A Novel Implementation of an AI-based Smart Construction Safety Inspection Protocol in the UAE

Mohammad Z. Shanti¹, Chung-Suk Cho¹, Young-Ji Byon¹, Member, IEEE, Chan Yeob Yeun², Senior Member, IEEE, Tae Yeon Kim¹, Song-Kyoo Kim²,³, Member, IEEE, Ahmed Altunajji²

¹Department of Civil, Infrastructure and Environmental Engineering, Khalifa University of Science and Technology, Abu Dhabi, 127788, UAE
²Department of Electrical Engineering and Computer Science, Khalifa University of Science and Technology, Abu Dhabi 127788, UAE
³School of Applied Sciences, Macao Polytechnic Institute, R. de Luis Gonzaga Gomes, Macao

Corresponding author: Chung-Suk Cho (e-mail: chung.cho@ku.ac.ae).

ABSTRACT The safety of workers at construction sites is one of the most important aspects that should be considered while performing their required tasks. Many rules and regulations have been implemented in the UAE to reduce injuries and fatalities in the jobsites. However, the number of accidents continues to increase. For instance, an accident category of fall–from-height is considered as the top cause of injuries and fatalities. Thus, this paper develops a novel technique that monitors the workers whether they are complying with a safety standard of the Personal Fall Arrest System (PFAS). This paper establishes a real time detection algorithm based on a Convolutional Neural Network (CNN) model in order to detect two main components of the PFAS that are, safety harness and life-line, in addition to a standard safety measure of using a safety helmet. The YOLOv3 algorithm is adopted for a deep learning network used to train the desired model. The model achieved an accuracy rate of 91.26% and around 99% precision. Moreover, the overall recall of the model was 90.2%. The obtained results verify the effectiveness of our proposed model in construction sites to control potential violations and to avoid unnecessary accidents. The main contribution of this paper is to provide an AI-based image detection framework to mitigate the likelihood of fall–from-height accidents.

INDEX TERMS Accidents, CNN, detection, fall from heights, PFAS, YOLOv3

I. INTRODUCTION

The construction sector is one of major industries in the world. However, due to architectural complexities of projects and high rates of progress required to minimize costs, it is also considered as one of the most dangerous industrial sectors. Many workers around the world die and get injured onsite due to the lack of safety measures and knowledge about risks of performing their tasks. Statistics show that around 108,000 workers are killed at construction sites around the world each year [1]. The Occupational Safety and Health Administration (OSHA) reported that one in ten construction workers are injured every year in the U.S.

There are many types of construction-related accidents, top five common site-accidents include falling from heights, struck by an object, electrocutions, caught in between, and slips/falls [2, 3]. These types of accidents tend to occur more in regions where the economy growth is very rapid due to a high number of construction sites, and this is the case in the gulf region. United Arab Emirates (UAE) for instance is one of the most developed countries in the region. The value of construction projects awarded in UAE in 2019 in the first 7 months was more than 700 million dollars as shown in Figure 1 [4]. Moreover, more than 1.64 million people work in the construction industry in the UAE, which is around 20% of the total population [5]. Most of the labor force are expats with different cultures, backgrounds, and varying levels of literacy.

Although UAE has implemented very strict rules regarding the construction safety, many accidents still occur. Lack of safety measures, poor supervision, and lack of communication are among the main reasons for these accidents. Statistics show that around 38% of the total construction-related deaths that have occurred in Dubai, UAE are due to the lack of supervision [2]. Thus, new techniques and technologies can be utilized and implemented in this industry to increase the supervision and awareness of the construction safety among the laborers. One of these techniques is to merge and integrate artificial intelligence in the field of construction safety.
Artificial Intelligence (AI) is a branch in computer science which focuses on developing software or machines which exhibit human intelligences. The main goals of AI include deduction and reasoning, knowledge representation, planning, natural language processing (NLP), learning, perception, and ability to manipulate and move objects [6].

There are mainly five common issues, often bundled in combinations, that can be considered as culprits for the increase in number of accidents at construction sites:
- Insufficient number of qualified authorities to conduct frequent on-site inspections over vast majority of construction projects at different locations
- Failing to comply with safety regulations (non-compliance)
- Lack of awareness regarding the risks associated with specific occupational activities
- Lack of safety management system and poor safety culture
- Inadequate planning and training

This paper aims to mitigate the first issue where the shortage of qualified site inspectors is augmented by adopting an artificial intelligence (AI)-based smart-construction safety-inspection protocol. Since the fall-from-height is a leading cause of serious and fatal injuries for construction workers, violation of the EHS regulations on fall protection, detecting violations of workers working at height without using a proper personal fall arrest system (PFAS) will be the primary focus of AI development and application in this research.

The main contributions of this paper are as follows.
- Introducing a novel technique that monitors workers working at heights by using the power of AI mainly deep learning.
- Preparation of a unique dataset for training the model.
- Training a CNN that detects the desired objects (safety harness, lifeline, and safety helmet).
- Testing and validation of the results on a unique dataset under different conditions.
- Demonstrating that AI can be a solution for reducing accidents on construction sites.

The rest of this paper is divided into multiple sections as follows. Section II discusses the background. Section III presents related works in terms of current developments in AI-based applications of construction safety. Section V discusses the main components of PFAS and the requirements associated with using it. Furthermore, Section IV shows our proposed deep learning algorithm. Section VII contains the testing and validation of the trained model. Section VIII discusses the obtained results. Finally, Section IX concludes this paper.

II. BACKGROUND

The outlook of the construction industry in UAE for the next five years is bright as the investment is expected to increase due to its government’s efforts to diversify the economy along with the recovery in crude-oil prices and rise in non-oil product exports [7]. The report further predicts in a section titled, ‘Construction in the UAE – Key trends and opportunities to 2023’ that the construction industry will rise at a compound annual growth rate of 4.64 per cent during 2019-23 while reaching up to $101.2 billion by 2023 in spending [7].

