

Article

Object Detection and Distance Measurement Algorithm for Collision Avoidance of Precast Concrete Installation during Crane Lifting Process

Yik Pong Yong ¹, Seo Joon Lee ¹, Young Hee Chang ¹, Kyu Hyup Lee ¹, Soon Wook Kwon ^{2,*}, Chung Suk Cho ³ and Su Wan Chung ⁴

¹ Department of Global Smart City, Sungkyunkwan University, Suwon 16419, Republic of Korea; yikpongyong@gmail.com (Y.P.Y.); sjlee3003@skku.edu (S.J.L.); yhyhchang@g.skku.edu (Y.H.C.); kyup@skku.edu (K.H.L.)

² School of Civil, Architectural Engineering & Landscape Architecture, Sungkyunkwan University, Suwon 16419, Republic of Korea

³ Department of Civil Engineering, American University of Bahrain, Riffa 942, Bahrain; chunguk.cho@aubh.edu.bh

⁴ Department of Future and Smart Construction Research, Korea Institute of Civil Engineering and Building Technology (KICT), Goyang 10223, Republic of Korea; suwan@kict.re.kr

* Correspondence: swkwon@skku.edu; Tel.: +82-031-290-7652

Featured Application: This system has the potential to be used as a guidance for the lifting path of any type of crane and a tool to help in preventing collision of a suspended or lifted load with any surrounding structure, facility, or human.

Abstract: In the construction industry, the process of carrying heavy loads from one location to another by means of a crane is inevitable. This reliance on cranes to carry heavy loads is more obvious when it comes to high-rise building construction. Depending on the conditions and requirements on-site, various types of construction lifting equipment (i.e., cranes) are being used. As off-site construction (OSC) is gaining more traction recently, cranes are becoming more important throughout the construction project as precast concrete (PC) members are major components of OSC calling for lifting work. As a result of the increased use of cranes on construction sites, concerns about construction safety as well as the effectiveness of existing load collision prevention systems are attracting more attention from various parties involved. Besides the inherent risks associated with heavy load lifting, the unpredictable movement of on-site workers around the crane operation area, along with the presence of blind spots that obstruct the crane operator's field-of-view (FOV), further increase the accident probability during crane operation. As such, the need for a more reliable and improved collision avoidance system that prevents lifted loads from hitting other structures and workers is paramount. This study introduces the application of deep learning-based object detection and distance measurement sensors integrated in a complementary way to achieve the stated need. Specifically, the object detection technique was used with the application of an Internet Protocol (IP) camera to detect the workers within the crane operation radius, whereas ultrasonic sensors were used to measure the distance of surrounding obstacles. Both applications were designed to work concurrently so as to prevent potential collisions during crane lifting operations. The field testing and evaluation of the integrated system showed promising results.

Keywords: object detection; computer vision; distance measurement sensor; precast concrete; off-site construction; collision avoidance; crane



Citation: Yong, Y.P.; Lee, S.J.; Chang, Y.H.; Lee, K.H.; Kwon, S.W.; Cho, C.S.; Chung, S.W. Object Detection and Distance Measurement Algorithm for Collision Avoidance of Precast Concrete Installation during Crane Lifting Process. *Buildings* **2023**, *13*, 2551. <https://doi.org/10.3390/buildings13102551>

Academic Editor: Yunchao Tang

Received: 6 September 2023

Revised: 1 October 2023

Accepted: 3 October 2023

Published: 9 October 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In recent years, the number of registered cranes in South Korea has grown tremendously as the importance of cranes at the construction site has continued to increase. As

of March 2022, 10,748 commercial lifting equipment and 5,996 commercial tower cranes were registered with the Ministry of Land, Infrastructure and Transport, Republic of Korea. These figures for commercial lifting equipment and tower cranes reflect an increase of 21.7% and 78.2%, respectively, over the past 10 years [1]. As off-site construction (OSC)—an umbrella term that gives a definition of assembling and installing volumetric construction components produced from a location different from the construction site where they will be used—is gaining more traction recently, cranes are becoming more important at construction sites [2]. While OSC, prefabricated construction, and modular construction are all similar concepts, modular construction is a subset of OSC, whereas OSC falls under prefabricated construction. Despite the higher construction cost per unit quantity (m^3) of precast concrete (PC), a major component of OSC, compared to traditional cast-in-place concrete, OSC application will increase in the coming years owing to its proven benefits, such as enhanced quality, productivity, and structural reliability, reduction in project duration, and waste of materials and manpower [3–5].

Jeong et al., however, claimed that the accident rate in OSC projects is much higher than in traditional on-site construction projects [6]. The OSC method can be divided into off-site manufacturing work and on-site construction processes. According to their research, 75.6% of safety accidents occurred during the on-site construction process, and 60.5% of them happened during the lifting and installation of materials. In addition, the statistics from 2010 to 2019, released by the Ministry of Employment and Labor, Republic of Korea (MOEL), show that a total of 162 out of 343 crane-related accidents occurred due to unauthorized access to crane lifting radius, inexperience in crane operation, and non-compliance with safety rules during load lifting operation [7]. Since the risk of accidents increases with the proliferation of OSC, extra safety measures should be considered and taken to safeguard on-site workers and maximize the benefits of OSC.

To cope with the safety issues of OSC and lifting operations in particular, the sensor integration method and object detection algorithm are mainly used in this study. A sensor is defined as a device that detects or measures a physical property, then converts the stimulus into a readable output. Specifically, the integration of sensors can be described as a process of combining multiple sensory datasets to obtain a more accurate and reliable representation. Object detection, on the other hand, is used to determine where objects are located in a given image or video, then categorize each object depending on how it was trained. Compared to other industries, such as finance, healthcare, entertainment, and education, the construction industry is slow in adopting new technologies and lacks advanced research on implementing AI and deep learning technologies on construction projects.

Akinosho et al. have summarized 44 research studies on deep learning applications in the construction industry and divided them into different functions [8]. As summarized in Table 1, their research identified that more than half of construction-related studies implementing deep learning involved applying convolutional neural networks (CNN). Most of these studies are related to construction equipment activity recognition, worker's protective equipment detection, as well as postural evaluation and activity assessment.

Table 1. Summary of deep learning applications in construction safety management.

Application	Reference
Construction equipment activity recognition	[9–13]
Worker's protective equipment detection	[14–16]
Worker's postural evaluation	[17–19]
Workforce activity assessment	[20,21]

Furthermore, a variety of research has been conducted on tower crane anti-collision systems over the years; however, not many studies specifically analyzed the lifting-load-centered anti-collision system. As an additional consideration for the increasing utilization of cranes on construction sites, especially with the growing adoption of OSC methods, a

need for a collision avoidance system that prevents lifted PC components from hitting other structures and workers is paramount.

With the goal of reducing collision accidents during the lifting operation of PC components, this study proposes a system that makes an objective judgment based on the collected data from computer vision and sensors rather than relying on the subjective judgments of field workers and crane operators. The following procedures were implemented in this study: (1) all risk factors associated with implementing the OSC method were analyzed and the causes of lifting equipment-related accidents were identified; (2) the commercialized crane collision avoidance systems were examined and recent related research results were compiled; (3) a smart system for load collision avoidance, especially during lifting operation of PC components, was designed using object detection and distance measurement sensors; (4) a real-time data prototype was created, and the performance of the smart system was field-tested by applying it to a crawler crane.

2. Background and Related Studies

Presently, a wide variety of cranes can be found at construction sites and can be used in different conditions or under varying requirements. During the lifting operation, hand signals are the main form of communication between the crane operator and signalman, both of which require professional certification. As such, training is needed to properly use hand signals, which are standardized in most countries. Under circumstances where hand signals are not clearly visible, voice signals are used as an alternative mode of communication. Regardless of which method is used, the lifting operations heavily rely on subjective human judgments and therefore warrant a collision avoidance system incorporating advanced technologies. Accordingly, relevant research studies have aimed at building an anti-collision objective decision support system in recent years as summarized in Tables 1 and 2.

Table 2. Related research on vision-based crane collision avoidance system.

Method	Research		
	[22]	[23]	[24]
AI-based image processing	No	No	No
Non-AI-based image processing	Yes	Yes	Yes
3D visualization	No	Yes	Yes
Worker detection	No	No	Yes

2.1. Vision-Based Crane Collision Avoidance System

Yang et al. used image processing technology to create a video surveillance system that tracks the target object automatically and maintains the predetermined distance between obstacles and the lifted load of an overhead crane. When a moving object or an obstacle is detected in the predetermined monitoring area, an alarm is set to trigger immediately. However, this method had limitations in application at construction sites with complex background environments, especially with massive moving objects or workers [22].

To improve safety in blind lifting, Fang et al. introduced a computer vision-based approach to track load positions for load sway monitoring on offshore platforms. In this study, an overhead camera mounted on the crane boom tracks the load position and orientation by using a color-based segmentation method. However, due to the limitations of this method for visual tracking, large errors occurred when it was applied in a complex background environment [23].

Price et al. also proposed a real-time crane monitoring system for blind lift operations, which consists of computer vision methods and 3D visualization modules. They installed a camera on the boom tip of a stationary crane and used the Hough transform and the mean-shift clustering method to track the position of the hanging load and nearby workers. Their research showed a good result in detecting potential collisions between the hanging load and the nearby workers; nonetheless, the approaching of other objects other than

workers could not be detected, and there was no warning alert mechanism for relevant individuals of potential collision hazards [24].

2.2. Sensor-Based Crane Collision Avoidance System

Gu et al. established a state equation along with an observation equation for detecting tower crane obstacles and proposed an anti-collision monitoring system that implements a distance measurement method based on ultrasonic sensors [25]. Three ultrasonic sensors installed on the tower crane arm were used to prevent the collision of the crane in rotation; however, they were not able to detect the collision of the lifted load.

Similar research was conducted by Li et al., where multiple ultrasonic sensors were placed on a tower crane arm. Through their proposed scheme, the position of stationary end barriers and moving objects within the crane operation area can be detected [26]. The maximum measurement distance of their proposed system was 20 m, which was mainly designed to avoid tower cranes colliding with end walls or other overhead equipment but not to evade the collision of the load being lifted by the crane.

Sleiman et al. also created a sensor-based tower crane anti-collision monitoring system [27]. They simulated the cranes' movements within a four-dimensional model by using measurements from multiple sensors, including an angular position transducer on the jib and two linear position sensors on the crane hook. Their four-dimensional simulation model was developed by continuously feeding all necessary parameters, such as the coordinates of the crane hook, jib dimensions, and tower height, to the three-dimensional crane model. This study also aimed to avoid collisions between cranes but not to prevent the collision of lifted loads.

Zhong et al. proposed an application combining a wireless sensor network and the Internet of Things to create an anti-collision management system for a tower crane group [28]. Their anti-collision controller processes the data collected from the angle sensor, displacement sensor, tilt sensor, load sensor, and wind speed sensor and runs the anti-collision algorithm. This system was also created to prevent collisions of tower cranes with other adjacent cranes or buildings, but not with nearby dynamic or moving objects or workers.

