.5 Evaluation Metrics

The assessment of the performance of YOLOv8, Detectron 2, and RCNN in the context of safety violation detection at construction sites requires a comprehensive set of evaluation metrics. These metrics provide a quantifiable and technical basis for comparing the models' capabilities. The following technical details outline the selected evaluation metrics:

3.5.1 The three matrices are Precision, Recall and F1 Score

Fundamental metrics in object detection tasks include precision, recall, and the F1 score, which provides insight into the models' capacity to detect safety violations while limiting false positives.

The recall metric (sometimes called sensitivity) measures how many positive instances were correctly predicted out of a total [80]. It evaluates the models' accuracy in detecting threats while keeping false-negative rates to a minimum. A high recall rate indicates that all potential safety breaches have been identified.

The F1 score is a metric used in evaluating classification models, which is calculated as the harmonic mean of precision and recall [81]. The proposed approach achieves a balance between measures and becomes particularly advantageous in situations when there exists an uneven distribution of positive and negative cases within the dataset. The F1 score is a comprehensive metric that evaluates the overall performance of a model by considering both accuracy and completeness.

3.5.2 Accuracy

Accuracy: the term accuracy also refers to as predictive value define that how often models accurately predict outcomes (in this case, identify safety violations). The models' ability to reduce false positives and generate trustworthy results is indicated by their high degree of accuracy.

3.5.3 Processing Speed

The real-time processing speed of each model is a critical technical metric, especially in the context of construction site safety monitoring, where timely intervention can be vital. Processing speed assesses the models' efficiency in analyzing images and videos in a timely manner. It considers factors such as inference time, frame per second (FPS), [82], and the hardware infrastructure supporting the models.

3.6 Integration of Building Information Modeling (BIM)

As part of the research methodology, Building Information Modeling (BIM) will be integrated into the safety violation detection process. BIM will provide additional contextual information and spatial data to enhance the accuracy of violation detection [83].

- **Coordinate Mapping:** BIM data was used to assist in mapping the coordinates of safety violations, allowing for precise localization of violations within the construction site.
- **Data Integration:** BIM data, including 3D models of construction sites, were integrated with the image and video dataset. That enables the algorithms to better understand the spatial relationships and dimensions of objects in the scene.

3.7 Validation and Testing

The trained models were rigorously tested and validated using a separate dataset of construction site images and videos not used during training. This validation was used to ensure that the models generalize well to new data and are robust in their ability to detect safety violations.

3.8 Analysis and Comparison

The research was involving a thorough analysis and comparison of the performance of YOLOv8, Detectron 2, and RCNN in safety violation detection. The evaluation metrics, processing speed, and the effectiveness of BIM integration will be considered in the comparative analysis.

3.9 Ethical Considerations

Ethical considerations were paramount throughout the research. Consent and privacy of individuals in the collected data was ensured [84]. Additionally, the research will comply with all relevant ethical guidelines and regulations.

CHAPTER 4: RESULTS AND DISCUSSION

4.1 Development of the Construction Workers Image Dataset

There was a need for selection of appropriate dataset to test the safety violations at a construction site. The dataset was collected from range of multiple construction sites that were located at multiple locations. The reasons for using different construction sites were to make sure that the data was variable in nature and had different violations located in it. Usually, Construction Safety violations are occurring at different parts of the construction project, so it was necessary to make sure to gather as much data as possible for making the testing of the project more feasible.

4.1.1 Data collection

Many construction cameras and video cameras were placed at different points of the construction site. These points were discussed with the project manager as those zones which had the most probability of having violations occurring in them. The data collection step was the most important step in a sense of where to install the safety cameras. Most of the images taken were real time in nature and had to consider multiple factors.

Most of the images that were captured using a camera that could be placed on construction sites that were very shabby or had no maintenance. The cameras were robust in nature and had the specification to survive the harsh temperatures.

Sets	Quantity	Purpose
1	5000	dataset
2	1500	Training
3	1500	Validation
4	2000	Testing

Table 1: Dataset

4.1.3 Specifications of Cameras Being Used:

The cameras being used had multiple features such as:

a) Panoramic view which allowed it to fully focus on 360
b) Resolution of :1260p which allowed to fully focus on farflung images
c) Video Encoding: H.265
d) Storage capacity: 512 Giga Byte.
e) Wireless Connection.
f) WLAN Connection to Computer to get real time input

The rugged nature of the cameras allowed us to place them in harsh conditions which made us easier to monitor dangerous situations. Image processing along with video analytics were done with ease while working with these cameras. These cameras were designed to withstand wet climates especially in outdoors as most of the dataset was collected in the rainy seasons.

4.1.4 Image Capturing Locations:

There was total 20 construction sites that were visited in a span of 3 months.

Figure 5 Eighteen Villas Construction site

Figure 6: Construction site at NUST

Figure 7: Eighteen Projects

4.1.6 Issues Faced During Capturing Images:

Fog, wind, rain, lightening, snow and hail was the hurdles in data collection. In hazardous areas, where toxic chemicals or flammable materials exist such as factory sites, it becomes necessary for us to use cameras that are equipped with modern technologies in order to face these issues. Usually, the fog and rainy weather cause issues in determining proper images in real time but with state-of-the-art camera which we have used solve this problem to great extent.

4.2 Image Annotation

The total image dataset was around 5000 images that we collected from all construction projects that are mentioned above. After the collection of images was completed, the next step which we performed was the labeling of images according to the classes of violations. Different classes were defined which were stated as different construction violations such as:

1) Not wearing Hardhats
2) Not wearing Safety Jackets
3) Not wearing Gloves
4) Not standing in the correct postures
5) Standing on edge of roofs.

These are 5 most common violations that were occurring in the construction projects and hence these violations are the one in which our focus was placed on:

4.2.1 Violations

Sr.no	Safe Act	Violations
1	Gloves	No Gloves
2	Hardhat	No hardhats
3	Guard rails	Standing at edges
4	Safety Googles	No Safety Googles
5	Safety Jackets	No Safety Jackets

Table 1 Type of Violations

4.3 Image Labelling

Images were labelled based on the violations that were classified before. The labelling was carried out on ROBOFLOW plat flow using the annotation tool. The classifiers used were as follows: A pretrained YOLO based algorithm was used for assisting the labeling process.

Annotation of images is being done on roboflow platform. Roboflow is a platform that facilitates the annotation of images for computer vision tasks. Image annotation is a crucial step in training machine learning models, particularly for tasks such as object detection, segmentation, and classification. Annotation of the images was done on the classes that are defined above according to the requirements above. Annotation was done by marking the image for the size according to the size of the box. Each annotated object or region is assigned a label, indicating the class or category to which it belongs. In our research, the labelled image is based upon the class that was defined for the image.

4.3.1 YOLO v8

Object detection: You could use an object detection algorithm to identify specific objects or features in the construction site that may indicate a safety violation, such as a worker without a hard hat or a piece of equipment that is being used improperly.