While the UAE construction industry continues to benefit from no shortage of construction projects and activities, assuring the utmost safety and health of the involved workers is not fully embedded in the culture of the construction industry relative to the large scale of the boom the industry is experiencing [8]. To address this issue, the Abu Dhabi Municipality (ADM) established the emirate’s environment, health, and safety (EHS) regulations prescribing safety requirements for all local construction sites. Accordingly, the regulatory authorities of the ADM frequently conduct enforcement activities, including establishment and implementation of project level EHS management system, on-site inspections and levying penalties where required. Despite efforts from the government level, studies suggest that work-related incidents and accidents leading to injuries and fatalities are still on the rise. The Health Authority Abu Dhabi (HAAD) has reported that occupational injuries accounted for 20% of all injuries in the emirate of Abu Dhabi during 2015, including 90 fatal occupational-related incidents [9]. The HAAD further reported that injuries related to falls and falling objects were the largest cause of occupational injury causing 48% of non-fatal and 64% of fatal occupational injuries [9]. Another study which analyzed the trauma registry data from a large local hospital, in agreement with the findings from the work of Baras et al. [10], the most frequent mechanism of injury was fall-from-height (52%) [11].

Considering a large proportion of the UAE workforce is employed in potentially hazardous occupations in the construction industry as can be seen in Figures 2 and 3, improving employee awareness of health and safety issues while complying to safety regulations is vital to prevent the

FIGURE 1. The values of projects in 2019 in the GCC countries until July 2019 [4].
injuries, fatalities and economic losses associated with common workplace accidents resulting from falls and falling objects. However, it is obvious from the figures that the EHS safety regulation for fall protection for unprotected edges (EHS RI CoP 23.0, Section 3.14.5) [12] is violated at all levels where no measures to protect the employee from falling (i.e., guardrail, safety net, personal fall arrest system) are in place.

With recent advancements in AI technologies, detecting specific image data has been frequently carried out by using deep learning (DL) techniques. Although applications of DL for processing images and video have experienced a tremendous growth, few studies have applied the DL approaches in smart construction inspections. Specifically, there is very few known applications of construction safety using DL-based image smart-recognition for checking any violations at construction sites, e.g., public safety, onsite workers’ safety violations. Adopting the recent advancements of AI technologies, this paper enhances the DL approaches so as to achieve reliable on-site detection and recognition of work-safety violations by directly feeding the raw images data to the established DL structure. YOLOv3 will be the main object detection algorithm that will be used in this research in order to perform the required work-at-heights detections. In the next section of this paper related research works and studies are presented to examine the current status in the AI safety field.

**III. RELATED WORKS**

Construction is one of the major industries in Abu Dhabi. In 2019, the construction sector contributed about 86 billion United Arab Emirates dirhams to the preliminary estimated value of the gross domestic product (GDP) of the emirate of Abu Dhabi [5]. Although this industry contributes a large share of money in the national GDP, a significant number of construction-related accidents still take place regardless of strict rules and regulations implemented by relevant authorities. Therefore, several new methods are being implemented worldwide to monitor and supervise labor workers during the execution of work. One of these methods is using Artificial Intelligence (AI) specifically a deep learning approach to supervise and detect safety violations. An article titled “Computer vision applications in construction safety assurance” examines the use of computer vision and deep learning for monitoring unsafe behaviors and conditions in construction sites. The research was conducted to identify the challenges in using such technology in complex environments such as construction sites. Associated main challenges include the difficulty in obtaining surveillance videos; and suitable and reliable training data sets for training the deep learning models [13].

Another article that demonstrates the importance of AI in construction industry is “Artificial Intelligence Enabled Safety for Construction Sites”. This paper presents a method for the real time detection of unsafe act and unsafe conditions of workers using Artificial Intelligence. The authors implement a computer vision in developing a safety model that is trained by a quantum number of images. The model helps in analyzing the safe and unsafe conditions in a construction site and thereby it will help in reducing the accidents at the construction sites to an extent [6]. Their model was trained with 1000 images, and attained around 90 % of precision. Their model mainly focuses on detecting labor workers without personal protective equipment (PPE).

Another supplementary research that proposes an approach for automatic detection of helmets in construction sites using computer vision and machine learning techniques is “Automatic Detection of Helmet Uses for Construction Safety”. The proposed system has two major parts: one part incorporates frequency domain information of the image with a popular human detection algorithm HOG for human (i.e., construction worker) detection; the other part works for helmet use detection combining color information and Circle Hough Transform (CHT) [14]. However, the developed system in the paper had some flaws in detecting different colors of worker’s helmets, and additionally it had some faults in differentiating a normal cap with the construction helmet.

A relevant study about falling-from-heights was conducted and discussed in an article titled, “Falls from heights: A computer vision-based approach for safety harness detection”. In this article a Faster-R-CNN and a deep CNN model were trained in order to detect workers and safety harness respectively. The CNN model trained to

**FIGURE 2. Construction worker working without proper fall protection system.**

**FIGURE 3. Unsafe work environment with inappropriate fall protection systems in place.**
to detect the safety harness in this research gave a relatively low accuracy. However, it was concluded that with more training datasets the accuracy of the model will increase. Thus, the use of such smart artificial intelligent techniques could decrease the construction accidents and fatalities significantly [15].

A recent paper titled, “Detection of Personal Protective Equipment (PPE) Compliance on Construction Site Using Computer Vision Based Deep Learning Techniques” developed a framework to detect in real-time, the safety compliance of construction workers with respect to PPE [16]. YOLOv3 was their adopted deep learning network that was used in their research, and it proved its ability to perform in a construction site environment with a relatively high accuracy.