Lee et al. developed a tower crane navigation system that provides three-dimensional information about the construction site and the location of objects being lifted in real-time using various sensors and a building information modeling (BIM) model [29]. This research showed a good result in preventing collisions between cranes and static structures, but not with moving objects or workers and dynamic structures.

Kim et al. conducted a study to develop and validate a virtual platform for autonomous tower crane technology [30]. To construct the virtual environment, a large amount of data, such as the surrounding environment data and the motion of each main part of the cranes, was collected by using a camera, Lidar, and other sensors. The crane operator was able to apprehend the surrounding conditions and crane movement by viewing the user interface of the virtual environment. However, the moving pattern and schedule dataset of a dynamic object had to be inputted in advance, which was not ideal for real-time collision avoidance, especially when workers were involved (Table 3).

Table 3. Related research on sensor-based crane collision avoidance system.

Method	Research					
	[25]	[26]	[27]	[28]	[29]	[30]
Sensor on hook block	No	Yes	Yes	Yes	No	Yes
Sensor on jib/ trolley	Yes	Yes	Yes	Yes	Yes	Yes
Crane anti-collision	Yes	Yes	Yes	Yes	Yes	Yes
Dynamic object anti-collision	No	Yes	No	No	No	No
3D Visualization	No	No	Yes	No	No	Yes

3. Risk Analysis for OSC Projects

3.1. Risk Factors of Implementing the OSC Method

Jeong et al. argued that intensive safety management is required for unit installations during the construction process because of the inherent risk of workplace accidents modular constructions pose [6]. A crush or crash of hanging loads constitutes one of the major accidents construction workers are commonly exposed to. During the offsite manufacturing and onsite construction processes, the ratio of crash or crush accidents each occupied 8% and 27%, respectively.

In OSC projects, a high-performance crane is commonly required to lift structures of different weights. During the crane operation, the larger the PC or prefabricated construction components, the higher the risk of collision when lifted, thereby posing more danger. Rosén et al. conducted an experiment on analyzing the impact of driving speed on pedestrian fatality risk. According to their research, forward-moving passenger vehicles hitting adult pedestrians with an impact speed of 40 km/h to 49 km/h cause a 2.9% fatality risk, whereas the fatality risk increases to 18% when the impact speed reaches 50 km/h to 59 km/h [31]. Besides the impact speed, crush-related injury risk is influenced by the object's mass, materials, impact area, height and weight of the individual facing injury risk, and other factors varying the risk level.

Jeong et al. and Kang et al., in their respective studies, claimed that falling from height occupied the highest proportion of accidents at construction sites: the former claimed 36% and the latter 44.6% [6,32]. According to the Korea Occupational Safety and Health Agency (KOSHA), the construction industry has the highest fatality rate among all industries while taking third place in non-fatal injury rate. Among the KOSHA fatality and injury rates, 37.2% of fatal accidents and 13.3% of injury accidents were due to falling from height [33].

Work at height (or high-place work)—any work that takes place above ground, usually 2 m in height, where a person could fall if precautions were not taken—commonly takes place at construction sites. Although the number of accidents leading to injury or fatality due to a crash or crush is relatively less than that of a fall, a crash and a fall can occur concurrently at the same spot, with one leading to another in any order. During the on-site PC assembly or inspection process, workers are frequently working at height, using the PC member itself as a platform. Making unexpected contact with the hoist line or colliding with a suspended load due to a lack of training and communication may cause workers to fall from height. Therefore, extra precautions should be taken when workers on OSC projects are exposed to the aforementioned safety hazards.

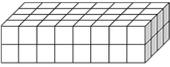
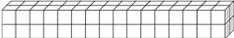
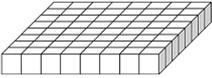
3.2. Analysis of Precast Concrete Lifting Operation

PC component refers to a construction product manufactured by casting reinforced concrete at a factory, which is then transported to the construction site for assembly. PC is usually classified into PC for buildings and PC for infrastructure. Within PC for buildings, there are architectural PC and structural PC. Architectural PC includes curtain walls and other non-load-bearing structures, whereas structural PC refers to load-bearing columns, girders, beams, walls, etc. that require lifting work, which is the main focus of this study. For lifting operations, mobile cranes and tower cranes are the two most frequently used lifting equipment: the former is commonly used when the lifting load requires a certain amount of transportation, while the latter is stationary and requires greater reach. Since both have similar lifting capacities, the decision on which to use depends on the project requirements as well as the jobsite conditions.

In general, the probability of a crash or struck-by hazard is relatively higher when lifting PC components than when lifting cast-in-place concrete buckets with a similar surface area on all sides. The PC components are cast into different shapes and sizes according to their function. Certain components, such as wall panels, slabs, and columns, have varying surface areas on each side. Where length, width, and height are measurements indicating object size, the object is said to have low compactness and an elongated shape if one of the measurements is distinct from the other two. As such, PC components with sides

with varying surface areas present a higher collision risk during lifting since the side with a large surface area is likely to increase the collision hazard with less space to evade. In short, two objects with the same volume but different shapes will have a different level of collision risk. Table 4 shows the relationship between object shape and the corresponding level of collision risk. A total of 64 cubes with the same volume are arranged in four different shapes or dimensions to grade their collision risk level.

Table 4. Relationship between shape of an object and its level of collision risk.

Shape	Dimension ¹	Largest Surface Area (unit ²)	Ratio of Largest Dimension to Smallest Dimension	Level of Collision Risk
	4 × 4 × 4	16	1:1	low
	8 × 4 × 2	32	4:1	medium
	16 × 2 × 2	32	8:1	high
	8 × 8 × 1	64	8:1	high

¹ The dimension of each cube is assumed to be 1 × 1 × 1 unit.

According to the safe operating procedures guide for mobile cranes prepared by South Korea's MOEL and KOSHA, the construction site supervisor has the responsibility to make sure no workers are present within the operation radius of a mobile crane in action and to check the clearance of the area under the suspended load. However, the presence of workers within the prohibited area during lifting operations (due to their negligence), inexperience in safe crane operation, and non-compliance with safety rules are the main causes of crane-related accidents in South Korea [7]. Accordingly, a hypothesis is drawn from this observation: knowing the presence of dangers, workers on construction sites intentionally choose to ignore the safety guidance so as to finish the work earlier or for their own convenience.

Additionally, effects of wind, over-rotating of mobile crane during lifting operation, and mismatch between the load's centre of gravity and hoist line are major factors leading to collisions between the lifted load and nearby structures or workers. These factors tend to enlarge the planned radius of the crane operation and therefore increase the risk of a crash with surrounding objects. Table 5 describes the cause factors of expected risk for each scenario.

Table 5. Factors that increase risk of collision between lifting load and surrounding objects.

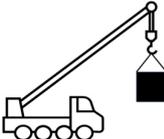
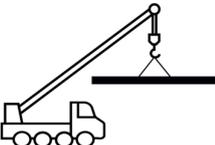
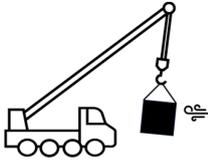
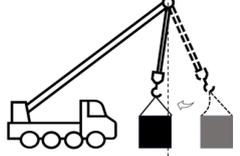
Scenario	Explanation of Scenario	Result
	Standard operating procedure is followed, and a high-compactness load is lifted. Load is transferred safely in a controlled environment. This scenario is taken as benchmark to compare performances of other scenarios.	Accident can still happen due to carelessness or negligence of workers and external factors.
	Standard operating procedure is followed. However, a low-compactness load which is flat or elongated shape is lifted.	Large dimension of load increases the risk of collision.

Table 5. Cont.

Scenario	Explanation of Scenario	Result
	High wind pressure occurs around the load and causes it to oscillate or swing.	Irregular and unpredictable movement of load increases the risk of collision.
	Rotation of mobile crane at high speed when a load is suspended causes the crane working radius to be increased instantaneously.	Centrifugal force and inertia of load increase the risk of collision.
	The load is lifted at a point far from the axis of hoist line. In case when the starting point is inside the hoist line axis, the operation radius will be increased instantaneously when the load is lifted.	Instant movement of load increases the risk of collision.

As in the first scenario, collision accidents can still occur even when standard operating procedures are followed by all parties involved. Thus, considering that PC components can come in various sizes and shapes, a more reliable approach to monitoring safe lifting operation conditions and providing an advance warning system for imminent danger recognition is required. With the aforementioned in mind, this study introduces a lifted-precast concrete collision avoidance system.

4. Precast Concrete Collision Avoidance in Crane Lifting Process

4.1. Overall Architecture of Lifting-Load Collision Avoidance System

For the anti-collision system of the PC component under lifting operation, this study implemented deep learning-based object detection for worker detection and ultrasonic sensors for surrounding obstacle detection. A downward-facing Internet Protocol (IP) camera was installed on the crane jib so as to capture the ground condition in real time, while ultrasonic sensors are laid horizontally on the hook block with the wave transmitter facing outwards. The collected video data and ultrasonic sensor data were then sent wirelessly to a computer for further processing. Subsequently, computer vision and deep learning techniques were applied to process all frames of the live video streaming for worker detection and their location prediction, while ultrasonic sensor data were processed and analyzed for obstacle distance measurement. Neither an IP camera nor an ultrasonic sensor can see or detect an object through an opaque material. Consequently, the implementation of these two methods increased the reliability and validity of the detection results. Figure 1 shows the overall process of the lifted-load collision avoidance system proposed in this study, whereas Figure 2 introduces the algorithm for merging computer vision and distance measurement sensors in constructing a lifted-load collision avoidance system.

Object detection used to detect the location of workers under crane operation refers to a computer vision and image processing task that identifies objects of a certain predefined class within an image.

Considering that the line of sight of a camera may be obscured while the PC component is being lifted, distance-measuring ultrasonic sensors were installed on the hook block, which work in parallel with the camera to detect any obstacles around the crane operation area. For early warning purposes, the sensors were oriented a few degrees downward, facing the ground. In addition to these sensors for measuring the distance of nearby obstacles, an additional distance-measuring sensor (i.e., distance between hook block and

ground) was used to prevent collisions between hook block and any underneath object or worker. All data collected by sensors is transmitted through a ZIGBEE device to a computer located inside the crane operator cab.