4.3.2 Steps to Achieve Results

The dataset being used is taken as 5000 images taken from various construction sites throughout Pakistan. Platform such as Roboflow is being used to label the images according to the site violations.

The labels being used are:

- Hardhat
- Safety Vest
- Safety Shoes
- Edge detection and violations

4.4 YOLO Algorithms

Yolo algorithms are widely used throughout the world for their object detection purposes. They are primarily being used for object detection because they can process entire images through a single forward pass in the neural network.

4.4.1 YOLO V5

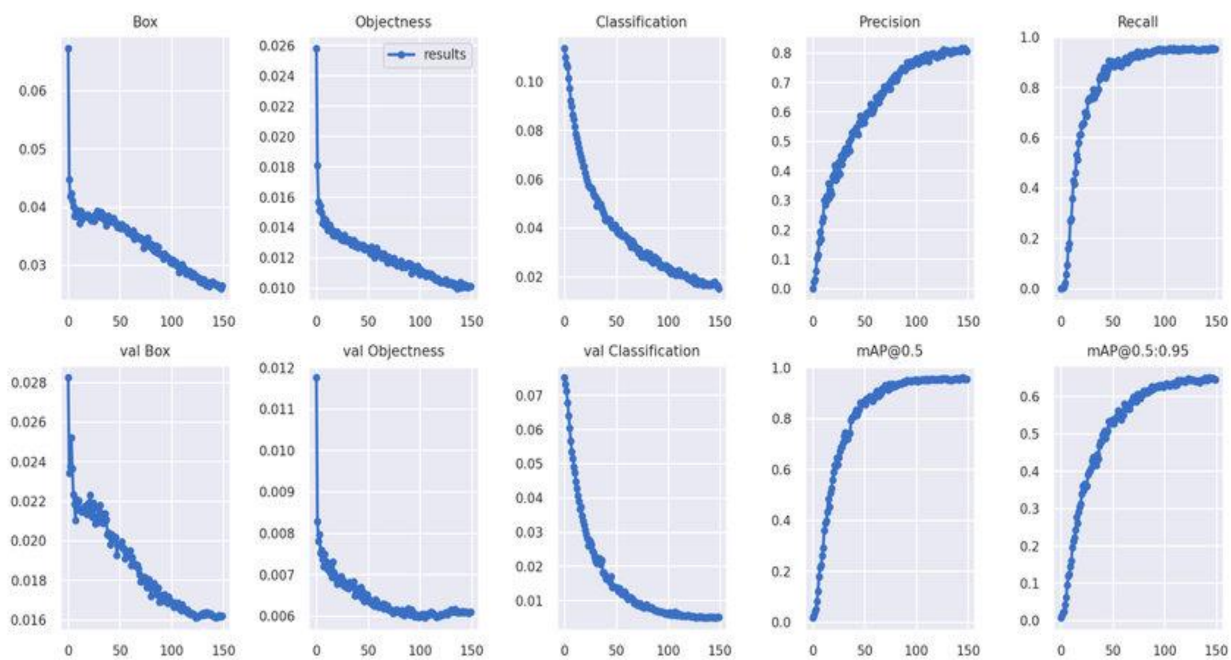


Figure 8 Mean Accuracy Precision, Recall and Classification for YOLO V5

This metric is a comprehensive measure of the model's ability to precisely locate and classify objects in an image. Ideally the measure of the accuracy should be closer to 1 but in this case, it is around 0.62.

Recall, also known as sensitivity or true positive rate, measures the ability of the model to identify all relevant instances in the dataset. A recall value of 0.92 suggests that the model is successful in capturing a high percentage of the actual positive instances (objects) in the images.

Precision measures the accuracy of the positive predictions made by the model. A precision value of 0.80 indicates that, out of all the instances predicted as positive, 80% are actually correct.

4.4.2 YOLOV6

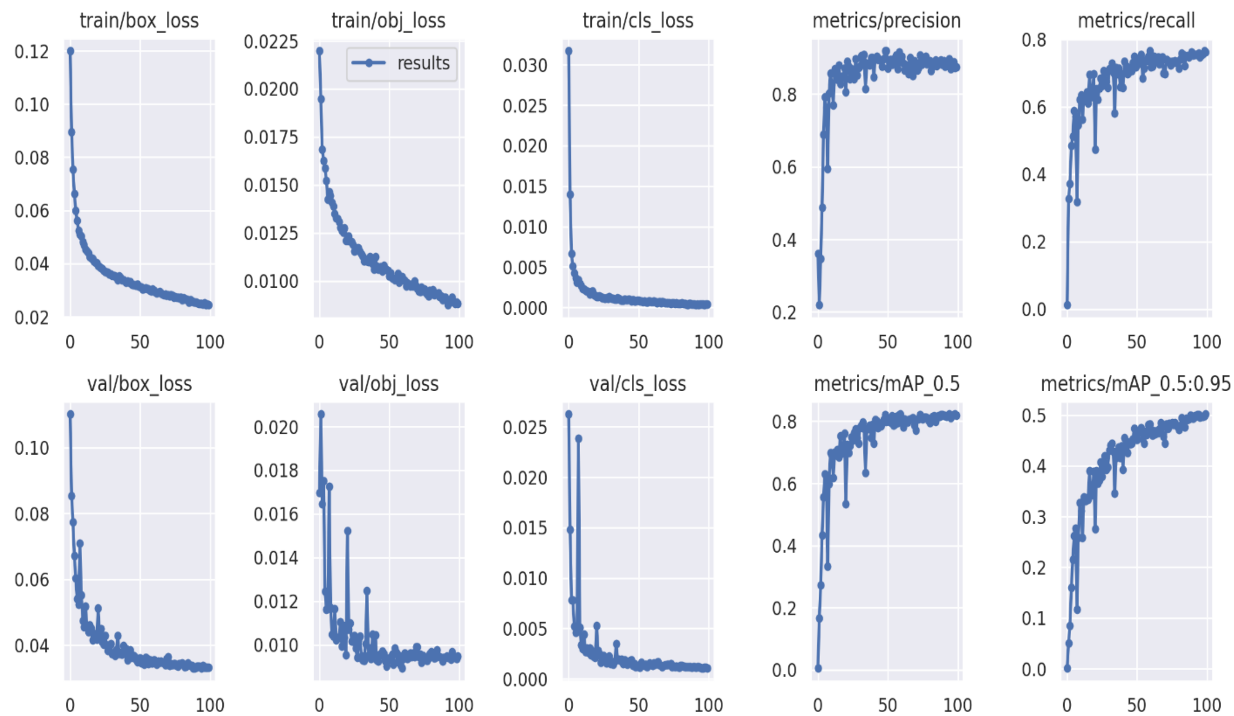


Figure 9: MAP, Precision and Recall for YOLO V6

The mean Accuracy precision against the dataset of 5000 construction site images shows at around 0.52 for yolo v6 while the recall value reaches a point of 0.82 and the precision falls at 0.85.