Using artificial intelligence for safety is not limited only for the construction sector. Many research works demonstrate the potential of using this technology in other industries; in the article “Improving Train Track Safety using Drones, Computer Vision and Machine Learning” the authors aimed to develop multiple algorithms that implement supervised and semi-supervised learning that accurately analyze whether a track is safe or unsafe based on simulated training data of train tracks. They developed a Convolutional Neural Network that can identify track defects using supervised learning without having to specify a particular algorithm for detecting those defects. The project demonstrated that by using computer vision and machine learning algorithms (artificial intelligence) with training datasets, the track defects can be identified objectively for maintaining safety [17].

In the civil infrastructure sector, more specifically in the field if Intelligent Transportation System (ITS), Artificial Neural Networks (ANNs) have been applied with Kalman filter-based data-fusion and Geographic Information System (GIS) for detecting/monitoring smartphone user’s mode of transportation for various ITS applications including traffic safety and road-condition monitoring [18, 19].

These previous studies and researches provide a strong basis for more advancements in the AI construction safety field. Table 1 above provides the most important findings from the related works discussed in this section.

| Serial No. | Paper’s Name | Targeted Object(s) for Detection | Findings |
|------------|--------------|---------------------------------|----------|
| 1          | Automatic Detection of Helmet Uses for Construction Safety | Safety Helmet | Diverse data set is needed in order to avoid detection errors due to the color and the shape of the objects desired for detection |
| 2          | A computer vision-based approach for safety harness detection | Safety Harness | CNN is a very effective approach in detecting construction objects. However, the best algorithm should be picked in order to avoid low accuracy |
| 3          | Artificial Intelligence Enabled Safety for Construction Sites | Personal Protective Equipments (PPE) | The model should be trained to encounter different light and image conditions |
| 4          | Detection of Personal Protective Equipment (PPE) Compliance on Construction Site Using Computer Vision Based Deep Learning Techniques | Personal Protective Equipments (PPE) | YOLOv3 is a reliable CNN network that can be used in a construction environment |

IV. PROPOSED RESEARCH DESIGN AND ARCHITECTURE

Our proposed research methodology of this paper was developed after taking into considerations of all the findings from existing literature in the previous section. A unique research design was established consisting of three main phases: image collection, machine learning with deep learning, and application of the deep learning algorithm.

A. IMAGE COLLECTION

An image database for Deep Learning using real construction site images and other web sources was developed. Images and recorded videos were collected from construction sites for training, testing, and validation of the deep learning algorithm in phases 2 and 3 respectively. Collected images were similar to photos in Figure 4 where compliance or violation of PFAS requirement is obvious. These collected images and videos were compiled into a database as a training data set and were used for conducting phase 2 with a deep learning model.

FIGURE 4. Database of images/videos showing the PFAS components.
B. MACHINE LEARNING WITH DEEP LEARNING

Deep Learning Algorithm based on the CNN: As for the deep learning, the CNN is one of the most useful and popular neural networks for AI algorithms that recently received attentions from researchers as an effective application involving images for training multi-layer network structure. The CNN can take in an input image, assign importance values according to various aspects/objects in the image and be able to differentiate one from the other. Since analyzing visual imagery is the primary interest of this research, the CNN algorithm, as shown in Figure 5, was adopted in this research so that implementing image-recognitions for detecting the PFAS classes became possible. The main goal of this task was to train a deep learning model with the ability to detect the safety harness, safety helmet, and the life line (PFAS components).

![Convolutional neural network classification process](image)

**FIGURE 5.** Convolutional neural network classification process [20].

C. APPLICATION OF THE DEEP LEARNING ALGORITHM

Testing and Validation of the Deep Learning Algorithm: A supervised learning, which is applied in this project, is a machine learning task of learning a function that maps an input to an output based on example input-output pairs [21]. It inverts a function from labeled training data consisting of a set of training examples [22]. The flow of the supervised learning in two steps is visually depicted in Figure 6. This technique is flexible enough to be adapted in the CNN. Applying the supervised deep leaning algorithm to this research, the training was executed by using a sufficient amount of data extracted from the compiled database. Once the deep learning was completed, a testing phase took place. Unique images were collected and various evaluation tests under different conditions were performed. The model was tested under a normal RGB color condition, grayscale, maximum brightness, dusting effect, and blur effect conditions. After executing the model, the AI spontaneously determined the PFAS components. These results of detections can easily allow the safety officer in charge to react spontaneously to any working-at-height violations in real-life construction sites.

![Graphical description of supervised machine learning process](image)

**FIGURE 6.** Graphical description of supervised machine learning process [23].

V. WORKING AT HEIGHTS REQUIREMENTS

There are three main safety requirements for fall protection that should be available when working at heights (1) Guardrail (2) Safety net (3) Personal Fall Arrest System (PFAS). The guard rail is a form of passive protection, it is the easiest and a basic way to keep the workers safe while achieving the compliance. The main purpose of guardrails is to prevent falls or unauthorized access to certain areas. Moreover, the safety nets are used at high-rise building’s construction sites for preventing accidental fall of people or objects from the site [24]. On the other hand, the PFAS which is the main focus of this research is considered as the most important component in protecting workers at heights. It provides an individual protection for each worker rather than a group protection which is the case in the safety net and guardrail.