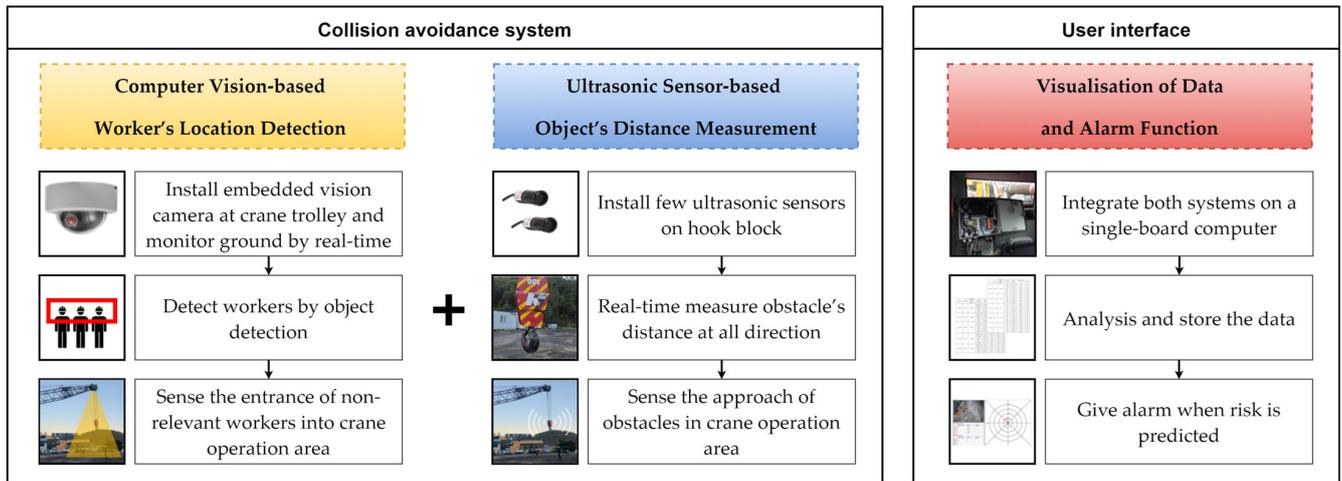


Figure 1. Computer vision for detecting workers' location and sensors for measuring obstacles' distance.

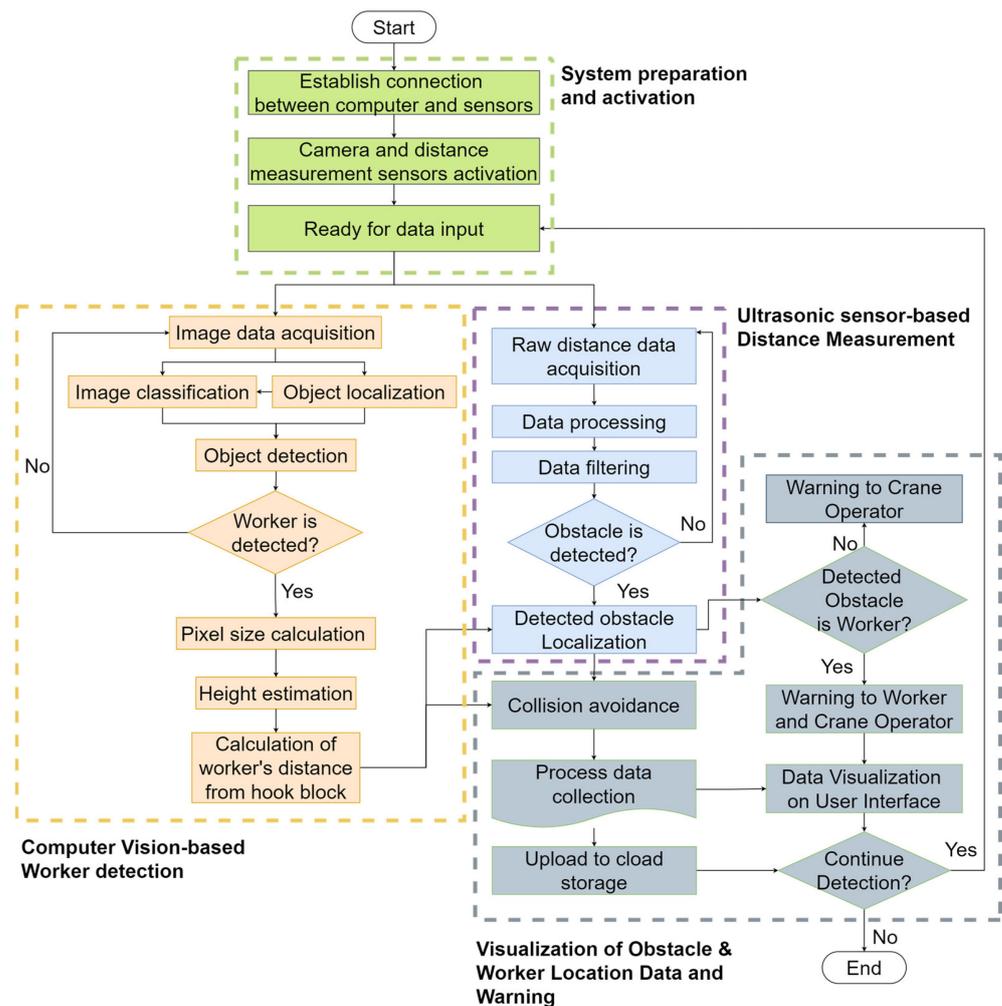


Figure 2. Algorithm of proposed crane-lifting-load collision avoidance system.

4.2. Laboratory Experiment of Prototype Model

A laboratory prototype was constructed to verify the operating principle of the proposed system and the soundness of its architecture as a whole. For the lab experiment, the NVIDIA Jetson Nano, a small but powerful single-board computer that can run multiple programs and neural networks in parallel, was used to execute both the object detection algorithm and sensor fusion. A Serial Peripheral Interface (SPI) camera was used to capture images instead of an IP camera, whereas economical sensors were used in place of industrial sensor modules for distance measurement. Video data from the SPI camera was transmitted to the single-board computer for information extraction, while distance data from the sensors was directed to Arduino for data acquisition. Figure 3 shows the devices used in the lab experiment, while their features and functions are listed in Table 6.

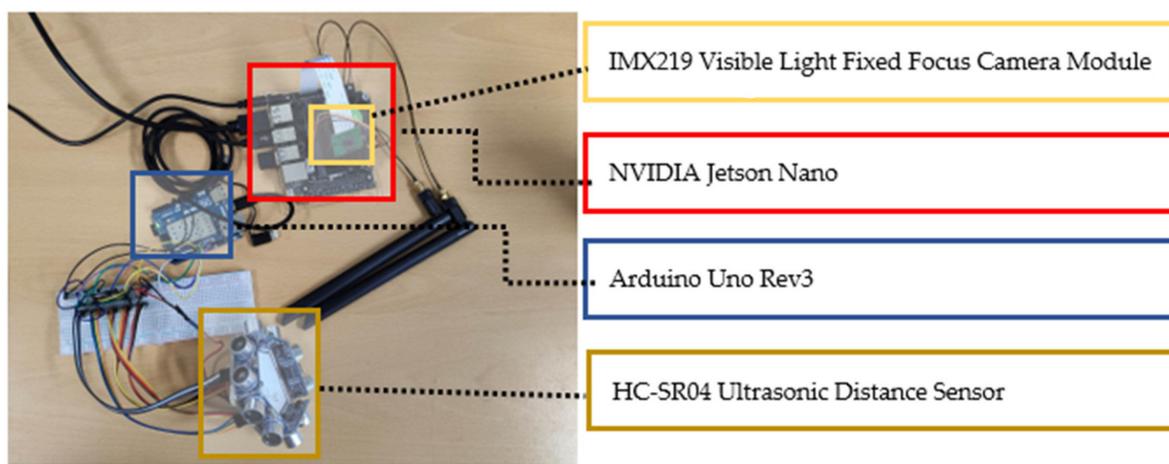


Figure 3. Prototype of lifting-load collision avoidance system for lab experiment.

Table 6. Features and functions of each device used in constructing laboratory prototype.

Device	Feature	Function
NVIDIA Jetson Nano	128-core Maxwell GPU; Quad-core ARM A57 @ 1.43 GHz CPU; 100 mm × 80 mm × 29 mm	Perform computing task and provide interfacing
Arduino Uno Rev3	ATmega328P microcontroller; 16 MHz of clock speed	Read raw data of sensors
IMX219 Visible Light Fixed Focus Camera Module	8 Megapixel IMX219 sensor; 21fps@8MP; 62.2° of horizontal viewing angle	Capture visual information
HC-SR04 Ultrasonic Distance Sensor	2 cm to 400 cm of ranging distance; 15° of effectual measuring angle	Non-contact distance measurement

4.2.1. Video Image-Based Object Detection

There is no straight answer as to which object detection model is best for a certain purpose. For real-life applications, it is important to select the appropriate model that balances detection accuracy and computing speed. Several factors should be considered that may have an impact on the model performance: type of object detection architecture, type of feature extractors, resolution of input image, data augmentation, training dataset, loss function, as well as training configuration, such as batch size, learning rate, learning rate decay, etc.

In order to compare and find the best-performing object detection model for this study, a training session was conducted five times using the same dataset, which consisted of 8,470 images with a total of 26,655 workers labeled for both training and evaluation. The workers in images were labeled according to the distance between camera location and workers' positions: within 5 m for 'very near', up to 10 m for 'near', up to 15 m for 'medium', up to 20 m for 'far', and up to 25 m and over for 'very far'. The dataset was then

divided into two parts, in which 80% was used for training and the remaining 20% was used for evaluating the accuracy of the model after training was completed.

For the first trial of training, a combination of a quantified Single Shoot Detector (SSD) model as a framework and MobileNet V2 as the backbone was used to perform object detection with a deep learning architecture. The SSD model was chosen first because of its fast computational performance, which meets our research goal of detecting the presence of workers in real-time. MobileNet was selected for the object detection model as a feature extractor due to its small size and ability to be used on a single-board computer. The first trial resulted in high performance in terms of computing speed, which processed 10.63 frames per second on average; however, the final mean average precision (mAP), a metric used to measure the accuracy of object detection (calculated by dividing true positives by total positive results), reached only 0.408, indicating the accuracy to be 40.8%.

In order to identify an object detection model that yields much higher accuracy, four additional trainings were conducted using the same dataset, data augmentation method, number of training steps, batch size, learning rate, and learning rate decay, but by varying the object detection algorithm, feature extractor, and resolution of input images. For the study, a momentum optimizer was applied to all models, and the batch size was lowered to two in order to train the model on GPUs with low memory sizes. All models were trained for 25,000 steps, and a dropout layer was brought into use. In addition, all five trainings were conducted using the cloud Nvidia Tesla T4 GPU in Google Colaboratory, where the training speed was affected by the number of simultaneous cloud GPU users. As such, the training time could not be taken into account when evaluating the efficiency of each model. However, the training steps, batch size, and other parameters were considered efficiency measures since the results will not vary significantly if the same values are used even with the fluctuating training speed. Evaluation of the object detection models was carried out based on the videos captured from the site of the field experiment.