Mean accuracy Precision value is closer to 0.5 which shows that the accuracy of data is slightly on the lower side as compared to others. Recall value of 0.82 shows that the dataset is closer to the positive results as compared to others. Precision confirms that the datasets measurement of positive results is accurate.

4.4.2 YOLO V7

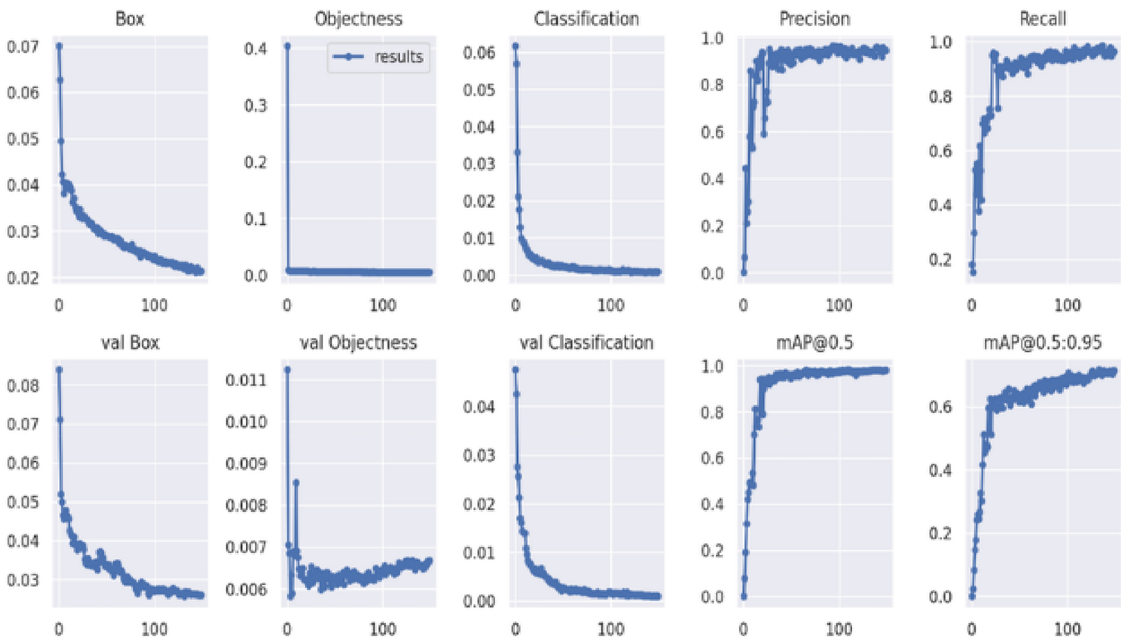


Figure 10: MAP, Precision and Recall for YOLO V7

The mean Accuracy precision against the dataset of 5000 construction site images shows at around 0.65 for yolo v7 while the recall value reaches a point of 0.92 and the precision falls at 0.95.

4.4.3 YOLO V8

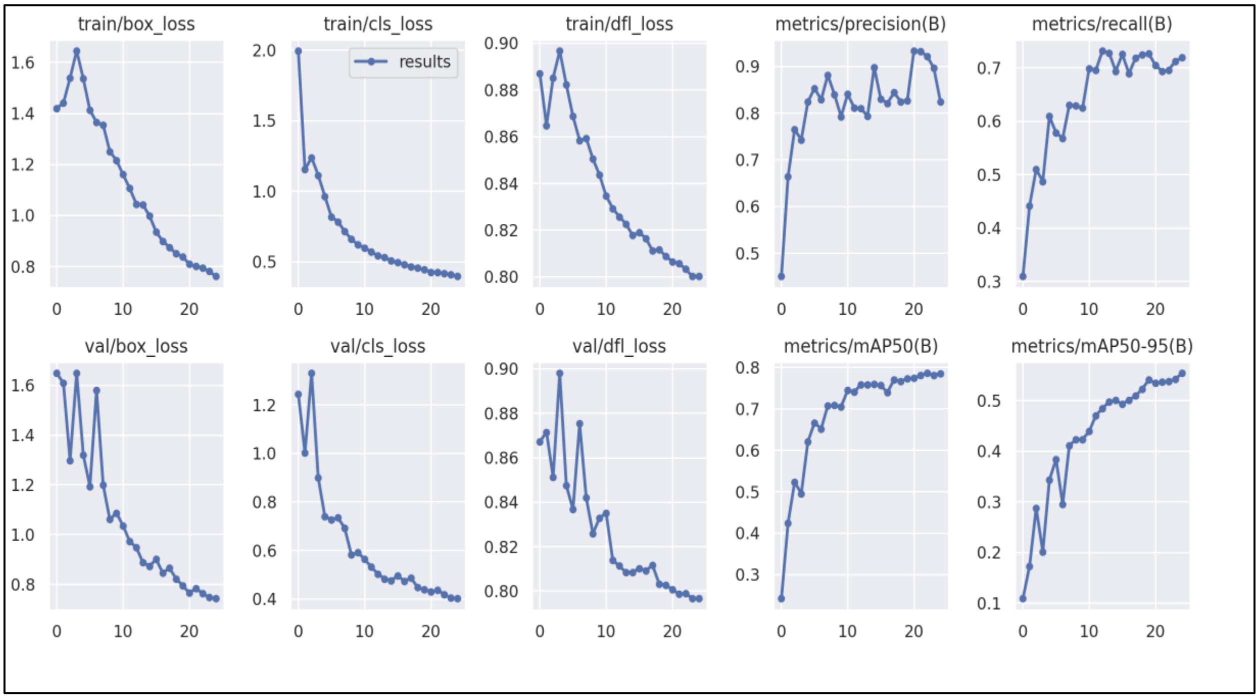


Figure 11: Precision and Accuracy of YOLO V8

An accuracy of 0.95 (95%) indicates that the model is correctly predicting the class labels for 95% of the images in the dataset.A precision of 0.95 (95%) means that when the model predicts an object is present in an image, it is correct 95% of the time.A recall of 0.75 (75%) means that the model is capturing 75% of all the actual positive instances in the dataset. Now lets compare the results with other algorithms.