The PFAS consists of three main components that are anchoring, body wear (safety harness), and connecting link (lifeline). Figure 7 gives a brief description of the components of the protection system. The Occupational Safety and Health Administration (OSHA) has set certain standards and criterions in order to control and maintain the safety of workers when using the PFAS. Some of these criterions include [25]:

- Anchorages used for attachment of personal fall arrest equipment shall be independent of any anchorage being used to support or suspend platforms.
- When vertical lifelines are used, each employee shall be provided with a separate lifeline.
- The attachment point of a body belt shall be located in the center of the wearer's back.
- Personal fall arrest systems shall be inspected prior to each use for mildew, wear, damage, and other deterioration.
- Personal fall arrest systems and components subjected to impact loading shall be immediately removed from service and shall not be used again for employee protection until inspected and determined by a qualified person to be undamaged and suitable for reuse.
- Training before using personal fall arrest equipment. Each employee shall be trained to understand the application limits of the equipment and proper hook-up, anchoring, and tie-off techniques.
All these strict standards and regulations should be implemented and applied in order to secure the safest working environment when working at heights.

**FIGURE 7. Personal fall arrest system** [26].

**VI. PROPOSED DEEP LEARNING ALGORITHMS**

The main performance indicators of any object detection models are detection accuracy and computational speed. Nowadays, the most common object detection algorithms are mainly based on deep learning models. These detection algorithms can be divided into two main categories: (1) two-stage detection algorithms, which divides the detection problem into two stages. First it generates candidate regions, and then classifies each candidate region. Well known forms of algorithms in this category are the R-CNN algorithms such as Rmnr CNN, Fast Rmnr, CNNQuest, and Faster R-CNN. (2) that are one-stage detection algorithms, which do not need a region-proposal stage. They directly produce the class probability and position coordinate values of the desired objects. The most common algorithms in this category include YOLO (You Only Look Once) and SSD (Single Shot multi-box Detector) [27-30].

**A. YOU ONLY LOOK ONCE (YOLO)**

YOLO is a clever convolutional neural network (CNN) for processing object detections in real-time. The algorithm applies a single neural network to the full image, and then divides the image into regions and predicts bounding boxes with associated probabilities for each region. These bounding boxes are weighted by the predicted probabilities [31]. YOLO divides the input image into an S x S grid. If the center of an object falls into a grid cell, that grid cell is responsible for detecting that object. Each grid cell predicts B bounding boxes and confidence scores for those boxes. These confidence scores reflect how confident the model is whether the box contains a desired object and also how accurate the detection is, that is whether the object it predicts is indeed the correct one or not. Finally, it predicts C conditional class probabilities as shown in Figure 8 [29].

There are several versions and upgrades for YOLO including YOLOv2, YOLOv3, and YOLOv4. YOLOv2 and YOLOv3 were an improvement over YOLO which introduced batch normalization, high-resolution classifiers, convolutions with anchor boxes, multi-scale training, and joint classification and detection. YOLOv3 uses SoftMax based predictions and enables multi-label classifications. It utilizes the Darknet-53 architecture as a feature extractor from images. YOLOv3 is an advanced CNN architecture whose mean AP (Accuracy precision) metric is comparable to other similar architectures like Faster R-CNN and SSD513, but offers a faster prediction rate when compared to other algorithms. The most noticeable feature of YOLOv3 is that it makes predictions at three scales, which are precisely given by down sampling the dimensions of the input image by 32, 16 and 8 respectively as shown in Figure 9. The detection is done by applying 1x1 detection kernels on feature maps of three different sizes at three different places in the network. Detections at different layers help address the issue of detecting small objects, a frequent issue with YOLOv2. The up-sampled layers concatenated with the previous layers help preserve fine-grained features which in turn aids detecting small objects. Moreover, YOLOv3 predicts 10x the number of boxes predicted by YOLOv2. YOLOv3 performs multi-label classification for objects detected in images which is a great upgrade from YOLOv2 [32]. YOLOv3 has proved to be effective at detecting multiple objects at construction sites [16]. Thus, YOLOv3 was chosen as the primary CNN architecture for this research.

**FIGURE 8. YOLO model** [29].

**FIGURE 9. YOLOv3 architecture** [32].

**B. TRAINING FLOW CHART**

The flow chart shown in Figure 10 summarizes the model’s training methods performed in this research. All the mentioned steps will be broadly explained in the upcoming sections and subsections of this paper.
custom data set, since there is no one standard color or shape for these protective equipments. Furthermore, some images were adjusted to have different orientations than the original to increase the diversity of the images. Similarly, the brightness of some images was adjusted to increase the flexibility of the model under different lighting conditions. After adjustments and filtration processes were completed, a total of around 1000 images were acquired for training and validating the model. The prepared custom data set included a sufficient number of images for each class to ensure that the data is well presented for the model and no class is predominant over others. Once the dataset was completed, each and every image was labeled using LabelImg interface. LabelImg is a free, open-source tool for graphically labeling images. It is written in Python and uses QT for its graphical interface. A sample of data labeling is shown in Figure 12. All annotated data was saved in TXT file format.

FIGURE 10. Training flowchart.

C. CUSTOM DATASET FORMATION

One of the most important aspects in developing an accurate deep-learning object-detector are quality and quantity of the data and images that will be used in training the model. The process is considered to be time consuming and very critical in determining the overall performance of the model. In this research the images used to train the model were obtained from two main sources including multiple local construction sites and web-based images as illustrated in Figure 11. Several visits were performed to local construction sites where photos were taken from ongoing works that involve working-from-height scenes. Furthermore, videos of workers performing their tasks on building edges and scaffolding were taken in order to increase the diversity of the training dataset. On the other hand, the process of obtaining images from the web was challenging since the images available online are very random and from different industries. So, a systematic procedure was made in order to select relevant photos. This procedure involved obtaining a bulk of images from the main source which is Google-provided images, and then a filtration/selection process was executed.

The filtration process involved eliminating images with huge watermarks and irrelevant texts. The photos obtained from the web were of two main formats JPG and PNG; any other photos in different formats were converted to either of the desired formats. Moreover, the images obtained or captured while creating the custom data set included mainly three classes: (1) Safety harness (2) Lifeline (3) Safety helmet. The safety helmet was added to the model since the helmet is one of the most important personal protective equipments that protects the workers during their job executions. Therefore, a failure of wearing a helmet is considered as a big failure in the safety system.

