applied sciences

MDPI

*Article*

# Design and Validation of a Camera-Based Safety System for Fenceless Robotic Work Cells

**Merdan Ozkahraman [1,*], Cuneyt Yilmaz [2] and Haydar Livatyali [2]**

[1] Department of Mechatronics Engineering, Faculty of Technology, Isparta University of Applied Sciences, Isparta 32260, Turkey
[2] Department of Mechatronics Engineering, Faculty of Mechanical Engineering, Yıldız Technical University, Istanbul 34220, Turkey; cuneyt@yildiz.edu.tr (C.Y.); hlivatya@yildiz.edu.tr (H.L.)
* Correspondence: merdanozkahraman@isparta.edu.tr; Tel.: +90-2462146822

**Abstract:** A two-dimensional (2-D) camera system with a real-time image processing-based safety technology is a cost-effective alternative that needs optimization of the cell layout, the number of cameras, and the camera's locations and orientations. A design optimization study was performed using the multi-criteria linear fractional programming method and considering the number of cameras, the resolution, as well as camera positions and orientations. A table-top experimental setup was designed and built to test the effectiveness of the optimized design using two cameras. The designs at optimal and nonoptimal parameters were compared using a deep learning algorithm, ResNet-152. To eliminate blind spots, a simple but novel 2-D image merging technique was proposed as an alternative to commonly employed stereo imaging methods. Verification experiments were conducted by using two camera resolutions with two graphic processors under varying illuminance. It was validated that high-speed entrances to the safety system were detected reliably and with a 0.1 s response time. Moreover, the system was proven to work effectively at a minimum illuminance of 120 lux, while commercial systems cannot be operated under 400 lux. After determining the most appropriate 2-D camera type, positions, and angles within the international standards, the most cost-effective solution set with a performance-to-price ratio up to 15 times higher than high-cost 3-D camera systems was proposed and validated.

**Keywords:** robotics; optimization; work-cell safety system; 2-D cameras

check for updates

**Citation:** Ozkahraman, M.; Yilmaz, C.; Livatyali, H. Design and Validation of a Camera-Based Safety System for Fenceless Robotic Work Cells. *Appl. Sci.* **2021**, *11*, 11679. https://doi.org/10.3390/app112411679

Academic Editor: Alessandro Gasparetto

Received: 15 November 2021
Accepted: 7 December 2021
Published: 9 December 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Copyright:** © 2021 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

As a result of rapid developments, new technologies are being introduced to the industry every day. Rapidly changing the technology and different customer preferences affect the production lines and create predictable or unpredictable changes in final products. Therefore, production systems are required to adapt to these changes [1]. However, flexibility in production is not natural to implement when considering mass production costs. Various production systems have been put forward, such as reconfigurable, fenceless, and hybrid systems [2,3]. Hybrid production systems involve human–robot cooperation, and they are also called collaborative systems, in which the human factor provides flexibility in the production to maximize operational efficiency [4].

Fenceless robotic production systems have been emerging for a couple of decades, and they are still being developed. As these systems consist of work cells without any separators between robot workspaces and neighboring fields, some active precautions are necessary to achieve tasks and workers' safety requirements. A fenceless and robot-assisted human-guided production system has significant advantages over full automated ones covered with fences. The most critical obstacle for a successful application in production is the inefficiency caused by the safety requirements [5]. Three-dimensional (3-D) cameras are the most common in image processing-based safety systems according to the current standards [6,7]. High-cost safety systems using a single 3-D camera have a significant

disadvantage due to the lack of intervention in blind spots. A system using multiple low-cost, two-dimensional (2-D) cameras instead of a 3-D camera may solve this problem. However, it is crucial to make the most suitable choice for the cameras' placements and orientations to scan a broader area eliminating blind spots.

Currently, the standards described in "ISO 10218: Robots for Industrial Environments—Safety Requirements" [8] apply to the primary robot cells' safety operations [9]. If the distance between the robot and the human is below a certain safety level, the robot arm is required to slow down and standstill. Thus, manipulators are prevented from hitting workers. In the designed safety system, there are three safety areas, which are the warning, alarm, and working zones. The robot arm is required to decelerate and stop when entries into the warning and alarm zones are detected, respectively. This study aims to design a safety system using multiple low-cost 2-D cameras for robotic work cells considering current international standards. It proposes an optimal calculation method of camera placements and a design of a safety camera system for fenceless robotic production fields.