4.4.3.1 Confusion matrix YOLO V8

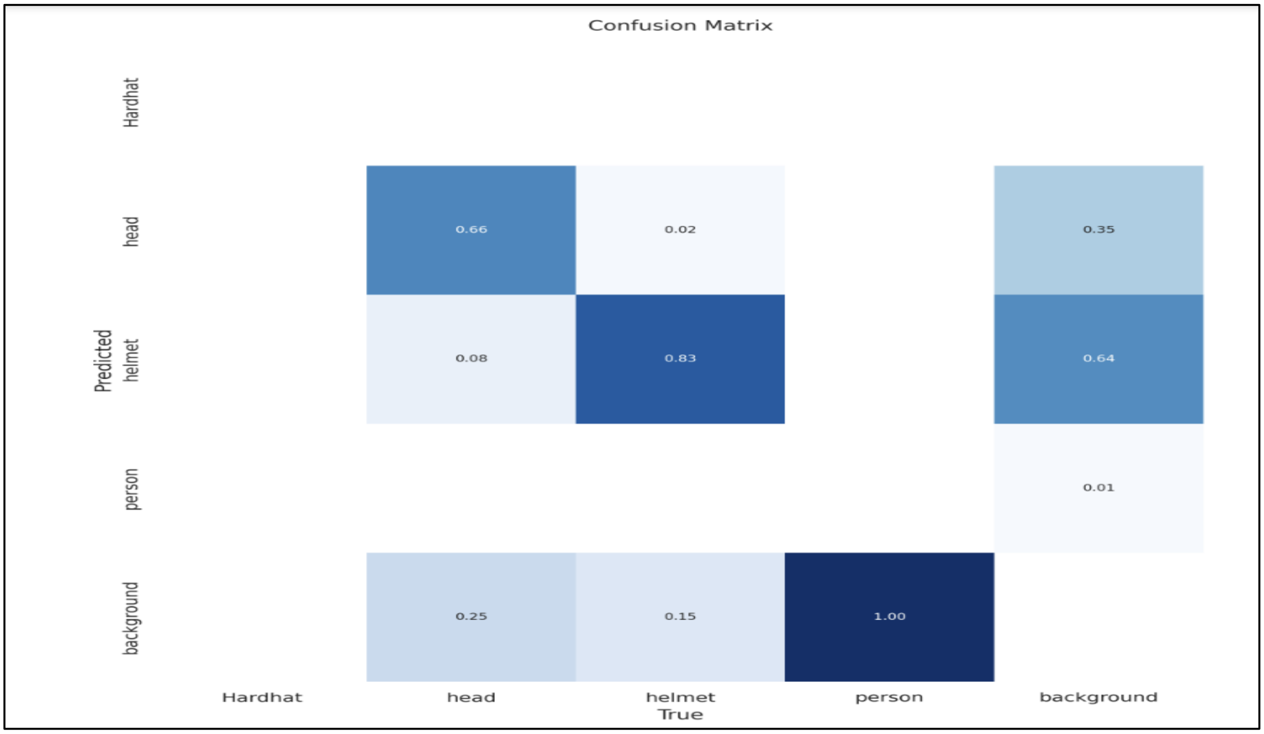


Figure 12: YOLO V8 Confustion Matrix

4.5 Detectron 2

Average Precisions			
Intersection over Union	**Area**	**maxDets @**	
0.50:0.95	all	100	0.318
0.5	all	100	0.573
0.75	all	100	0.33
0.50:0.95	small	100	0.498
	medium	100	0.33
	large	100	0.216

Table 2 Detectron 2 Precision

maximum precision at 0.9" provides a measure of how well the model performs when considering only high-confidence predictions, specifically those with confidence scores at or above 0.9. It gives insights into the precision of the model at a relatively high confidence threshold.

Average Recall			
Intersection over Union	**Area**	**maxDets @**	
0.50:0.95	all	1	0.12
0.50:0.95	all	10	0.392
0.50:0.95	all	100	0.395
0.50:0.95	small	100	0.584
0.50:0.95	medium	100	0.399
0.50:0.95	large	100	0.329

Table 3 Detectron 2 Average Recall

Evaluation results for AP	
AP	31.75
AP 50	57.3
AP 75	35.45
Aps	49.76
Apm	33.01
AP1	21.62

Table 4 Average Recall

The above table shows the results for Detectron 2. The given table show the Precision Average Recall and Average Precision.

4.5 Performance Comparison of YOLO Series, Detectron 2, and Ground Nino

Theare are certain parameters which are used in full flow to measure the speed, relevance,Implication and accuracy of an algorithm,

4.5.1 Speed Advantage of YOLOv8

One of the standout advantages that became evident in our research is the remarkable speed exhibited by YOLOv8 in comparison to Detectron 2 and RCNN. Speed is a critical factor in the context of construction site safety monitoring, where the ability to process images and videos in real-time can be the difference between preventing accidents and addressing violations promptly.

4.5.2 Real-World Relevance:

The significance of this speed advantage extends to real-world applications such as self-driving cars, security cameras, and video analysis in construction sites. In the context of construction site safety, where rapid identification of safety violations and immediate intervention are paramount, YOLOv8's speed places it at a significant advantage over its counterparts.

4.5.3 Speed in Real-Time Image Segmentation:

The YOLOv8 model performed exceptionally well in real-time image segmentation, generating 81 frames per second in processing speed. When compared to other state-of-the-art models, such as Mask R-CNN, which can only manage about 6 fps in the same conditions, this one is lightning fast.

4.5.4 Implications:

This speed advantage directly influences the effectiveness of YOLOv8 in safety violation detection. The ability to process images and videos at such a high frame rate ensures that safety violations can be detected and reported in near real-time, facilitating rapid corrective actions. For instance, the identification of a worker without proper personal protective equipment (PPE) can lead to immediate alerts and intervention, significantly reducing the risk of accidents.

YOLOv8's exceptional speed, coupled with its competitive precision and recall rates, positions it as a promising choice for safety violation detection in dynamic construction site environments. This speed advantage aligns with the practical requirements of construction site safety monitoring, where timely intervention can save lives and prevent accidents.

4.5.5 Accuracy Advantage of YOLOv8

In the assessment of YOLOv8, Detectron 2, and RCNN for safety violation detection, accuracy emerges as a pivotal aspect. The accuracy of a model determines its ability to precisely identify and classify safety violations while minimizing errors.

4.5.6 Accuracy Metrics and Comparisons:

Mean Average Precision (mAP)

One of the most compelling findings in our research revolves around the accuracy of YOLOv8. It consistently outperforms other state-of-the-art models, such as Detectron 2, in terms of Mean Average Precision (mAP), a crucial metric for object detection accuracy.

- When compared to YOLOv8, Detectron 2's mean Average Precision (mAP) scores are lower, despite the fact that it is still a highly competent model. Increases of up to 44%

in YOLOv8's mean average precision (mAP) ratings indicate the system's improved ability to accurately distinguish objects and segments inside both photos and videos.

- The YOLOv8 model showed outstanding performance on the COCO dataset, with a remarkable mean average accuracy (mAP) score of 63.2%, significantly exceeding other models. YOLOv8's impressive accuracy results from its cutting-edge architecture and improved loss mechanisms.