As shown in Table 7, the accuracy in mAP and processing speed in frames per second varied under different deep learning algorithms, feature extractors, and input image sizes. From this comparison, the SSD algorithm had an advantage over the Faster RCNN in processing speed but a lesser ability to detect target objects in an image compared to the Faster RCNN. The Faster RCNN algorithm, on the other hand, performed better in detecting tiny objects, resulting in a higher mAP compared to the SSD model under the same feature extractor and input image size. Furthermore, it was noticed that the size of the input image had a significant impact on the detection accuracy as well as the image processing time.

Table 7. Parameters and functions for each object detection training process.

	Training Model				
	First	Second	Third	Fourth	Fifth
Deep learning algorithm	SSD	SSD	SSD	Faster RCNN	Faster RCNN
Feature extractor	MobileNet V2	Resnet50 V1	Resnet50 V1	Resnet50 V1	Resnet50 V1
Input image size	640 × 640	640 × 640	1024 × 1024	640 × 640	1024 × 1024
Frame per second	10.63	4.85	3.03	3.20	2.64
mAP	0.749	0.707	0.791	0.751	0.831

4.2.2. Sensor-Based Distance Measurement

Prior to constructing the full-scope PC collision avoidance system, ultrasonic sensors HC-SR04 were used and tested on the laboratory prototype. Five HC-SR04 ultrasonic sensors were arranged in a pentagon shape with their signal transmitter and receiver facing outward, which were then connected to a microcontroller board as shown in Figure 4.

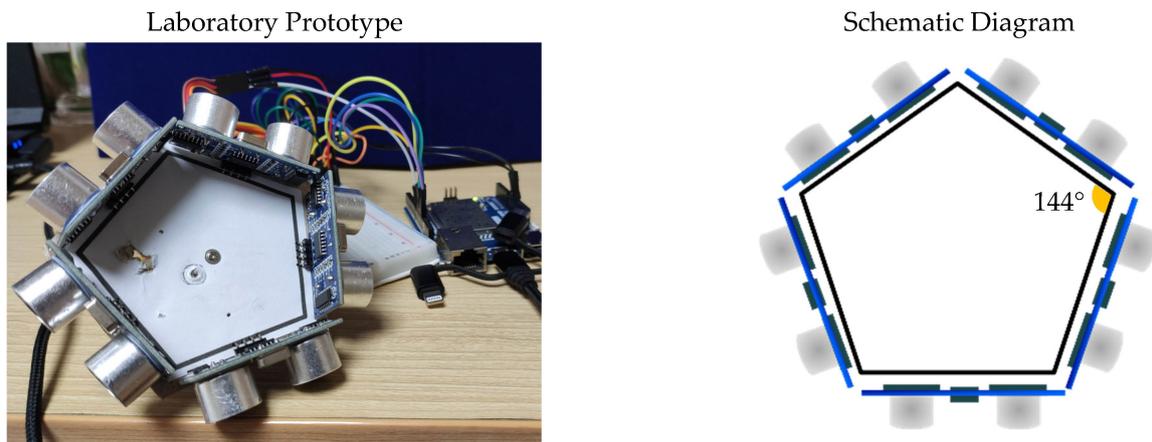


Figure 4. Laboratory prototype of sensor-based distance measurement and its schematic diagram.

The HC-SR04 ultrasonic sensor, having an effective detection angle of within 15° , emits a frequency of 40 kHz. Distance measurement data collected from five sensors was sent through a serial port of the Arduino Uno microcontroller board to the computer for further processing. Each sensor provided one measurement reading at a time. However, a common limitation of ultrasonic sensors to consider is the data fluctuation resulting from the interference of echo signals, which occurs when more than one ultrasonic wave is emitted from different ultrasonic sensors simultaneously at the same frequency. To prevent this irregular fluctuation of data due to wave interference, an interval of 65 ms was set between each ultrasonic emission time. As a result, the average time of each iteration for giving a new line of reading was around 0.6 s.

Subsequent lab experiments were conducted to test the stability and reliability of sensor data. Five ultrasonic sensors connected to a microcontroller board were placed in a space where there was no moving object within their measurable distance. Figure 5 shows the stability test result of the measurement reading obtained from the five HC-SR04 ultrasonic sensors.

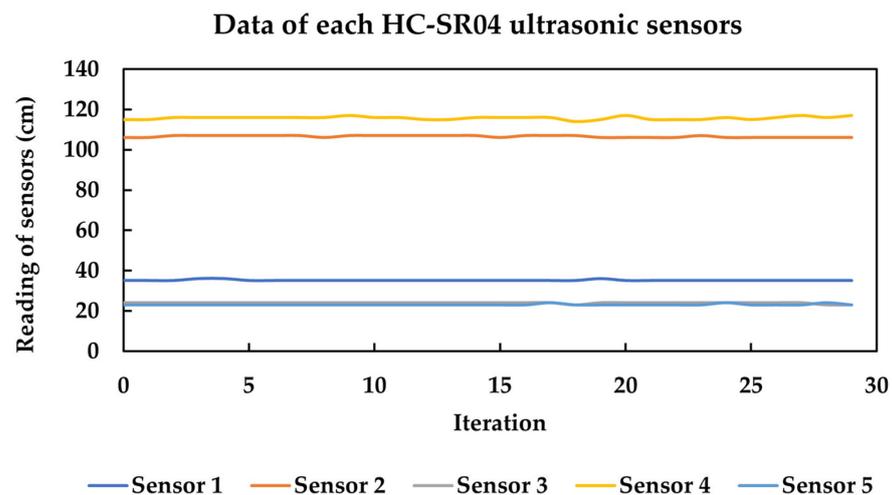


Figure 5. Graph of stability test result of measurement reading obtained from multiple ultrasonic sensors tested under a motionless environment.

Besides differentiating the time intervals between the emission of ultrasonic waves from the sensors, another variable manipulated and tested was the surface variation or angle of the object. A sensor emits an ultrasonic wave that reflects back to the sensor after hitting an obstacle. The distance from the sensor to the obstacle is measured by calculating

the time between emission and reception. Considering that various types of objects, such as temporary and permanent structures with different materials, construction equipment, and workers, are present at the construction site, this test was conducted to account for the fact that ultrasonic waves can be absorbed or deflected away from the sensor due to the irregular surface of the detected object or its angle to the sensor, which leads to an unreliable distance measurement result.

Similarly, an additional experiment was set up to test how the angle of an object affects the reliability of sensor readings. As shown in Figure 6, a hard, flat plastic board was set perpendicularly 10 cm away from the direction of the ultrasonic wave transmitted, then twenty sets of distance readings were recorded, and the average distance was calculated. The steps were repeated five additional times, each time reducing the angle of the board by 15°. The same procedures were applied to 20 cm and 30 cm of distance from the ultrasonic sensor to the board.

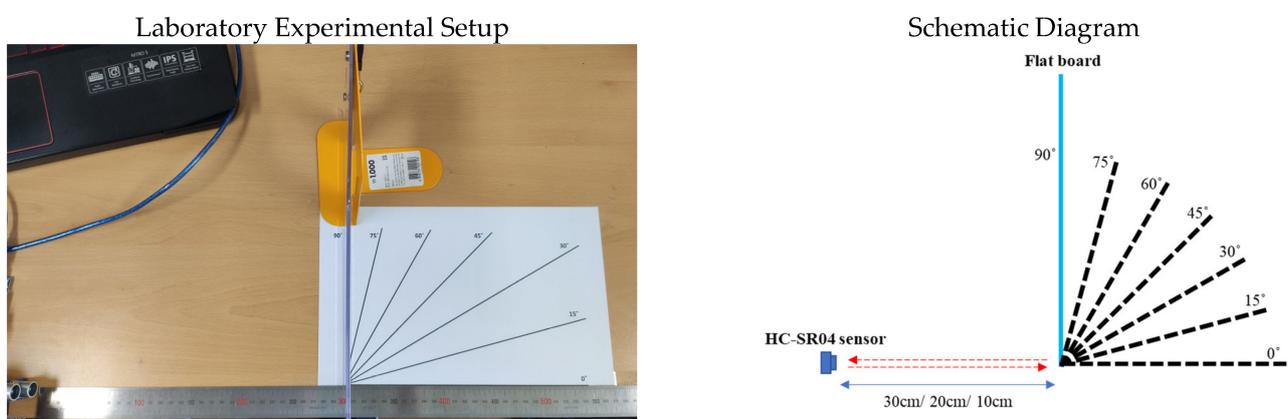


Figure 6. Laboratory experiment setup to test the relationship between the reliability of ultrasonic sensor readings and the angle of detected object.

It was observed from this experiment that when the surface of a target is not exactly perpendicular to the HC-SR04 sensor, the target object deflects the incident ultrasonic pulses away from the sensor, making the sensor unable to catch the reflected wave immediately. As can be seen in Figure 7, the larger the inclination of the object surface, the lesser the accuracy of the sensor reading until the inclination angle decreases from 90° to 30°. However, the effect of inclination on accuracy becomes smaller starting from 30° to 15°.

Effect of object angles on sensor reading

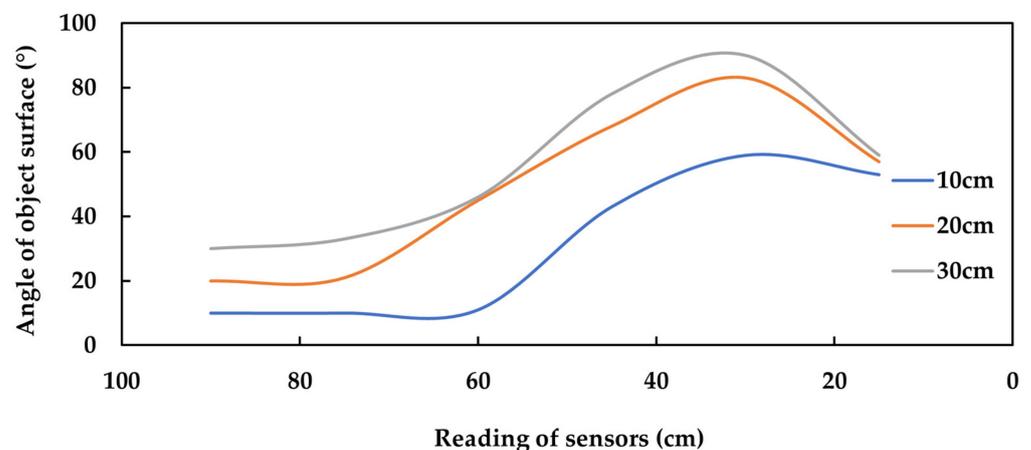


Figure 7. Laboratory test results on accuracy of ultrasonic sensor readings depend on angle of object surface reflecting ultrasonic wave.