After this brief introduction, some background information is summarized. Design and optimization of the proposed system are followed by a description of the image processing method to avoid blind spots. Next, testing of the optimized design using a deep learning algorithm is presented. The experimental setup design section is followed by results and discussions of the validation tests. The article ends with some concluding remarks.

## 2. Background

In this section, safety standards in work cells using robot arms, shortcomings of existing safety systems, comparison of known sensor-based and camera-based safety systems, and some proposed solutions are summarized. Safety standards applied in robotic work cells have been an active research topic in the past few decades. Considering work cells using robot arms, the most critical factor is ensuring work and worker safety [10]. Some studies combine various design approaches for the hardware and software of safety-related systems [11,12].

ISO 12100: 2010 Machine Safety—General Principles for Design—Risk Assessment and Risk Reduction Standard [13] describes a typical workflow that focuses on the safety of the work cell. This workflow can be listed as explaining the system, identifying the hazards, and determining the emerging risks and risk reduction strategies. In addition to the documentation of the system and the associated risks, it is also necessary to validate the risk reduction strategies. A study by Bessler et al. was conducted within the framework of these standards [14], which provides guidance for the design and implementation of a robotic safety system.

There are many commercially developed safety systems for work cells using industrial robot arms. These can be classified into sensor- and camera-based systems. Sensor-based ones are mostly preferred in simpler and controlled industrial environments because of disruptive factors such as industrial ambient noise that affects measurements' sensitivity and accuracy. These effects are also valid for camera-based safety systems. Developing camera-based safety systems with software technology has a considerable advantage of supporting flexibility in production. Consequently, various camera image-based safety systems are available in the industry [15]. Considering that camera technology is developing day by day, it is predicted that it will be a more practical and widely used choice in the future, as compared with sensor-based safety systems [16].

In the past few decades, IP cameras have been used to ensure safety in robotic work cells in which monitoring systems are based on the human body's posture and position [17]. Another approach is tracking the position and speed of robot arms' end effectors [18]. There are also commercial systems that visually track undesirable objects in work cells [19]. Knowing these systems and with developments in camera technology, various systems using 2-D cameras have been proposed [20]. Some commercial systems have been developed by companies such as Kuka, Reis, ABB, Fanuc, and Pilz. The one by Pilz is a safety system called SafetyEYE, in which a 3-D camera is used [3]. This system works with the principle

of slowing down and stopping the robot arm when there is an undesired entry into the scanned volume of the work cell in 3-D. However, using one camera, it is believed that the robot arm may create a blind spot depending on the working angles. These blind spots affect system reliability. In addition, because additional 3-D cameras increase the cost, it is a necessity to introduce a more cost-effective and high-reliability system.

Machine vision (MV) is an imaging-based application whose technology is used in the field of industrial automation [21]. In these systems, where software and hardware work in an integrated manner, images can be used in areas such as guidance, control, and safety of robot arms. While the input and output are both images in image processing, in MV technology, the output is an information or command corresponding to the image input. MV automatically converts the information it extracts from the image into a command. Thus, a camera-based safety system, using either 2-D or 3-D cameras and no artificial-intelligence-based image characterization is considered within the framework of MV.

Consequently, there have not been any studies to find the most appropriate number of cameras to cover the blind spots in work cells. Instead of using well-established but costly 3-D cameras, more affordable 2-D cameras are considered a viable option. A 2-D camera-based safety system design was performed using a basic optimization method considering the number of cameras, the resolution, as well as camera positions and orientations.

With the availability of multiple cameras, which are provided to monitor the work cells at various angles, robot arms' safe working areas can be continuously monitored. The worker's safety is ensured by slowing down and then stopping the robot arm according to the directives on entry of undesired foreign objects given in the safety standards. In the following sections of this paper, various formulations were made according to a scaled-down robot arm model, and a scenario created was tested by using the most suitable 2-D cameras. The results were compared with the ones from alternative 3-D camera systems.