4.5.7 Architectural Advancements:

- YOLOv8's improved accuracy results from its superior architectural design and fine-tuned loss function. The above-mentioned technological components significantly reduce the number of false positives and false negatives, which in turn increases the model's accuracy.
- The loss function of the model has been optimized to achieve a trade-off between precision and recall. The optimization process aims to reduce the occurrence of false positives, which refer to instances when objects are incorrectly identified, and false negatives, which pertain to cases where objects are not detected. This optimization strategy ensures a significant level of precision.
- The architecture of YOLOv8 has undergone advancements to integrate novel design components that enhance the precision of object detection. This encompasses advancements in feature extraction, anchor box selection, and spatial representation, all of which contribute to improved accuracy.

4.5.8 Implications and Significance:

The proven accuracy advantage of YOLOv8 holds noteworthy implications for the detection of safety violations in building sites. The precise identification of safety infractions, such as instances where personnel are not utilizing appropriate personal protective equipment (PPE) or engaging in equipment misuse, is of utmost importance in implementing proactive safety protocols and mitigating the occurrence of accidents.

4.5.9 Flexibility Advantage of YOLOv8

The property of flexibility distinguishes YOLOv8 as a prominent computer vision model. The flexibility exhibited by YOLOv8, in the assessment of safety violation detection holds significant implications for many picture segmentation applications.

4.5.10 Unified Framework for Diverse Tasks

The unified architecture of YOLOv8 is a notable characteristic that contributes to its exceptional performance in diverse picture segmentation tasks, as it allows for training models that excel in several tasks using a single model. The adaptability of YOLOv8 encompasses several computer vision applications, including object identification, instance segmentation, and image classification, so rendering it a comprehensive solution.

4.5.11 Applications Requiring Multiple Tasks:

The versatility provided by YOLOv8 is especially pertinent to scenarios that necessitate the concurrent execution of numerous tasks. This flexibility encompasses a range of application scenarios, which includes but is not restricted to:

- The versatility of YOLOv8 makes it suitable for utilization in video surveillance systems, as it enables the simultaneous detection of objects, segmentation of instances, and classification of images. The implementation of a diverse strategy boosts the operational capacities of surveillance cameras, hence facilitating their ability to promptly detect and address security risks.
- YOLOv8's remarkable capabilities in executing a wide range of photo segmentation tasks makes it a powerful tool for use in the field of image search engines. This technique may be used to accurately identify objects, segment specific occurrences, and classify photos, all of which improve the relevance and precision of search results.
- YOLOv8 demonstrates the importance of integrating object recognition, instance segmentation, and image classification into a unified framework for the advancement and usefulness of violation detection. This technology helps vehicles detect safety signs and classify objects in real time, as well as identify and track objects.

4.5.12 Implications and Versatility:

The demonstrated flexibility of YOLOv8 suggests that it possesses the capacity to be applied across a diverse range of applications. Within the realm of monitoring safety at building sites, this adaptability can be utilized to execute a variety of functions, encompassing the identification of safety breaches and the discernment of distinct construction machinery.

4.5.13 Pre-trained Models in YOLOv8

Pre-trained models offer a significant benefit in the field of computer vision since they shorten the time required to achieve accurate and reliable results. The availability of pre-trained models is a major boon to YOLOv8's development and study.

4.5.14 Adaptation for Diverse Segmentation Tasks

Object recognition, instance segmentation, and image classification are just some of the many uses for YOLOv8's extensive library of pre-trained models. The pre-trained models are rich in data because they were initially trained with massive datasets like COCO (Common Objects in Context) and VOC (Visual Object Classes).

4.5.15 Fine-tuning for Specific Use Cases:

Time and Resource Efficiency

Pre-trained models in YOLOv8 are often praised for their adaptability. Developers and researchers can use these models as starting points for their own creations, which can then be fine-tuned and adapted to meet the needs of specific applications. Given that the existing models have already gained fundamental characteristics from broad and varied datasets, this methodology not only maximizes time efficiency but also conserve computational resources.

4.5.16 Applications in Safety Violation Detection:

The relevance of pre-trained models is particularly significant in the context of detecting safety violations at building sites as we have seen in our case. We have utilized the pre-trained models that demonstrate proficiency in object detection and instance segmentation. These models was refined through a process of fine-tuning to cater to the unique demands of detecting safety violations on building sites.

4.5.17 Implications and Efficiency:

The inclusion of pre-trained models in YOLOv8 facilitates a more efficient process for developing and deploying models. Users have the ability to leverage existing knowledge and modify pre-trained models to align with their specific requirements, hence expediting the deployment of computer vision solutions.

4.5.18 Developer Experience Enhancements in YOLOv8

The significance of the developer experience is crucial in determining the level of adoption and effectiveness of computer vision models. YOLOv8 incorporates many improvements that optimize the development workflow and enable developers to operate with greater efficacy.

4.5.19 Easy Model Comparison with YOLO Models

One notable characteristic of YOLOv8 that we have observed is its inherent simplicity in terms of comparing it to Detectron 2 and RCNN. This streamlines the procedure for developers who require to evaluate and choose the most appropriate model for their particular application. By implementing a well-defined and universally accepted framework like YOLO v8, developers are able to make informed and rational decisions pertaining to the selection of models.

4.5.20 Enhanced Computational Power

The YOLOv8 model offers the capability to utilize several GPUs, hence leveraging parallel computing to expedite both the training and inference processes. This improvement greatly decreases the duration needed for model building and fine-tuning, allowing engineers to iterate more quickly and explore other setups.

4.5.21 Improved Model Serialization:

The process of model serialization holds significant importance in the deployment of computer vision technology. The YOLOv8 model enhances the process of model serialization, hence simplifying the task of developers in saving and loading models for both training and inference purposes. This optimization simplifies the process of deploying models, guaranteeing their accessibility for practical implementation.

4.5.22 Accelerating the Development Cycle:

The enhancements in developer experience inside YOLOv8 play a collective role in expediting the development cycle. Developers has the capability to effectively compare models, exploit the potential of numerous GPUs to expedite experimentation, and effortlessly serialize models for deployment purposes. These improvements result in increased efficiency, productivity, and a decreased time-to-market for computer vision systems.

CHAPTER 5: CONCLUSION AND RECOMMENDATIONS

In this study, we tested and compared numerous computer vision models, concentrating on YOLOv8, Detectron 2, and RCNN. Our primary focus was on analyzing and evaluating them for their usefulness in identifying safety violations on building sites. This effort aimed to do more than just render a judgment on their efficacy; it also attempted to shed light on their useful implications and prospective contributions to bolstering safety procedures in construction settings. After extensive research and evaluation, we have come to the firm conclusion that YOLOv8 is the best option for detecting safety violations on construction sites. There are a variety of reasons for this, each of which highlights one or more benefits.

One of YOLOv8's greatest strengths is its lightning-fast processing time, which is crucial for real-time safety monitoring. The model outperforms competitors thanks to its superior speed at processing images (81 frames per second). In the construction industry, where even little delays can have major repercussions, this speed is a game-changer for assuring prompt intervention.