4.3. Field Testing of the Prototype

After verifying the soundness of the detection architecture and the design of the obstacle avoidance system prototype, the model was taken to the field and applied to the Huvance Hydraulic Crawler Crane H900 for field testing. For the real-life application with enhanced reliability and practicality, a higher computational performance single-board computer, six units of industrial ultrasonic sensors, and an industrial high-definition IP camera were used for field testing. Additionally, for convenient installation and scalability, a wireless network that fit into the overall network infrastructure of the system was built to replace the completely wired network used in the laboratory prototype. A Wi-Fi transmitter and a ZIGBEE transmitter were used for transporting image and distance data, respectively.

The IP camera was installed at the end of the crane jib, facing downward perpendicularly to the ground, even though the angle of the crane jib varies from time to time during the lifting and loading operations. Six units of ultrasonic sensors, arranged in a hexagonal shape where each vertex is occupied by one sensor unit, were placed on the crane hook with the wave transmitter and receiver facing outward. These ultrasonic sensors were connected to a ZIGBEE transmitter, which sends the distance data wirelessly to the computer located on the crane cab. In addition to this single-board computer, a monitor that displays the user interface was also installed so that the crane operator could monitor the surroundings of the crane during lifting and loading operations. Figure 8 shows the overall setting of the field experiment, while Table 8 lists the devices that were used.

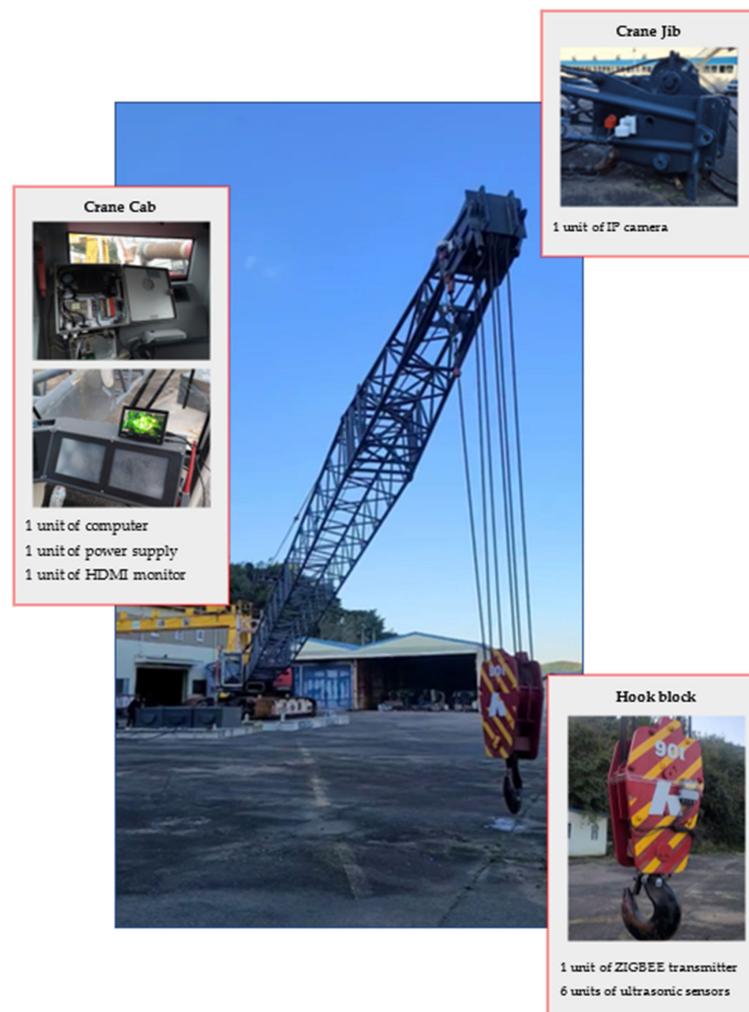


Figure 8. Setting of field experiment to test the function of proposed lifting load collision avoidance system.

Table 8. Devices used for object detection and distance measurement during field experiment.

Device	Feature	Function
NVIDIA Jetson AGX Xavier	512-core Volta GPU; 8-core ARM v8.2 64-bit CPU; 105 mm × 105 mm × 65 mm of module size	Performs computing task and provides interfacing
Hikvision PTZ IP Dome Camera	3 Megapixel CMOS sensor; 30fps@3MP; 102° of horizontal viewing angle; 1/2.8 of sensor size; 4× Optical Zoom; water, dust, and vandal resistant	Captures visual information for object and worker detection
STMA-503 Ultrasonic Sensor Module	30 cm to 200 cm of ranging distance; 90° of effectual measuring angle	Non-contact distance measurement and moving object counter
HDMI monitor	7-inch HD display	Displays vision for crane operator
ZIGBEE SH-WST24 router with XBEE3 module	Radio frequency 250 Kbps; Serial 1 Mbps; Up to 1200 m of outdoor and 60 m of indoor transmission range	Transmits sensor data wirelessly to computer for processing
LRS-100-24 Switching Power Supply	24 V DC voltage, 4.5 A rated current, 108 W rated power	Supplies electric power to sensors and camera

The aforementioned field-testing setup enabled quantitative evaluation of the performance and effectiveness of the proposed lifting-load anti-collision system. The distance measurement system used for detecting the presence of any obstacle around the hook block was updated at 1.67 Hz, whereas the vision-based object detection for locating workers within the working radius of the crane was updated within 7 to 10 Hz.

4.3.1. Data Transmission

The convergence of computer vision-based worker detection and sensor-based obstacle detection was proposed to enhance sensing limits and minimize blind spots in the detection of potential collisions during loading and lifting operations. Figure 9 illustrates the overall architecture of the proposed PC component collision avoidance system. The real-time video data are transmitted to a single-board computer placed on the crane operator cab via an IP network, whereas sensor data are sent to the computer over the ZIGBEE network for further processing.

In order to reduce network congestion, isolate security threats, and facilitate easier maintenance, the computer and the IP camera were addressed to the same subnet so that the IP camera could communicate directly with the computer without going through any routers. The multithreading method was applied—a model execution that allows multiple instructions to be scheduled and executed independently while sharing the same process resources—to collect, process, and visualize the image and distance measurement data simultaneously, although they are updated at different rates. While the video feed to be accessed, processed, and displayed relies on the speed of the decoding and encoding processes, the expected data latency of the IP camera used in this study was at least hundreds of milliseconds, which was considered still within the tolerance for real-time detection. The latency of distance measurement data, on the other hand, was always below 200 milliseconds, making it slightly lower than the video feed latency. The collected distance measurement data were sent in a group of 3 bytes per data iteration.

4.3.2. Accuracy Evaluation of Distance Measurement System

The accuracy of the measured distance between the crane hook and a nearby object was evaluated by comparing the true distance to the measurement reading obtained via ultrasonic sensors. Specifically, a concrete block was placed in a random position, and the true distance between the block and the crane hook was measured using a measuring tape. Then, all ultrasonic sensors were activated to detect the presence of the concrete block and to estimate its direction with respect to the corresponding sensor that transmitted the ultrasonic wave, which was then reflected back to the sensor after hitting the block. The

same procedures were carried out by replacing the concrete block with a worker to evaluate the sensor performance in the case of a dynamic detection target.

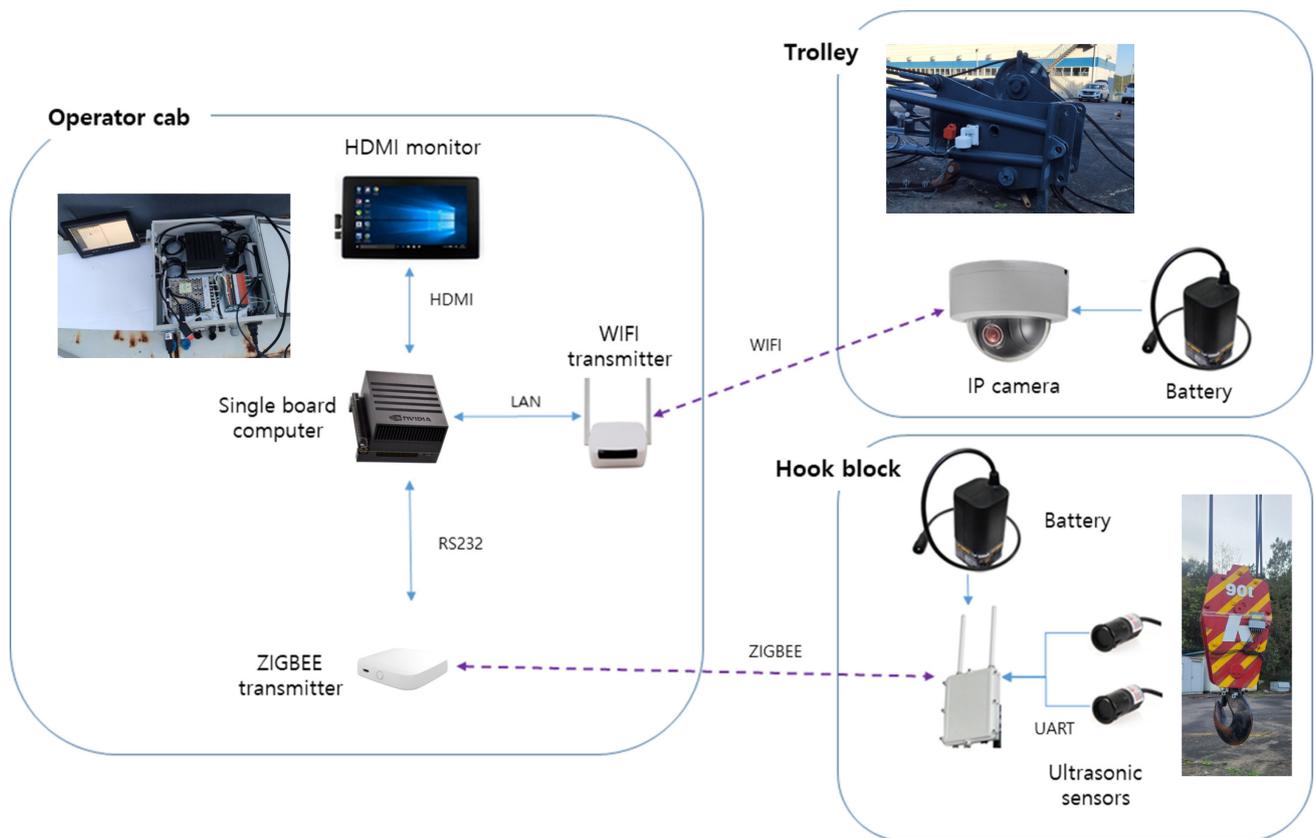


Figure 9. Data transmission of the proposed system.

Each set of experiments with different target objects was repeated five times after altering the target location. Table 9 and Figure 10 show the results of measuring the distance of the target object and worker from the crane hook using multiple ultrasonic sensors. As can be seen in Table 9, the further the target object location, the larger the error margin due to the energy loss and changes in intensity during the propagation of ultrasound waves in a medium. However, this was neglectable since the maximum detection range in this testing was two meters from the sensors. Consequently, the maximum absolute error that occurs within the area is deemed small enough to be ignored when this ultrasonic sensor-based anti-collision system is used on a real project. Overall, the system proposed in this study is considered acceptable for detecting the presence of any obstacle within the crane operating region.