## 3. Design of the Image-Based Safety System

### *3.1. Locating the Camera(s)*

In the design of the safety system, the dimensional parameters of a work cell are primarily considered based on the safety distances per the standards. The number of cameras, camera lens angle, camera orientation angle, and camera height values are considered in the optimization problem. The safety standards used in this study are shown in Figure 1.

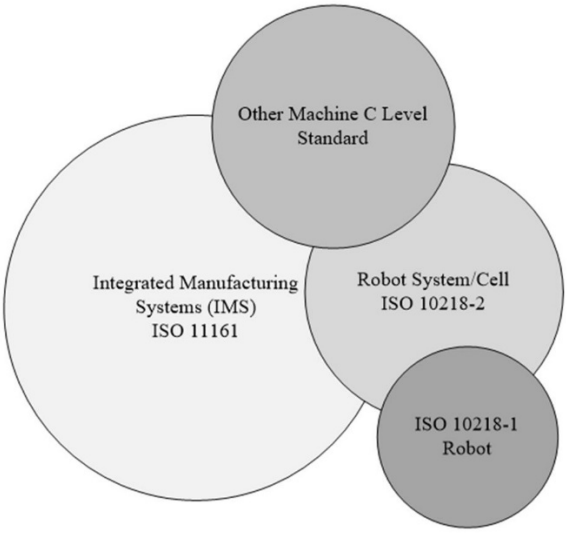



**Figure 1.** Schematic representation of the relationships between the robotic systems and related international standards [8].

Considering the standards (EN ISO 13849-1: 2008 and EN ISO 13849-2: 2013) [22], collaboration modes can be defined to ensure human and robot collaboration in a safe manner. These include the following modes:

- Safe posture;
- Safe speed and distance monitoring;
- Manual robot orientation [23,24].

Safety distance (S) in work cells with robot arms defined in the related standards is stated as

$$S = K(t_1 + t_2) + C \quad (1)$$

where $t_1$ is the reaction time of the measurement, $t_2$ is the reaction time of the machine, C is the possible approach to the alarm zone, and K is the estimated approach speed of the human body or body parts. The ISO 13855 standard identifies K with a walking speed of 1600 mm/s and a hand speed of 2000 mm/s [13].

The optimization problem is considered a linear fractional function. The properties of linear fractional functions facilitate the usage of a well-known scalar optimization problem. If the constraints are linear, the linear fractional programming (LFP) function is obtained as follows [25]:

$$\max \left[ f(x) = \frac{p(x)}{q(x)} \right] \text{ such that } x \in S \subset R^n \quad (2)$$

Figure 2 shows a sketch of a system consisting of two cameras and one robot planned to be realized. In addition, Figure 2 shows three safety areas, which are the warning, alarm, and working zones. The robot arm is required to decelerate and stop when entries into the warning and alarm zones are detected, respectively.

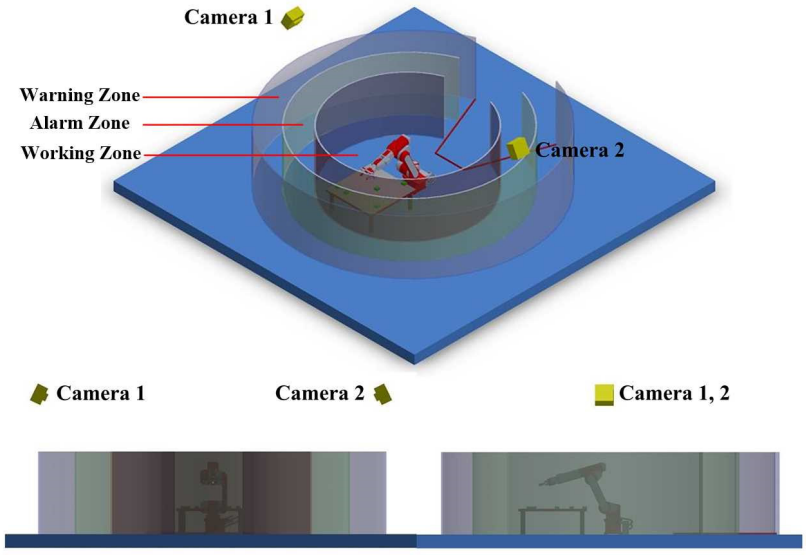



**Figure 2.** Illustration of a system with two cameras and one robot (**upper**) and front (**lower left**) and side (**lower right**) views.