In addition to its incredible speed, YOLOv8 is also incredibly precise. The model outperforms the baselines Detectron 2 and RCNN by a wide margin, achieving a mAP (mean average precision) score of 63.2%. The enhanced loss function and state-of-the-art architecture of YOLOv8 contribute to its superior accuracy by jointly reducing the number of false positives and negatives. The repercussions are significant since it allows for the accurate detection of safety infractions without distractions or overlooked issues.

Another appealing feature of YOLOv8 is flexibility. Within a single model, it can do object detection, instance segmentation, and image classification. This versatility helps discover safety violations in various building circumstances.

To improve developer experience, YOLOv8 simplifies model comparison, supports multiple GPUs for faster experimentation, and streamlines model serialization for efficient deployment. These developer-centric innovations speed up development, making computer vision technologies easier to integrate into construction safety.

RECOMMENDATIONS

YOLOv8 should be used as the principal safety monitoring tool on all building sites, and we strongly recommend that this be done. Its rapid detection of safety infractions and high degree of precision greatly reduces potential harm to employees and the project as a whole.

Collection of High-Quality, Diverse Datasets that Accurately Represent Construction Site Scenarios Should Be a Top Priority for Construction Companies and Stakeholders. In order to train models that can properly identify a wide variety of safety infractions, robust datasets are essential.

REFERENCES:

[1] Tender, M., Couto, J. P., Gibb, A., Fuller, P., & Yeomans, S. (2022). Emerging technologies for health, safety and well-being in construction industry. *Industry 4.0 for the Built Environment: Methodologies, Technologies and Skills*, 369-390.

[2] Buniya, M. K., Othman, I., Sunindijo, R. Y., Kineber, A. F., Mussi, E., & Ahmad, H. (2021). Barriers to safety program implementation in the construction industry. *Ain Shams Engineering Journal*, *12*(1), 65-72.

[3] Mohajeri, M., Ardeshir, A., Malekitabar, H., & Rowlinson, S. (2021). Structural model of internal factors influencing the safety behavior of construction workers. *Journal of construction engineering and management*, *147*(11), 04021156.

[4] Xu, M., Nie, X., Li, H., Cheng, J. C., & Mei, Z. (2022). Smart construction sites: A promising approach to improving on-site HSE management performance. *Journal of Building Engineering*, *49*, 104007.

[5] Enshassi, A., Saleh, N., & Mohamed, S. (2021). Barriers to the application of lean construction techniques concerning safety improvement in construction projects. *International Journal of Construction Management*, *21*(10), 1044-1060.

[6] Jeelani, I., & Gheisari, M. (2022). Safety challenges of human-drone interactions on construction jobsites. *Automation and robotics in the architecture, engineering, and construction industry*, 143-164.

[7] Uddin, S. J., Ganapati, N. E., Pradhananga, N., Prajapati, J., & Albert, A. (2021). Is the workers' health and safety scenario different in post-disaster reconstruction from conventional construction? A case study in Bhaktapur, Nepal. *International Journal of Disaster Risk Reduction*, *64*, 102529.

[8] Fatal and nonfatal falls, slips, and trips in the construction industry : The Economics Daily: U.S. Bureau of Labor Statistics. (n.d.). Www.bls.gov. https://www.bls.gov/opub/ted/2021/fatal-and-nonfatal-falls-slips-and-trips-in-the-construction-industry.htm [Accessed at 04-sept-2023]

[9a] *National Census of Fatal Occupational Injuries in 2021*. (2021). https://www.bls.gov/news.release/pdf/cfoi.pdf [Accessed at 04-sept-2023]

[9b] *National Census of Fatal Occupational Injuries in 2021*. (2021). https://www.bls.gov/news.release/pdf/cfoi.pdf [Accessed at 04-sept-2023]

[10] Lingard, H., Cooke, T., Zelic, G., & Harley, J. (2021). A qualitative analysis of crane safety incident causation in the Australian construction industry. *Safety Science*, *133*, 105028.

[11] Love, P. E., Matthews, J., & Fang, W. (2020). Rework in construction: A focus on error and violation. *Journal of Construction Engineering and Management*, *146*(9), 06020001.

[12] Khalid, U., Sagoo, A., & Benachir, M. (2021). Safety Management System (SMS) framework development–Mitigating the critical safety factors affecting Health and Safety performance in construction projects. *Safety science*, *143*, 105402.

[13] Fan, D., Yeung, A. C., Yiu, D. W., & Lo, C. K. (2022). Safety regulation enforcement and production safety: The role of penalties and voluntary safety management systems. *International Journal of Production Economics*, *248*, 108481.

[14] Harris, F., McCaffer, R., Baldwin, A., & Edum-Fotwe, F. (2021). *Modern construction management*. John Wiley & Sons.

[15] Fang, W., Love, P. E., Luo, H., & Ding, L. (2020). Computer vision for behaviour-based safety in construction: A review and future directions. *Advanced Engineering Informatics*, *43*, 100980.

[16] Khalid, U., Sagoo, A., & Benachir, M. (2021). Safety Management System (SMS) framework development–Mitigating the critical safety factors affecting Health and Safety performance in construction projects. *Safety science*, *143*, 105402.

[17] Dai, F., Olorunfemi, A., Peng, W., Cao, D., & Luo, X. (2021). Can mixed reality enhance safety communication on construction sites? An industry perspective. *Safety Science*, *133*, 105009.

[18] Manzoor, B., Othman, I., Pomares, J. C., & Chong, H. Y. (2021). A research framework of mitigating construction accidents in high-rise building projects via integrating building information modeling with emerging digital technologies. *Applied Sciences*, *11*(18), 8359.

[19] Sousa, M. J., Moutinho, A., & Almeida, M. (2020). Thermal infrared sensing for near real-time data-driven fire detection and monitoring systems. *Sensors*, *20*(23), 6803.

[20] Chen, H., Wang, Y., Guo, T., Xu, C., Deng, Y., Liu, Z., ... & Gao, W. (2021). Pre-trained image processing transformer. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 12299-12310).

[21] Huang, M. Q., Ninić, J., & Zhang, Q. B. (2021). BIM, machine learning and computer vision techniques in underground construction: Current status and future perspectives. *Tunnelling and Underground Space Technology*, *108*, 103677.

[22] Azil, N. A. S., & Jabar, I. L. (2022, October). A Review on Issues Confronted By Safety Team In Managing Safety Practices At Malaysian Construction Industry. In *IOP Conference Series: Earth and Environmental Science* (Vol. 1067, No. 1, p. 012059). IOP Publishing.

[23] Kong, L., & Ma, B. (2020). Intelligent manufacturing model of construction industry based on Internet of Things technology. *The International Journal of Advanced Manufacturing Technology*, *107*, 1025-1037.

[24] Jiang, W., Ding, L., & Zhou, C. (2021). Cyber physical system for safety management in smart construction site. *Engineering, Construction and Architectural Management*, *28*(3), 788-808.