Table 9. Evaluation of the distance measurement errors on ultrasonic sensor-based system.

Test	Number of Worker	True Distance (cm)	Index of Sensor	Measured Distance (cm)	Percent Error (%)
1	1 worker. 1 is detected	189	2	193	2.07
2	1 worker. 1 is detected	187	4	181	3.32
3	2 workers. 1 is detected	185 231	2	182 Not Detected; Over 200 cm	1.65
4	2 workers. 2 are detected	180 231	1	176 Not Detected; Over 200 cm	2.27

Test 1 refers to (a), test 2 refers to (b), test 3 refers to (c), and test 4 refers to (d) in Figure 10.

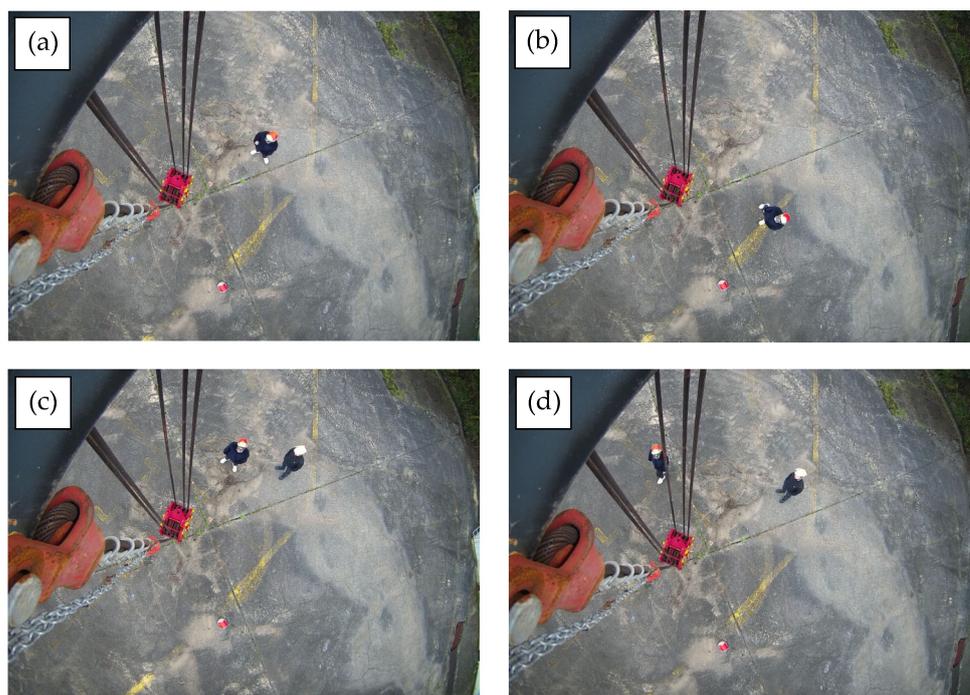


Figure 10. Sample images taken at crane yard through the camera installed at the tip of crane jib. All (a–d) were images captured at the height of 7 m from ground.

4.3.3. Accuracy Evaluation of Computer Vision System

The accuracy evaluation of a computer vision-based worker detection system was carried out in the same testing field. First, an IP camera was installed on the crane trolley, which is usually positioned at a height of 10 m or above with 2.8 mm of default focal length, allowing an overwide field view of 102 horizontally and 74° vertically, so that features are captured effectively on the video feed. In addition, the video feed was split into 16 equal parts so as to eliminate the distortion effect on the image and maximize the feature extraction performance. Only the middle four parts were cropped for image processing and were determined as a region of interest. Then, a random number of one to five workers wearing safety helmets were placed around the crane hook within the field of view. The direct distance from the worker was assumed to be tangent to the camera, making the vertical distance from the camera to the ground opposite. Additionally, the distance between the worker and the intersection point of the vertical axis of the camera and the horizontal axis of the worker was assumed to be adjacent. Finally, the crane trolley was operated at 7 m (i.e., 2 m from the ground to the crane arm and 5 m from the crane arm to the camera) vertically from the ground with the object detection running concurrently. The perimeter of the bounding box drawn on the detected workers was recorded. The same procedure was repeated for 10 m, 15 m, 20 m, and 25 m of trolley height. Table 10 and Figure 11 show the collected data and a graphical depiction showing the relationship between camera distance and the perimeter of the bounding box.

Table 10. Average perimeter of bounding box drawn around each worker in an image when captured at different heights.

Test	Height of Camera (m)	Worker	Actual Adjacent Distance from Hook, D_A (cm)	Average Perimeter of Bounding Box, P (Pixel)
1	7	1	241	238
		2	268	207

Table 10. Cont.

Test	Height of Camera (m)	Worker	Actual Adjacent Distance from Hook, D_A (cm)	Average Perimeter of Bounding Box, P (Pixel)
2	12	1	241	123
		2	268	116
3	17	1	241	96
		2	268	88
4	22	1	241	79
		2	268	74
5	27	1	241	65
		2	268	61

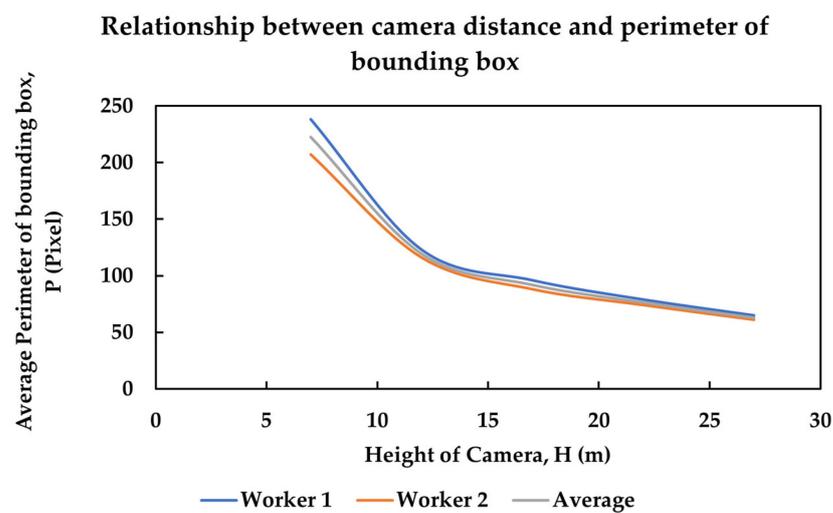


Figure 11. Graph shows the relationship between camera distance and perimeter of bounding box.

Based on the result shown in Figure 11, an exponentially decreasing function was observed, which led to the derivation of a regression equation as below. This equation shows the relationship between the perimeter of the bounding box drawn on the detected worker and the estimated height of the camera from the ground.

$$H = 2346P^{-1.084} \quad (P \neq 0) \quad (1)$$

The direction of the worker with respect to the crane hook, θ , can be calculated as the coordinate of the bounding box drawn on the worker is known. The coordinates of the bounding box are defined as y_{max} , x_{max} , y_{min} , and x_{min} , where the starting point is at the top-left corner of the frame. In this research, the video feed obtained from the IP camera was resized to 800 pixels of horizontal resolution, R_H , and 600 pixels of vertical resolution, R_V for better object detection performance. Subsequently, the coordinate at the top-left corner of the frame was set at (0,0), while the bottom-right corner was (800, 600). In the calculation of the direction of the worker, the crane hook was assumed to lie on the same vertical axis as the center point of the camera field of view. The middle point was defined as M_x and M_y , respectively. Equation (2) below shows the formula to calculate the direction of the worker with respect to the crane hook.

$$\theta = \tan^{-1} \frac{M_y - \left(y_{min} + \frac{y_{max} - y_{min}}{2} \right)}{M_x - \left(x_{min} + \frac{x_{max} - x_{min}}{2} \right)} \quad (2)$$

If θ is less than 90° , then the worker is in the first quadrant; the second quadrant if θ is between 90° and 180° ; the third quadrant between 180° and 270° ; and the fourth quadrant between 270° and 360° .

As the field of view (FOV) of the camera is known, the distance of the worker from the hook, D_E , can be calculated by substituting the estimated height of the camera, H , into Equation (3) below. The IP camera used in this research with a horizontal field of view, FOV_H , of 102° was installed at the right side of the tip of the crane jib, therefore not exactly on the same axis as the crane hook. Accordingly, 0.3 m was added to Equation (3) when the worker was standing on the right side of the FOV, while 0.3 m was deducted when the worker was on the left side of the FOV. The Equation (3) is as below:

$$D_E = \tan\left(\frac{FOV_H^\circ}{2}\right) H \times \frac{\sqrt{|M_x - x|^2 + |M_y - y|^2}}{R_H} \quad (3)$$

$$x = \begin{cases} x_{max}, & \text{third and fourth quadrant} \\ x_{min}, & \text{first and second quadrant} \end{cases}$$

$$y = \begin{cases} y_{max}, & \text{first and fourth quadrant} \\ y_{min}, & \text{second and third quadrant} \end{cases}$$

In order to verify the soundness of the equations above, the percent error was calculated by using Equation (4) as below:

$$\text{Percent error} = (|D_A - D_E| / D_E) \times 100\% \quad (4)$$

According to Table 11, the percent errors were within the range of 0.55% to 12.02%. It is noticeable that the percentage error increases as the worker stands further from the center point of FOV. In addition, it was found to be a challenging task to detect tiny objects with the object detection method, especially when using the Single Shot Multi-Box Detector algorithm, which is designed for faster processing speed at the expense of its accuracy. However, the validity of Equation (3) in estimating the adjacent distance of workers from the camera was verified. Thus, the proposed method of using the object detection algorithm for locating the workers based on the size of the bounding box drawn is proven to be valid as well.

Table 11. Calculation of percent error of the estimated adjacent distance from hook.

Test	Perimeter of Bounding Box, P (Pixel)	Height of Camera (m)	Estimated Adjacent Distance from Hook, D_E (m)	Actual Adjacent Distance from Hook, D_A (m)	Percent Error (%)
1	245	6.03	2.61	2.55	2.30
2	122	12.84	1.81	1.80	0.55
3	112	14.14	1.30	1.41	8.46
4	83	19.50	10.36	8.52	11.55
5	62	26.75	12.31	10.83	12.02

4.3.4. Graphical User Interface (GUI) Development

For the purpose of integrating and visualizing both the distance data obtained from ultrasonic sensors and computer vision to help the crane operator prevent any collision accidents, a user interface was created using the Python graphical user interface (GUI) framework. Figure 12 shows the sample views of the real-time GUI screenshot showing an actual situation during the crane operation.