In addition, the visualization of the projected areas of these cameras is shown with parametric expressions in Figure 3. The outermost edge of the scanning area of the cameras is placed so that it intersects the outermost edge of the safety area. The area scanned by the cameras is geometrically in the form of an angular truncated cone. In Figure 3, the camera lens angle is given as $2\alpha$. The angle of the camera with the horizontal axis (orientation angle) is given as $\theta$. The height of the camera from the ground is shown as h.

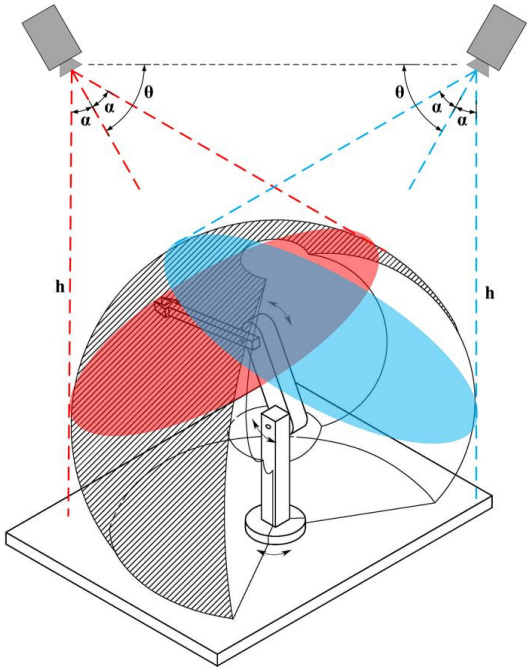



**Figure 3.** The intersection of projections in a two-camera system.

The projection of a camera with specific angle parameters ($\alpha$ and $\theta$) to a working space of a robot is a truncated cone base area (Figure 4). The major and minor axis (a and b) of truncated cone base geometry is given in Equations (3) and (4), respectively [26]. Depending on these two equations, the area scanned by the camera can be expressed in a parametric function with two variables, a and b.

$$a = h(\sin 2\alpha) / \left(\sin^2 \theta - \sin^2 \alpha\right) \quad (3)$$

$$b = 2h(\sin \alpha) / \sqrt{\left(\sin^2 \theta - \sin^2 \alpha\right)} \quad (4)$$

where $\alpha$ is the one-half of the camera's view angle, $\theta$ is the camera's orientation angle concerning the x-axis, and h is the camera's vertical position (Figure 4). The image formed on the camera sensor is rectangular. Therefore, the image outside the circular area is removed by masking in the program.

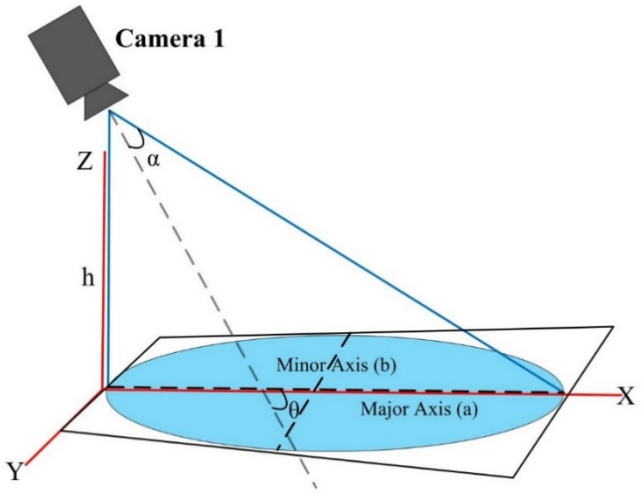



**Figure 4.** Display of skewed cone parameters.

The camera's scanning depends on camera lens angle, camera orientation angle, and camera height as the variables. The area scanned by the camera is the objective function

that is intended to be maximized. The camera height (h) is the optimization constraint. The objective function to maximize the scanned area ($A_{max}$) is

$$A_{max} = \pi h^2 \sin(\alpha) \sin(2\alpha) / \left(2\left(\sin^2(\theta) - \sin^2(\alpha)\right)^{3/2}\right) \quad (5)$$