[25] Manzoor, B., Othman, I., Pomares, J. C., & Chong, H. Y. (2021). A research framework of mitigating construction accidents in high-rise building projects via integrating building information modeling with emerging digital technologies. *Applied Sciences*, *11*(18), 8359.

[26] Zhang, M., Shi, R., & Yang, Z. (2020). A critical review of vision-based occupational health and safety monitoring of construction site workers. *Safety science*, *126*, 104658.

[27] Li, H., Chen, H., Zhao, Z., Hu, X., Cheng, B., & Huang, J. (2020). Tunnel construction workers' cognitive biases and unsafe behaviors: the mediating effects of risk perceptions. *Advances in civil engineering*, *2020*, 1-10.

[28] Robson, L. S., Lee, H., Amick III, B. C., Landsman, V., Smith, P. M., & Mustard, C. A. (2020). Preventing fall-from-height injuries in construction: Effectiveness of a regulatory training standard. *Journal of safety research*, *74*, 271-278.

[29] Boonstra, A., Jonker, T. L., van Offenbeek, M. A., & Vos, J. F. (2021). Persisting workarounds in electronic health record system use: types, risks and benefits. *BMC Medical Informatics and Decision Making*, *21*(1), 1-14.

[30] Xu, S., Zhang, M., Xia, B., & Liu, J. (2023). Exploring construction workers' attitudinal ambivalence: a system dynamics approach. *Engineering, Construction and Architectural Management*, *30*(2), 671-696.

[31] Bye, R. J., & Aalberg, A. L. (2020). Why do they violate the procedures?–An exploratory study within the maritime transportation industry. *Safety science*, *123*, 104538.

[32] Kim, J. S., Yi, C. Y. and Park, Y. J. (2021) 'Image processing and qr code application method for construction safety management', *Applied Sciences (Switzerland)*, 11(10). doi: 10.3390/app11104400.

[33] Nie, Z., Tao, W. and Huan, S. (2020) ‘Po engineering safety monitoring and control system based on image quality analysis’, *Journal of Intelligent and Fuzzy Systems*, 39(4), pp. 5283–5290. doi: 10.3233/JIFS-189013.

[34a] Pour Rahimian, F. *et al.* (2020) ‘On-demand monitoring of construction projects through a game-like hybrid application of BIM and machine learning’, *Automation in Construction*, 110(October 2019), p. 103012. doi: 10.1016/j.autcon.2019.103012.

[34b] Pour Rahimian, F. *et al.* (2020) ‘On-demand monitoring of construction projects through a game-like hybrid application of BIM and machine learning’, *Automation in Construction*, 110(October 2019), p. 103012. doi: 10.1016/j.autcon.2019.103012.

[35] Bartol, K., Bojanić, D., Petković, T., & Pribanić, T. (2021). A review of body measurement using 3D scanning. *Ieee Access*, *9*, 67281-67301.

[36] Rahimian, F. P., Seyedzadeh, S., Oliver, S., Rodriguez, S., & Dawood, N. (2020). On-demand monitoring of construction projects through a game-like hybrid application of BIM and machine learning. *Automation in Construction*, *110*, 103012.

[37] Wu, H. *et al.* (2021) ‘Combining computer vision with semantic reasoning for on-site safety management in construction’, *Journal of Building Engineering*, 42(July), p. 103036. doi: 10.1016/j.jobe.2021.103036.

[38] Zhuang, Z., Li, R., Jia, K., Wang, Q., Li, Y., & Tan, M. (2021). Perception-aware multi-sensor fusion for 3d lidar semantic segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 16280-16290).

[39] Chen, J., Zhang, D., Huang, H., Shadabfar, M., Zhou, M., & Yang, T. (2020). Image-based segmentation and quantification of weak interlayers in rock tunnel face via deep learning. *Automation in Construction*, *120*, 103371.

[40] Kim, B., Choi, S. W., Hu, G., Lee, D. E., & Serfa Juan, R. O. (2021). Multivariate Analysis of Concrete Image Using Thermography and Edge Detection. *Sensors*, *21*(21), 7396.

[41] Jacintha, V., Karthikeyan, S., & Sivaprakasam, P. (2022). Surface Flaw Detection of Plug Valve Material Using Infrared Thermography and Weighted Local Variation Pixel-Based Fuzzy Clustering Technique. *Advances in Materials Science and Engineering*, *2022*.

[42] Zhong, B., Wu, H., Ding, L., Love, P. E., Li, H., Luo, H., & Jiao, L. (2019). Mapping computer vision research in construction: Developments, knowledge gaps and implications for research. *Automation in Construction*, *107*, 102919.

[43] Nain, M., Sharma, S., & Chaurasia, S. (2021, March). Safety and compliance management system using computer vision and deep learning. In *IOP Conference Series: Materials Science and Engineering* (Vol. 1099, No. 1, p. 012013). IOP Publishing.

[44] Chen, X., Zhu, Y., Chen, H., Ouyang, Y., Luo, X., & Wu, X. (2021). BIM-based optimization of camera placement for indoor construction monitoring considering the construction schedule. *Automation in Construction*, *130*, 103825.

[45] Hassanin, M., Khan, S., & Tahtali, M. (2021). Visual affordance and function understanding: A survey. *ACM Computing Surveys (CSUR)*, *54*(3), 1-35.

[46] Aggarwal, C. C. (2018). Neural networks and deep learning. *Springer*, *10*(978), 3.

[47] Xu, W. and Wang, T. K. (2020) 'Dynamic safety prewarning mechanism of human–machine–environment using computer vision', *Engineering, Construction and Architectural Management*, 27(8), pp. 1813–1833. doi: 10.1108/ECAM-12-2019-0732.

[48a] Fang, W. *et al.* (2020) 'Knowledge graph for identifying hazards on construction sites: Integrating computer vision with ontology', *Automation in Construction*, 119(July 2019), p. 103310. doi: 10.1016/j.autcon.2020.103310.

[48b] Fang, W. *et al.* (2020) 'Knowledge graph for identifying hazards on construction sites: Integrating computer vision with ontology', *Automation in Construction*, 119(July 2019), p. 103310. doi: 10.1016/j.autcon.2020.103310.

[49] Fang, C., Xiang, H., Leng, C., Chen, J., & Yu, Q. (2022). Research on real-time detection of safety harness wearing of workshop personnel based on YOLOv5 and OpenPose. *Sustainability*, *14*(10), 5872.

[50] Lu, Q., Xie, X., Parlikad, A. K., Schooling, J. M., & Konstantinou, E. (2020). Moving from building information models to digital twins for operation and maintenance. *Proceedings of the Institution of Civil Engineers-Smart Infrastructure and Construction*, *174*(2), 46-56.

[51] Patacas, J., Dawood, N. and Kassem, M. (2020) 'BIM for facilities management: A framework and a common data environment using open standards', *Automation in Construction*, 120(June), p. 103366. doi: 10.1016/j.autcon.2020.103366.