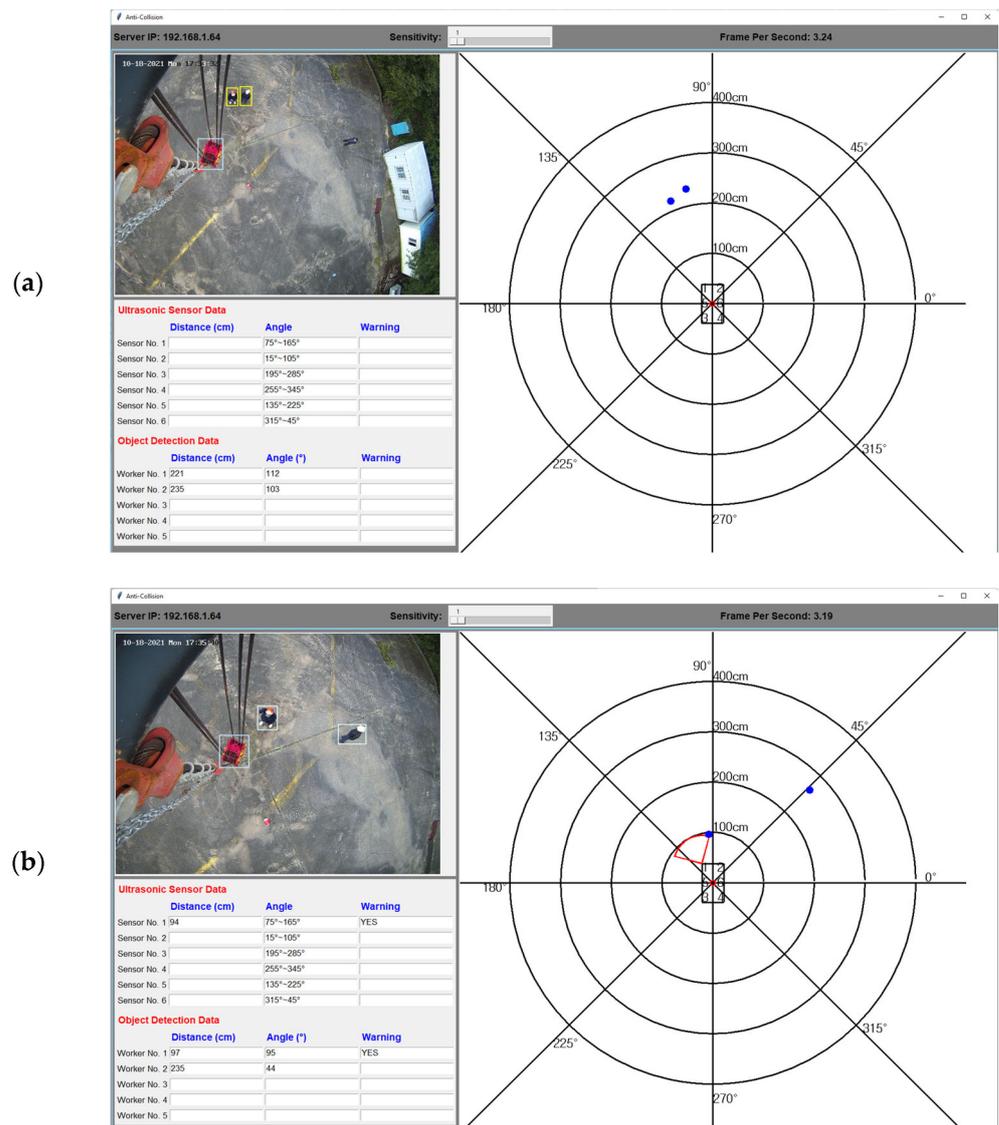


Figure 12. Sample views of GUI on monitor inside the crane cab.

On Figure 12a, as the maximum detection range of the ultrasonic sensors used in this research was up to 200 cm, no workers were detected by the distance sensors, and consequently, no result was shown. However, the object detection method applied was able to detect the workers successfully, but no warning sign was given since they were located outside the 2 m radius of the crane hook. In Figure 12b, however, a worker standing inside the detection range of the ultrasonic sensor labeled '1' was successfully detected by the object detection method. Since the worker was within a 2 m radius of the crane hook, a warning sign was also provided. Another worker standing outside the detection range of the ultrasonic sensor was only detected by the object detection method, with no warning sign given. The red pie slice on (b) indicates the ultrasonic distance sensor data, and the blue dots on both (a) and (b) indicate the location of workers. Both data were shown on a single user interface.

The user interface and the anti-collision system were updated concurrently so that the crane operator could identify hazards resulting from human error or equipment failure in real-time. The five nearest objects detected by ultrasonic sensors and the positions of the five workers with the highest object detection accuracy were shown in the bottom left frame. Both datasets were updated continuously and independently using different frequencies. When an object or a worker was detected within a 2 m radius of the crane

hook, the interface alerted the crane operator with the warning message displayed behind the row of data, respectively.

An interview was conducted with the involved crane operators to collect their feedback on the proposed collision avoidance system based on their first-hand experience. Two crane operators were asked to rank the computer vision system, the multi-ultrasonic sensor system, and the overall system separately based on how effective they perceived the system to be for safe lifting operations. Table 12 shows the results of the interview.

Table 12. Crane operator’s feedback on the proposed system.

Criteria	Score	Additional Comment
Computer vision system	6	Helped in tracking the workers; however, choppy video due to low frame rate bring irritating experience and there is still a room for improvement in its accuracy.
Multi-ultrasonic sensors system	6	Measured distance in all directions was helpful in detecting workers and other obstacles; however, the detectable distance was somewhat too short.
Computer vision system + multi-ultrasonic sensors system	7	Integrating two techniques together raised the reliability of this collision avoidance system.
Overall system	8	Wireless architecture eased the installation of devices. Clear GUI made important information obvious and was user-friendly.

Score: From 1 to 10, where 10 is most satisfying and 1 is most unsatisfying.

5. Discussion

5.1. Challenges and the Solutions

This study presents the implementation of computer vision and multi-ultrasonic sensors used in crane lifting operation monitoring for collision avoidance. Although the field testing and evaluation showed promising results, some challenges were identified and corresponding solutions suggested so as to enhance accuracy and user-friendliness.

First, the ultrasonic sensors used for laboratory and field testing were HC-SR04 and STMA-503, respectively. Regardless of the type, ultrasonic sensors have the common limitation of creating acoustical crosstalk when two or more ultrasonic sensors with the same frequency are located close to each other. To reduce the effects of wave interference on the validity of the data, the time interval of ultrasonic wave emissions from each sensor was set at 65 ms.

In the field experiment, a few ultrasonic sensors were aligned on every side of the crane hook. When an object is moving near the hook, or vice versa, the sensors detect the approaching object, and the crane operator will receive a warning when it is inside the prohibited area. Nevertheless, if the object is extremely close to the sensors but beyond their measuring coverage area, which is also called the blind zone, the object would not be detected by the sensors. In this situation, the utilization of computer vision can highly reduce the blind spot, as mentioned previously.

The proposed method estimates the adjacent distance of workers to the camera by using the perimeter of the bounding box drawn on the detected workers with the computer vision algorithm. However, this method does not provide the tangent distance or the opposite distance between the worker and the crane hook. When the adjacent distance of the worker to the hook is within the predefined crane-monitoring radius of 2 m, a collision warning will be given to the crane operator. To solve this problem, when a worker is detected by the object detection algorithm, the result is double-checked using the distance data provided by the respective ultrasonic sensor facing in the direction of the worker. In the event of a discrepancy in both results, the collision warning is expected to be a false positive result.

Another challenge was caused by the common limitations of computer vision algorithms. An object detection model is trained to detect the presence and location of one or multiple classes of objects. The common convolutional neural network is trained with a

large amount of labeled image data. However, the artificial neural network used in this study is unlike the human brain and is not designed for interpreting, associating, and analyzing complex data when several pieces of information are given. For example, a child is taught to identify different colors and humans. After learning these two pieces of information, he/she is able to recognize a human even though the helmet worn is different in color. However, a common artificial neural network does not work this way. Variation in size, brightness, color, direction of view, and the completeness of the object in view will altogether affect the decision of the artificial neural network in recognizing an object. Since the core concept of this study is to build a collision avoidance system by implementing computer vision techniques and multi-ultrasonic sensors, an artificial neural network that may solve this limitation was not proposed and applied, although the object detection model was trained with 8,470 images with a total of 26,655 labeled workers within those images wearing safety helmets in different colors, but most were red.

5.2. Future Work

Besides the aforementioned challenges, several known limitations will be solved in future work. The method adopted to predict the distance between a detected worker and the camera on the crane trolley was based on the perimeter of the bounding box drawn on the worker. Referring to Table 12, this method can provide an estimation of distance with a relatively large range. However, when a distance estimation with high accuracy is required, the marker should be an object that is static in the real environment rather than a dynamic bounding box that slightly varies in size and shape constantly. Two alternative methods will be explored in future work: (1) applying the edge detection method within the region inside the bounding box drawn on the worker, then estimating the distance of the worker to the camera by using the radius of the helmet; and (2) utilizing the same technique, but the target is defined as the body of the crane itself. Since the body size of the crane is always constant, the vertical distance between the camera installed on the crane trolley and the ground can be calculated with the known length of the crane body as long as the crane body is within the field of view. However, since these two methods require a highly accurate edge detection result, additional steps will be carried out to train the model to recognize the marker and measure the distance between two predefined points. In addition, the calibration has to be carried out when cameras of different focal lengths, sizes of regions of interest, and sensor sizes are used without applying the zoom function. Even with an additional distance measurement sensor with high accuracy installed at the same level of the camera to measure the vertical distance between the camera and the ground, those problems mentioned above could not easily be solved since the estimation of horizontal distance is still left unsolved. Currently, the stereovision depth method still has limitations in long-range estimation.

An ultrasonic sensor works well on targets with large, flat, and solid surface materials, while a soft or irregular surface will deflect the wave. To achieve a result with high validity, the target should always be perpendicular to the ultrasonic sensor, which is an extremely difficult condition to satisfy at a construction site. As such, the future work will extend the current study by fusing the 3D LiDAR technique. Specifically, the addition of 3D LiDAR will augment the collision avoidance system by obtaining the position of lifted loads, workers, surrounding objects, and structures, regardless of whether they are static or dynamic. With the addition of 3D LiDAR, which will further verify the results obtained through computer vision and ultrasonic sensors, future work will strive to improve the accuracy, validity, and reliability of the current work.

6. Conclusions

This study proposes and presents a crane lifting-load collision avoidance system to ensure the safety of workers on construction sites, especially those working near the crane operation area, and to protect the lifted load from damage due to a collision. In order to ensure every static and dynamic object exposed to potential collision hazards is detected,

computer vision-based object detection and distance measurement sensors were combined as an integrated system. Object detection was trained to detect the presence of workers, while the distance measurement sensors were configured to measure and provide the location of all objects or structures near the hook block. A graphical user interface was created to provide the crane operator with real-time visualization of the objects' location data present in the crane operation area. This information shown on the GUI monitor can assist the crane operator in keeping track of any object within the crane operation area and identifying potential collision hazards. Laboratory experiments and field testing with the prototype model were conducted to verify the soundness and validity of the proposed system.