The images collected for training the model were taken to consider the construction working conditions. Different types, shapes, and colors of the items were added to the

FIGURE 11. Source of images.

FIGURE 12. Labeling.

D. MODEL TRAINING

The next step after creating a rich dataset is to train the model. In this research the model was trained on Google colab notebook. Colaboratory, or “colab” for short, is a product from a Google research group. Colab allows users...
to write and execute arbitrary python codes through a browser, and is especially well suited to machine learning, data analysis and education. More technically, colab is a hosted Jupyter notebook service that requires no extensive setup procedure to use, while providing a free access to computing resources including GPUs [33].

In order to create a custom YOLOv3 detector, the following components were needed (1) Labeled custom dataset (2) Custom configuration (.cfg) file (3) data.txt and names.txt files. The first step in training the model was to enable GPUs within the colab notebook. This increased the rate of training and the rate at which detections were performed. The second step was to move the custom data set into the Google Cloud VM. The third step in training the model was configuring specific files for training. This step involved properly configuring the custom .cfg file, data.txt, names.txt, and train.txt files. The configuration file (.cfg file) contains information such as how many classes (objects) will be trained in this model and how many iterations the model will perform— the greater number of iterations the more accurate the model; the number of iterations in this research was chosen to be 2000 for each class which means a total of 6000 iterations were performed for this model. Figure 13 shows a cross-entropy plot representing the performance of the model in the training process. As the number of iterations increases, the loss decreases which means the model is becoming more accurate. After the 6000 iterations were performed the final loss value was minimal and had a value of 0.207. Furthermore, a names.txt file was created which included the names of the classes of the model. In this research three classes were added to this file that are safety harness, lifeline, and safety helmet. Similarly, a data.txt file was created and this file included information about the location paths to the names.txt file and the location where the final trained model weights will be saved. Finally, the train.txt file was generated and it contained relative paths to all training images.

After generating all required files for training the model, pre-trained weights for the convolutional layers were downloaded. This step downloads the weights for the convolutional layers of the YOLOv3 network. This step is optional but was carried out in order to increase the accuracy and reduce the time it takes to train the model. Finally, the model training was performed. The model training was completed in around 30 continuous hours. A sample of the results obtained after the model was trained is shown in Figure 14.

![Figure 13. Loss vs Iteration.](image1)

![Figure 14. Model's sample result.](image2)

### E. DETECTIONS USING TENSORFLOW

Once the model was trained on the Google colab platform, image detections were performed as mentioned earlier. However, the Google colab does not offer a permanent platform for making detections, since it has a short-term virtual memory that deletes all stored files after a specific period of time. So, a more relevant software was utilized which is TensorFlow. TensorFlow provides a collection of workflows to develop and train models using Python or JavaScript, and to easily deploy in the cloud, in the browser, or on-device no matter what language is being used [34]. One of the main advantages of using TensorFlow is that it is a free open-source package that contains thousands of machine learning libraries. This great advantage made it easier to develop the required object detector. This was achieved by obtaining a prewritten code for object detection and modifying certain parameters in order to match our research objectives. The most important parameters that was modified in the code are model weights. These weights are the output obtained...
after training the model in colab. They include all the data required to detect the desired classes. Another parameter that was altered was the names of the classes required for detection. Once the weights and the names were downloaded and implemented to the code, the object detector was ready to start detecting new images for validation. 75% of the overall produced dataset was used for training the model and the other 25% was optimized for evaluating and validating its results. The detector was trained to perform detections on both images and videos.

VII. TESTING AND VALIDATION
Before implementing the object detection model, the model should be tested and validated to evaluate its performance. In this research four main parameters were identified to test the performance of the model (1) the accuracy of the detections (2) the confidence in the detections (3) precision of the detections (4) recall. The accuracy of the detections was measured in a way that if the model detected the desired object a binary outcome of 1 (i.e., 100%) accuracy value is recorded. On the other hand, if the object is not detected a value of 0 (i.e., 0%) is recorded. In case of partial detections (e.g., 1 of 2 helmets detected) an average value is recorded. Moreover, the confidence in the detections were directly obtained from the developed application via the TensorFlow platform. The confidence score reflects the probability that an anchor box contains the desired object. These evaluation parameters are further demonstrated in Figure 15. The values in the yellow boxes below show the confidence scores, and since the model detected the desired objects (helmet, harness, lifeline) correctly, each object received a 100% accuracy in this specific situation.

In this research 250 unique images were collected for performing the validation tests. The images used for testing the model were completely different from the images used to train the model. The 250 images used for validation accounted for 25% of the overall created dataset. The model was then validated by testing the model under six different conditions (1) RGB (2) Grayscale (3) Maximum brightness (4) Dusting effect (5) Blur effect (6) Videos.

A. RGB AND GRAYSCALE
Various tests were performed to evaluate the performance of the model under different conditions and circumstances. One of these tests was to evaluate the performance of the model under different color conditions. The testing images were initially evaluated in a full color condition, after that the same images were converted into grayscale state. This can be visualized clearly in Figure 16 along with the detection results. Upon the completion of the testing performed on the discussed conditions. A bar chart was generated showing the difference between the results obtained, and this is shown in Figure 17.