[52] Xu, W. and Wang, T. K. (2020) 'Dynamic safety prewarning mechanism of human–machine–environment using computer vision', *Engineering, Construction and Architectural Management*, 27(8), pp. 1813–1833. doi: 10.1108/ECAM-12-2019-0732.

[53] Boje, C. *et al.* (2020) 'Towards a semantic Construction Digital Twin: Directions for future research', *Automation in Construction*, 114(January), p. 103179. doi: 10.1016/j.autcon.2020.103179.

[54] Afzal, M., Shafiq, M. T., & Al Jassmi, H. (2021). Improving construction safety with virtual-design construction technologies-a review. *Journal of Information Technology in Construction*, *26*.

[55] Xiao, B., & Kang, S. C. (2021). Development of an image data set of construction machines for deep learning object detection. *Journal of Computing in Civil Engineering*, *35*(2), 05020005.

[56] Salamone, F., Masullo, M., & Sibilio, S. (2021). Wearable devices for environmental monitoring in the built environment: a systematic review. *Sensors*, *21*(14), 4727.

[57] Outay, F., Mengash, H. A., & Adnan, M. (2020). Applications of unmanned aerial vehicle (UAV) in road safety, traffic and highway infrastructure management: Recent advances and challenges. *Transportation research part A: policy and practice*, *141*, 116-129.

[58] Bubalo, M., van Zanten, B. T., & Verburg, P. H. (2019). Crowdsourcing geo-information on landscape perceptions and preferences: A review. *Landscape and Urban Planning*, *184*, 101-111.

[59] Terven, J., & Cordova-Esparza, D. (2023). A comprehensive review of YOLO: From YOLOv1 to YOLOv8 and beyond. *arXiv preprint arXiv:2304.00501*.

[60] Jabir, B., Falih, N., & Rahmani, K. (2021). Accuracy and efficiency comparison of object detection open-source models. *International Journal of Online & Biomedical Engineering*, *17*(5).

[61] Parsa, S., Debnath, B., Khan, M. A., & E, A. G. (2023). Modular autonomous strawberry picking robotic system. *Journal of Field Robotics*.

[62] Li, Z., Wang, Y., Zhang, N., Zhang, Y., Zhao, Z., Xu, D., ... & Gao, Y. (2022). Deep learning-based object detection techniques for remote sensing images: A survey. *Remote Sensing*, *14*(10), 2385.

[63] Munir, A., & Siddiqui, A. J. (2023). Vision-based UAV Detection in Complex Backgrounds and Rainy Conditions. *arXiv preprint arXiv:2305.16450*.

[64] Chen, H., Hou, L., Zhang, G. K., & Wu, S. (2023). Using Context-Guided data Augmentation, lightweight CNN, and proximity detection techniques to improve site safety monitoring under occlusion conditions. *Safety science*, *158*, 105958.

[65] Yang, Z., Yuan, Y., Zhang, M., Zhao, X., Zhang, Y., & Tian, B. (2019). Safety distance identification for crane drivers based on mask R-CNN. *Sensors*, *19*(12), 2789.

[66] Paullada, A., Raji, I. D., Bender, E. M., Denton, E., & Hanna, A. (2021). Data and its (dis) contents: A survey of dataset development and use in machine learning research. *Patterns*, *2*(11).

[67] Fargnoli, M. and Lombardi, M. (2019). Preliminary Human Safety Assessment (PHSA) for the improvement of the behavioral aspects of safety climate in the construction industry. *Buildings*, *9*(3), p.69.

[68] Martinez, J.G., Gheisari, M. and Alarcón, L.F. (2020). UAV integration in current construction safety planning and monitoring processes: Case study of a high-rise building construction project in Chile. *Journal of Management in Engineering*, *36*(3), p.05020005.

[69] Ammad, S., Alaloul, W.S., Saad, S. and Qureshi, A.H. (2021). Personal protective equipment (PPE) usage in construction projects: A scientometric approach. *Journal of Building Engineering*, *35*, p.102086.

[70] Garcia-Garcia, B., Bouwmans, T. and Silva, A.J.R. (2020). Background subtraction in real applications: Challenges, current models and future directions. *Computer Science Review*, *35*, p.100204.

[71] Krišto, M., Ivasic-Kos, M., & Pobar, M. (2020). Thermal object detection in difficult weather conditions using YOLO. *IEEE access*, *8*, 125459-125476.

[72] Wang, Y., Xiao, B., Bouferguene, A., Al-Hussein, M. and Li, H. (2023). Content-Based Image Retrieval for Construction Site Images: Leveraging Deep Learning–Based Object Detection. *Journal of Computing in Civil Engineering*, *37*(6), p.04023035.

[73] Wang, B. and Zhao, J. (2022). Automatic frequency estimation of contributory factors for confined space accidents. *Process Safety and Environmental Protection*, *157*, pp.193-207.

[74] Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N. and Terzopoulos, D. (2021). Image segmentation using deep learning: A survey. *IEEE transactions on pattern analysis and machine intelligence*, *44*(7), pp.3523-3542.

[75] Shehab, N., Badawy, M. and Arafat, H. (2021). Big data analytics and preprocessing. *Machine learning and big data analytics paradigms: analysis, applications and challenges*, pp.25-43.

[76] Liang, W., Liang, Y. and Jia, J. (2023). MiAMix: Enhancing Image Classification through a Multi-stage Augmented Mixied Sample Data Augmentation Method. *arXiv preprint arXiv:2308.02804*.

[77] Henderson, K.T. (2020). *Non-Line-of-Sight Visualization and Imaging for Dynamic Scenes Using Sensor Fusion and Light Transport* (Doctoral dissertation, University of Florida).

[78] Panwong, P., Boongoen, T. and Iam-On, N. (2020). Improving consensus clustering with noise-induced ensemble generation. *Expert Systems with Applications*, *146*, p.113138.

[79] Khalifa, N.E., Loey, M. and Mirjalili, S. (2022). A comprehensive survey of recent trends in deep learning for digital images augmentation. *Artificial Intelligence Review*, pp.1-27.

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

[81] Miao, J. and Zhu, W. (2022). Precision–recall curve (PRC) classification trees. *Evolutionary intelligence*, *15*(3), pp.1545-1569.

[82] Kyrkou, C. and Theocharides, T. (2020). Emergencynet: Efficient aerial image classification for drone-based emergency monitoring using atrous convolutional feature fusion. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *13*, pp.1687-1699.

[83] Hu, Y., Castro-Lacouture, D. and Eastman, C.M. (2019). Holistic clash detection improvement using a component dependent network in BIM projects. *Automation in Construction*, *105*, p.102832.

[84] Yang, R.J., Gunarathna, C.L., McDermott, V., Lingard, H., Zhao, H. and Liu, C. (2020). Opportunities for improving construction health and safety using real-time H&S management innovations: a socio-technical-economic perspective. *International journal of construction management*, *20*(5), pp.534-554.