The key tasks carried out and the significance of this study can be summarized as follows: (a) analyzed the crane-related collision risk factors and reviewed the previous research regarding the collision avoidance system; (b) developed a prototype model for the collision avoidance system with integration of computer vision and multi-ultrasonic sensors; (c) designed a unique graphical user interface for crane operators that can visualize the location information of objects around the crane hook; (d) identified several technical challenges, such as the mechanism of ultrasonic waves, the accuracy of the distance prediction using computer vision technique, and the learnability of an artificial neural network model that can affect the system performance; (e) optimized the proposed system by devising solutions, such as placing a time interval between wave emissions and double-checking the result of computer vision-based distance prediction by processing the ultrasonic sensor data.

Author Contributions: Conceptualization, Y.P.Y.; Data curation, Y.P.Y., S.J.L. and Y.H.C.; Formal analysis, Y.P.Y., S.J.L. and Y.H.C.; Investigation, Y.P.Y., S.J.L., Y.H.C. and K.H.L.; Methodology, Y.P.Y. and S.W.K.; Resources, Y.P.Y., S.J.L., Y.H.C. and K.H.L.; Software, Y.P.Y.; Supervision, K.H.L. and S.W.K.; Writing—original draft, Y.P.Y.; Writing—review & editing, C.S.C. and S.W.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the Korea Agency for Infrastructure Technology Advancement (KAIA) grant funded by the Ministry of Land, Infrastructure and Transport (Grant RS-2020-KA158109). This work is financially supported by the Korean Ministry of Land, Infrastructure, and Transport (MOLIT) under the “Innovative Talent Education Program for Smart City”.

Data Availability Statement: The data presented in this study are available on request from the corresponding author and the first author. The data are not publicly available due to confidentiality reasons.

Acknowledgments: This work is supported by the Korea Agency for Infrastructure Technology Advancement and the Korean Ministry of Land, Infrastructure, and Transport (MOLIT).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ministry of Land, Infrastructure and Transport, Republic of Korea. *Current Status of Major Statistics for Construction Equipment*. Construction Policy Bureau, Seoul; MOLIT: Sejong, Republic of Korea, 2022.
2. Ministry of Land, Infrastructure and Transport, Republic of Korea. *Final Report on the Development of Innovative Technology for the Apartment Housing Development System Based on Off-Site Construction*. Korea Agency for Infrastructure Technology Advancement; MOLIT: Sejong, Republic of Korea, 2020.
3. Cao, X.; Li, X.; Zhu, Y.; Zhang, Z. A comparative study of environmental performance between prefabricated and traditional residential buildings in China. *J. Clean. Prod.* **2015**, *109*, 131–143. [[CrossRef](#)]
4. Boyd, N.; Khalfan, M.M.; Maqsood, T. Off-Site Construction of Apartment Buildings. *J. Archit. Eng.* **2013**, *19*, 51–57. [[CrossRef](#)]
5. Yun, W.; Bae, B.; Kang, T. Comparison of Construction Cost Applied by RC and PC Construction Method for Apartment House and Establishment of OSC Economic Analysis Framework. *Korean J. Constr. Eng. Manag.* **2022**, *23*, 20–42.
6. Jeong, G.; Kim, H.; Lee, H.; Park, M.; Hyun, H. Analysis of safety risk factors of modular construction to identify accident trends. *J. Asian Archit. Build. Eng.* **2021**, *21*, 1040–1052. [[CrossRef](#)]
7. Ministry of Employment and Labor, Republic of Korea. *Tower Crane Safety Operation Manual*. Korea Occupational Safety and Health Agency; MOEL: Sejong, Republic of Korea, 2021.

8. Akinosho, T.D.; Oyedele, L.O.; Bilal, M.; Ajayi, A.O.; Delgado, M.D.; Akinade, O.O.; Ahmed, A.A. Deep learning in the construction industry: A review of present status and future innovations. *J. Build. Eng.* **2020**, *32*, 101827. [[CrossRef](#)]
9. Kolar, Z.; Chen, H.; Luo, X. Transfer learning and deep convolutional neural networks for safety guardrail detection in 2D images. *Autom. Constr.* **2018**, *89*, 58–70. [[CrossRef](#)]
10. Rashid, K.M.; Louis, J. Times-series data augmentation and deep learning for construction equipment activity recognition. *Adv. Eng. Inform.* **2019**, *42*, 100944. [[CrossRef](#)]
11. Hernandez, C.; Slaton, T.; Balali, V.; Akhavian, R. A Deep Learning Framework for Construction Equipment Activity Analysis. In Proceedings of the ASCE International Conference on Computing in Civil Engineering 2019, Atlanta, Georgia, 17–19 June 2019; pp. 479–486.
12. Slaton, T.; Hernandez, C.; Akhavian, R. Construction activity recognition with convolutional recurrent networks. *Autom. Constr.* **2020**, *113*, 103138. [[CrossRef](#)]
13. Sherafat, B.; Rashidi, A.; Song, S. A Software-Based Approach for Acoustical Modeling of Construction Job Sites with Multiple Operational Machines. In Proceedings of the Construction Research Congress 2020, Tempe, AZ, USA, 8–10 March 2020; pp. 886–895.
14. Fang, Q.; Li, H.; Luo, X.; Ding, L.; Luo, H.; Rose, T.M.; An, W. Detecting non-hardhat-use by a deep learning method from far-field surveillance videos. *Autom. Constr.* **2018**, *85*, 1–9. [[CrossRef](#)]
15. Wu, J.; Cai, N.; Chen, W.; Wang, H.; Wang, G. Automatic detection of hardhats worn by construction personnel: A deep learning approach and benchmark dataset. *Autom. Constr.* **2019**, *106*, 102894. [[CrossRef](#)]
16. Nath, N.D.; Behzadan, A.H.; Paal, S.G. Deep learning for site safety: Real-time detection of personal protective equipment. *Autom. Constr.* **2020**, *112*, 103085. [[CrossRef](#)]
17. Zhang, H.; Yan, X.; Li, H. Ergonomic posture recognition using 3D view-invariant features from single ordinary camera. *Autom. Constr.* **2018**, *94*, 1–10. [[CrossRef](#)]
18. Yu, Y.; Li, H.; Yang, X.; Kong, L.; Luo, X.; Wong, A.Y. An automatic and non-invasive physical fatigue assessment method for construction workers. *Autom. Constr.* **2019**, *103*, 1–12. [[CrossRef](#)]
19. Yang, K.; Ahn, C.R.; Kim, H. Deep learning-based classification of work-related physical load levels in construction. *Adv. Eng. Inform.* **2020**, *45*, 101104. [[CrossRef](#)]
20. Luo, H.; Xiong, C.; Fang, W.; Love, P.E.; Zhang, B.; Ouyang, X. Convolutional neural networks: Computer vision-based workforce activity assessment in construction. *Autom. Constr.* **2018**, *94*, 282–289. [[CrossRef](#)]
21. Luo, X.; Li, H.; Cao, D.; Dai, F.; Seo, J.; Lee, S. Recognizing Diverse Construction Activities in Site Images via Relevance Networks of Construction-Related Objects Detected by Convolutional Neural Networks. *J. Comput. Civ. Eng.* **2018**, *32*, 04018012. [[CrossRef](#)]
22. Yang, J.S.; Huang, M.L.; Chien, W.F.; Tsai, M.H. Application of Machine Vision to Collision Avoidance Control of the Overhead Crane. In Proceedings of the 2015 International Conference on Electrical, Automation and Mechanical Engineering, Phuket, Thailand, 26–27 July 2015; pp. 361–364.
23. Fang, Y.; Chen, J.; Cho, Y.K.; Kim, K.; Zhang, S.; Perez, E. Vision-based load sway monitoring to improve crane safety in blind lifts. *J. Struct. Integr. Maint.* **2018**, *3*, 233–242. [[CrossRef](#)]
24. Price, L.C.; Chen, J.; Park, J.; Cho, Y.K. Multisensor-driven real-time crane monitoring system for blind lift operations: Lessons learned from a case study. *Autom. Constr.* **2021**, *124*, 103552. [[CrossRef](#)]
25. Gu, L.; Kou, X.; Jia, J. Distance measurement for tower crane obstacle based on multi-ultrasonic sensors. In Proceedings of the 2012 IEEE International Conference on Automation Science and Engineering, Seoul, Republic of Korea, 20–24 August 2012; pp. 1028–1032.
26. Xi-Ping, L.; Li-Chen, G.; Jia, J. Anti-collision method of tower crane via ultrasonic multi-sensor fusion. In Proceedings of the International Conference on Automatic Control and Artificial Intelligence (ACAI 2012), Xiamen, China, 3–5 March 2012; pp. 522–525.
27. Sleiman, J.; Zankoul, E.; Khoury, H.; Hamzeh, F. Sensor-Based Planning Tool for Tower Crane Anti-Collision Monitoring on Construction Sites. In Proceedings of the Construction Research Congress 2016, San Juan, Puerto Rico, 31 May–2 June 2016; pp. 2624–2632.
28. Zhong, D.; Lv, H.; Han, J.; Wei, Q. A Practical Application Combining Wireless Sensor Networks and Internet of Things: Safety Management System for Tower Crane Groups. *Sensors* **2014**, *14*, 13794–13814. [[CrossRef](#)] [[PubMed](#)]
29. Lee, G.; Cho, J.; Ham, S.; Lee, T.; Lee, G.; Yun, S.; Yang, H. A BIM- and sensor-based tower crane navigation system for blind lifts. *Autom. Constr.* **2012**, *26*, 1–10. [[CrossRef](#)]
30. Kim, M.; Yoon, I.; Kim, N.; Park, M.; Ahn, C.; Jung, M. A Study on Virtual Environment Platform for Autonomous Tower Crane. *Korean J. Constr. Eng. Manag.* **2022**, *3*, 3–14.
31. Rosén, E.; Sander, U. Pedestrian fatality risk as a function of car impact speed. *Accid. Anal. Prev.* **2009**, *41*, 536–542. [[CrossRef](#)] [[PubMed](#)]

32. Kang, Y.; Siddiqui, S.; Suk, S.J.; Chi, S.; Kim, C. Trends of Fall Accidents in the U.S. Construction Industry. *J. Constr. Eng. Manag.* **2017**, *143*, 04017043. [[CrossRef](#)]
33. Ministry of Employment and Labor, Republic of Korea. *Status of Industrial Accidents Occurrence at the End of December 2020*. Korea Occupational Safety and Health Agency; MOEL: Sejong, Republic of Korea, 2021.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.