B. HIGH BRIGHTNESS AND DUSTING EFFECT
Another parameter that was modified to evaluate the performance of the object detector was the brightness of the images. The same images that were used earlier in testing the detection model were used to test the new images condition. The brightness for each and every image was maximized in order to test the model under different brightness and light conditions. This test was done in order to evaluate the performance of the model in a very sunny day. Similarly, the images were altered to test if the model could cope under dusty weather conditions. A sample of the detection results for both categories is demonstrated in Figures 18 and 19. A bar chart was generated to illustrate the performance of the model under both environments. This is pictured in Figure 20.
The weather in the gulf region especially UAE is considered to be hot and humid in the summer and moderate in the winter. So, the deep learning object detector has to cope with these weather conditions. For that one of the most important weather conditions was set into action, which is humidity. The testing images were modified into three main categories the 0%, 4%, and 8% blur. These categories resemble the clear normal weather (0%), the average humid weather (4%), and the high humid weather (8%). Figure 21 shows a sample of the three different categories.

It was noticed from the results obtained that the model did not get affected significantly when the parameters were modified. This shows that the trained object detector is consistent in its results regardless of the available conditions. This is demonstrated clearly in the charts generated from the obtained results.

C. BLURRING EFFECT

The weather in the gulf region especially UAE is considered to be hot and humid in the summer and moderate in the winter. So, the deep learning object detector has to cope with these weather conditions. For that one of the most important weather conditions was set into action, which is humidity. The testing images were modiﬁed into three main categories the 0%, 4%, and 8% blur. These categories resemble the clear normal weather (0%), the average humid weather (4%), and the high humid weather (8%). Figure 21 shows a sample of the three different categories.

It was noticed from the results obtained that the model did not get affected significantly when the parameters were modified. This shows that the trained object detector is consistent in its results regardless of the available conditions. This is demonstrated clearly in the charts generated from the obtained results.
D. VIDEO DETECTIONS
Upon completing the detections on static images, detections on several videos were performed to evaluate the performance of the model under dynamic conditions. The videos used were short videos of workers performing their tasks on real construction site. The videos were then uploaded to the TensorFlow model and detections were performed on the videos. To evaluate the results of the detections, the videos were slowed down to a rate of five frames per second (FPS) to record the accuracy and confidence of the detections. Figure 23 shows a screenshot taken from one of the videos that were tested using the trained model.

FIGURE 23. Video screenshot.

Several videos were tested to evaluate and validate the performance of the model. The overall accuracy and confidence in the detections is summarized in Figure 24.

VIII. DISCUSSION OF RESULTS
All the tests done in this research were performed to ensure that the deep learning model - object detector - is well suited for the use in construction sites. For that, multiple performance indicators were identified and checked throughout the evaluation process. The main image categories that were tested include RGB images (images with normal colors), grayscale images (images with no colors), images with maximum brightness, images with dusting effect, and blurred images. Moreover, the model was tested under dynamic conditions which is the MP4 videos. This test verified the model's efficiency under various and multiple conditions.

The model showed a very high accuracy rate in detecting the three desired objects. The model achieved around 94% overall accuracy in object detection and 83% confidence in the detected objects in normal RGB conditions. These values slightly decreased when the testing data set was converted into the grayscale condition. The model's accuracy in this case decreased to 87.70% and confidence to 77.30%. This decrease in numbers was expected since the three classes in this model are highly color dependent. The object with the major decrease in accuracy in this category was the lifeline. This is because the lifeline is a sort and a type of rope, so when the colors are removed from the images the detector cannot clearly identify if the object was a lifeline or not. Moreover, when the brightness of the images was maximized, the model achieved very high accuracy and confidence results similar to those of the RGB. The accuracy almost stayed the same with 93.5% and the confidence slightly decreased to around 80%.

The images in the testing set were altered in a manner to represent different weather conditions. Two main weather conditions were targeted (1) dusty weather (2) Humid/rainy weather. The results obtained from both tests demonstrated that the model is very efficient under any weather condition. Finally, the model’s performance under the video environment was also very accurate. The model’s achieved around 90% accuracy rate. The main advantage in the video detection that can be compromised in the construction sites is that if the model detected the object desired in one of its frames and then missed it in one of any other frames, the overall decision of the safety in charge is no violation. Since the desired PFAS is available and worn by the worker. The summary of all results is shown in Table 2. Finally, the model was very accurate and confident in the detections performed. Increasing the number of training images in the future will also lead to better results and outcomes.

FIGURE 24. Detection's accuracy and confidence (videos).
Table 2. Summary of results.

|                  | RGB     | Grayscale | Bright  | Dusty   | Blur    | Videos  | Average |
|------------------|---------|-----------|---------|---------|---------|---------|---------|
| Helmet's Accuracy| 95.20%  | 93.40%    | 94.10%  | 92.20%  | 95.20%  | 90.61%  | 93.45%  |
| Helmet's Confidence| 87.60% | 74.60%    | 83.30%  | 82.10%  | 84.80%  | 85.22%  | 82.94%  |
| Harness’s Accuracy| 96.65%  | 93.10%    | 96.50%  | 95.10%  | 92.70%  | 91.85%  | 94.32%  |
| Harness’s Confidence| 91.50% | 89.90%    | 90.25%  | 84.00%  | 90.52%  | 83.71%  | 88.31%  |
| Lifeline’s Accuracy| 89.80% | 76.50%    | 89.80%  | 87.50%  | 85.25%  | 87.15%  | 86.00%  |
| Lifeline’s Confidence| 70.10% | 67.40%    | 65.60%  | 65.10%  | 68.90%  | 76.18%  | 68.88%  |
| Overall Model’s Accuracy | 91.26% |           |         |         |         |         |         |
| Overall Model’s Confidence | 80.04% |           |         |         |         |         |         |

A. PRECISION AND RECALL

Another two parameters that were used to evaluate the model’s performance were the precision and recall. These measures are widely used to determine the ability of a system to classify and detect objects [15]. Precision attempts to answer the question: What proportion of positive identifications was actually correct? On the other hand, recall attempts to answer the question: What proportion of actual positives was identified correctly [35]? For that, the results were divided into three main categories True Positives (TP) which contains the number of correct objects detected. False Positives (FP) contains the number of mistakenly detected objects. False Negative (FN) contains the number of undetected objects. The precision and recall were defined as follows:

\[
\text{Precision} = \frac{TP}{TP + FP} \quad (1)
\]

\[
\text{Recall} = \frac{TP}{TP + FN} \quad (2)
\]

Upon the completion of calculating all the outcomes for precision and recall for each object, a table summarizing the results were generated. Table 3 shows all the values obtained after calculating the TPs, FPs, and FNs.

Table 3. Precision and recall results.

|         | Helmet | Harness | Lifeline |
|---------|--------|---------|----------|
| TP      | 352    | 357     | 259      |
| FP      | 0      | 2       | 6        |
| FN      | 20     | 39      | 43       |
| Precision | 1.0000  | 0.9944  | 0.9774  |
| Recall  | 0.9462 | 0.9015  | 0.8576  |

The results obtained showed a very high precision rates for all three objects. The model had a very minimal error degree in mis-detecting the desired objects. On the other hand, the recall results varied between 0.94 and 0.85 having the helmet as the highest and the lifeline as the lowest. Although the lifeline achieved the lowest recall value it’s still considered acceptable due to the novelty of performing such as detection on very random and irregular shaped object. Furthermore, the obtained recall values were for all 5 images conditions (RGB, Grayscale, etc.) discussed earlier. So, achieving such values in different light, color, and weather conditions can be considered a success and can lead to extreme reduction in construction sites accidents.

B. RESEARCH’S NOVELTY

As mentioned earlier in the literature review section several researches attempted to use artificial intelligence in the construction safety sector. For instance, some used machine learning to detect safety helmets, others used computer vision and deep learning to detect the safety harness. However, none of the studies and researches performed combined the detection of the safety helmet, safety harness, and most importantly the lifeline. The lifeline is a major component in the PFAS that could save thousands of lives if used properly. In this research a model combining all these three objects were trained in order to perform the detections.

The results obtained in this research were as high and accurate as those results obtained in the earlier studies. For instance, according to [16] a similar YOLO v3 model were trained to detect the availability of safety helmets and safety vests in a construction site. Their model achieved an overall precision of 96%. In this research an overall precision of 99% were achieved in detecting the desired objects. Furthermore, the novelty of this research resides in its ability to detect the desired objects under different color, light, and weather conditions. Most of the studies performed earlier in utilizing the use of AI in construction sites focused on only one specific image aspect which is in most cases the RGB photos. Furthermore, the uniqueness of this research lies in the idea of detecting the lifeline. The
A recent study attempted to use deep learning in detecting only the safety harness in construction sites. According to [15] Faster R-CNN was the deep learning algorithm used to perform the detections rather than YOLO. Their model achieved an 80% precision in detecting the harness. In this research the safety harness achieved a 99.44% precision rate. Thus, YOLO proved in theory and practice to be a very powerful tool in detecting construction related objects and can be used reduce construction accidents.

IX. CONCLUSION
Preventing accidents and ensuring the safety of workers on the jobsite is one of main goals when performing any construction projects. In this paper, a novel procedure was introduced in order to reduce accidents when working at heights. A deep learning model was trained and developed in order to detect the main components of the PFAS. The YOLOv3 algorithm was applied to a CNN model in this research which is verified theoretically and experimentally with real-life image data for practical applications and implementations at construction sites. The validation process of the model experienced multiple tests to ensure that the model can perform in any color, light, and weather conditions. The model previewed very capable results. The model scored 91.26% accuracy rate and around 99% precision. Furthermore, the overall recall of the model was 90.2%. These numbers are considered to be promising when compared to previous related works. Despite having some errors and mis detections the model can be considered reliable. Future studies and researches could aid in better performance.

APPENDIX
A unique dataset was generated especially for the purpose of this research. The dataset produced for this research is available upon request to the corresponding author.

REFERENCES
[1] G. E. Gürcanli and U. Mungen, “Analysis of Construction Accidents in Turkey and Responsible Parties,” Industrial Health, vol. 51, no. 6, pp. 581–595, 2013.
[2] M. K. Zekri “Construction safety and health performance in Dubai “, 2013.
[3] K. Kavya and T. Pradeep, “Causes and Effects of Construction Accidents,” International Journal of Innovative Technology and Exploring Engineering Regular Issue, vol. 9, no. 2. pp. 1129–1133, 2019.
[4] T. Umar, “Applications of drones for safety inspection in the Gulf Cooperation Council construction,” Engineering, Construction and Architectural Management, vol. ahead-of-print, no. ahead-of-print, 2020.
[5] A. Puri-Mirza, “UAE: Abu Dhabi contribution of construction to the GDP 2019,” Statista, 08-Dec-2020. [Online]. Available: https://www.statista.com/statistics/818149/uae-contribution-of-construction-to-the-gdp-in-abu [Accessed: 13-Jan-2021].
[6] M. Mohan and S. Varghese, “Artificial Intelligence Enabled Safety for Construction Sites,” International Research Journal of Engineering and Technology (IRJET), vol. 06, no. 06, Jun. 2019.
[7] Construction in the UAE – Key Trends and Opportunities to 2023 [Internet]. Research and Market, 2019 [cited 2019 Sep 5]. Available from: https://www.globaldata.com/store/report/gdcn0454mr--construction-in-the-uae-keytrends-and-opportunities-to-2023/.
[8] Middle Eastern promise: prioritizing safety [Internet]. Atkins, 2015 [cited 2019 Sep 2]. Available from: https://www.atkinsglobal.com/en-GB/angles/all-angles/middle-eastern-promise-prioritising-safety.
[9] Health Statistics 2015 [Internet]. Health Authority Abu Dhabi, 2016 [cited 2019 Sep 2]. Available from: https://www.haad.ae/HAAD/LinkClick.aspx?fileticket=t=-IF2Jo3nP%E3%&tabid=1516.
[10] P. Baras, K. Addley, M. Grivna, C. Stanculescu, and F. Abu-Zidan, “Occupational injury in the United Arab Emirates: epidemiology and prevention”. Occup Med (Lond), Vol. 59, pp. 493–498, 2009.
[11] T. Loney, R.F. Cooling, and T.C. Aw, “Lost in Translation? Challenges and Opportunities for Raising Health and Safety Awareness among a Multinational Workforce in the United Arab Emirate”. Safety and Health at Work, Vol. 3 (4), pp. 298-304, 2012.
[12] EHS RI – Code of Practice 23.0 – Working at Heights, Version 2.0, Abu Dhabi EHSMS Regulatory Framework (AD EHSMSRF).
[13] W. Fang, L. Ding, P. E. D. Love, H. Luo, H. Li, F. Peña-Mora, B. Zhong, and C. Zhou, “Computer vision applications in construction safety assurance,” Automation in Construction, vol. 110, p. 103013, 2020.
[14] A. H. M. Rubaiyat, T. T. Toma, M. Kalantari-Khandani, S. A. Rahman, L. Chen, Y. Ye, and C. S. Pan, “Automatic Detection of Helmet Uses for Construction Safety,” 2016 IEEE/WIC/ACM International Conference on Web Intelligence Workshops (WIW), 2016.
[15] W. Fang, L. Ding, H. Luo, and P. E. D. Love, “Falls from heights: A computer vision-based approach for safety harness detection,” Automation in Construction, vol. 91, pp. 53–61, 2018.
[16] V. S. Delhi, R. Sankarlal, and A. Thomas, “Detection of Personal Protective Equipment (PPE) Compliance on Construction Site Using Computer Vision Based

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/
Deep Learning Techniques,” Frontiers in Built Environment, vol. 6, 2020.

[17] Kumar, K., & Kaashyap, A. “Improving Train Track Safety using Drones, Computer Vision and Machine Learning”, 2020.

[18] Y.J. Byon, A. Shalaby, B. Abdulhai, C.S. Cho, H. Yeo and S. El-Tantawy, “Traffic Condition Monitoring with SCAAT Kalman Filter-based Data Fusion in Toronto, Canada,” KSCE Journal of Civil Engineering, vol. 23, no. 2, pp. 810-820, doi: 10.1007/s12205-018-0132-5, 2019.

[19] Y.J. Byon, J.S. Ha, C.S. Cho, T.Y. Kim and C.Y. Yeun, “Real-time Transportation Mode Identification using Artificial Neural Networks Enhanced with Mode Availability Layers: A Case Study in Dubai,” Applied Sciences, vol. 7, no. 9, pp. 923, doi: 10.3390/app7090132, 2017.

[20] “Towards data science,” Towards Data Science. [Online]. Available: https://towardsdatascience.com/. [Accessed: 03-Sep-2021].

[21] S.J. Russell, P. Norvig “Artificial Intelligence: A Modern Approach”. Third Edition, Prentice Hall ISBN 9780136042594, 2010.

[22] M. Mohri, A. Rostamizadeh, A. Talwalkar, “Foundations of Machine Learning”. The MIT Press ISBN 9780262018258, 2012.

[23] Medium. [Online]. Available: https://medium.com/topic/machine-learning. [Accessed: 03-Sep-2021].

[24] “United States,” 3M in the United States. [Online]. Available: https://www.3m.com/3M/en_US/fall-protection-us/products/netting-and-guardrail-systems/. [Accessed: 20-Jun-2021].

[25] “Department of Labor logo UNITED STATESDEPARTMENT OF LABOR,” 1915.159 - Personal fall arrest systems (PFAS). | Occupational Safety and Health Administration. [Online]. Available: https://www.osha.gov/lawsregs/regulations/standardnumber/1915/1915.159. [Accessed: 27-Jun-2021].

[26] “6.2 Safety - sitechsolar,” Google Sites. [Online]. Available: https://sites.google.com/site/sitechsolar/Installation/safety. [Accessed: 07-Jul-2021].

[27] R. Girshick, "Fast R-CNN," 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, 2015. pp. 1440-1448, doi: 10.1109/ICCV.2015.169.

[28] S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137-1149, 1 June 2017, doi: 10.1109/TPAMI.2016.2577031.

[29] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, 2016, pp. 779-788, doi: 10.1109/CVPR.2016.91.

[30] S. Zhang, Y. Wu, C. Men and X. Li, "Tiny YOLO Optimization Oriented Bus Passenger Object Detection," in Chinese Journal of Electronics, vol. 29, no. 1, pp. 132-138, 1 2020, doi: 10.1049/cje.2019.11.002.

[31] “Overview of the YOLO Object Detection Algorithm,” Open Data Science - Your News Source for AI, Machine Learning & more, 26-Jun-2019. [Online]. Available: https://opendatascience.com/overview-of-the-yolo-object-detection-algorithm/. [Accessed: 02-Jun-2021].

[32] “Object Detection · datadocs,” · datadocs. [Online]. Available: https://polakowo.io/datadocs/docs/deep-learning/object-detection#yolov3. [Accessed: 07-Jul-2021].

[33] “Colaboratory,” Google. [Online]. Available: https://research.google.com/colaboratory/faq.html. [Accessed: 05-Jun-2021].

[34] “Introduction to TensorFlow,” TensorFlow. [Online]. Available: https://www.tensorflow.org/learn. [Accessed: 08-Jun-2021].

[35] “Classification: Precision and Recall &nbsp;&nbsp; Machine Learning Crash Course,” Google. [Online]. Available: https://developers.google.com/machine-learning/crash-course/classification/precision-and-recall. [Accessed: 04-Jul-2021].