In a 3000 mm diameter workspace, the camera projection area is expected to be larger not exceeding 4000 mm in diameter; that is, the first constraint is that the major axis (a) should be longer than 3000 mm and less than 4000 mm. According to the working environment's constraints, the camera placement height is expected to remain below a certain distance, and it was assumed as 10,000 mm as the second constraint for the experimental study. The third constraint is that the camera lens angle ($2\alpha$) should be between 30° and 70°. The extremum points of a continuous objective function are the points where the derivative is zero. For f($\alpha$, h), the function's derivatives are taken with respect to $\alpha$ and h. The solution set for this point is shown in Equation (6).

$$\begin{gathered} \alpha_1 = \pi n,\ n \in Z \\ \alpha_2 = \pi n - 1/2(\cos^{-1}(-3)),\ n \in Z \\ \alpha_3 = \pi n + 1/2(\cos^{-1}(-3)),\ n \in Z \end{gathered} \quad (6)$$

Since $\alpha$ values 0 and over $\pi$ obtained from the equalities of $\alpha_1$ and $\alpha_3$ are not reasonable, we considered using $\alpha_2$. Therefore, the optimal $\alpha$ was obtained as 0.505 radians (28.94°), and this value represents half of the camera lens angle.

Figure 5 shows the graphical representation of the area that the objective function scans for the given constraints. In this graph, the area of the safety zone is directly proportional to h and $\alpha$. Therefore, although the yellow area shown in Figure 6 has the highest area for the safety zone, it is seen that the optimal point should be sought in this region due to the height constraints. Therefore, the optimal region for the optimal angle of the camera lens and h height values between 4000 to 10,000 mm is marked in red color.

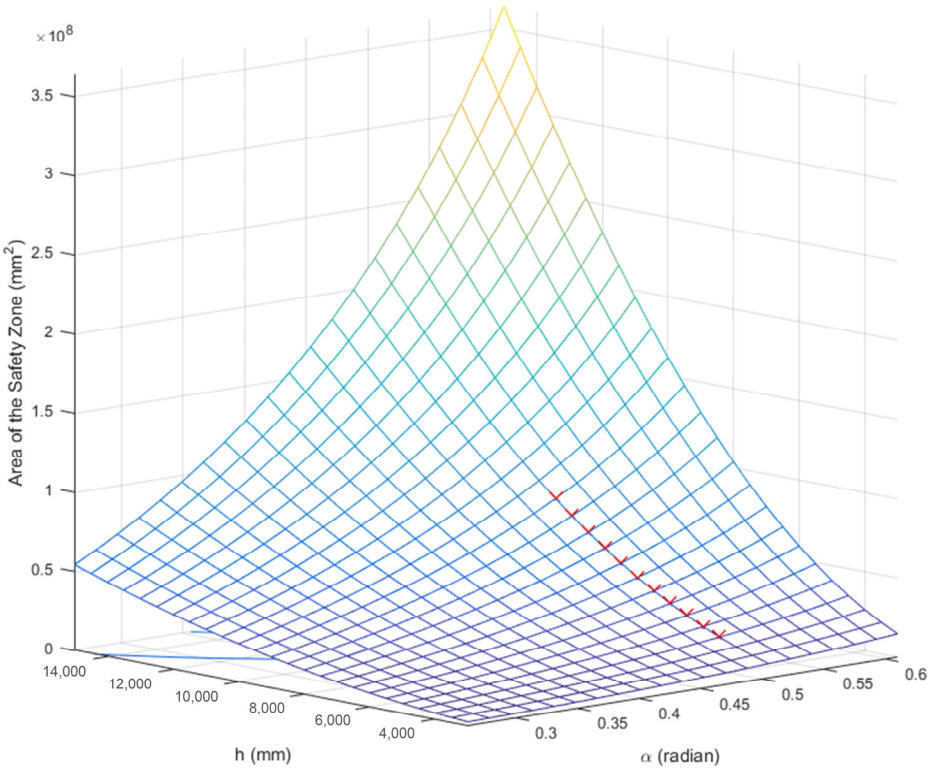



**Figure 5.** Projected area by one camera versus camera height and semi-lens angle.

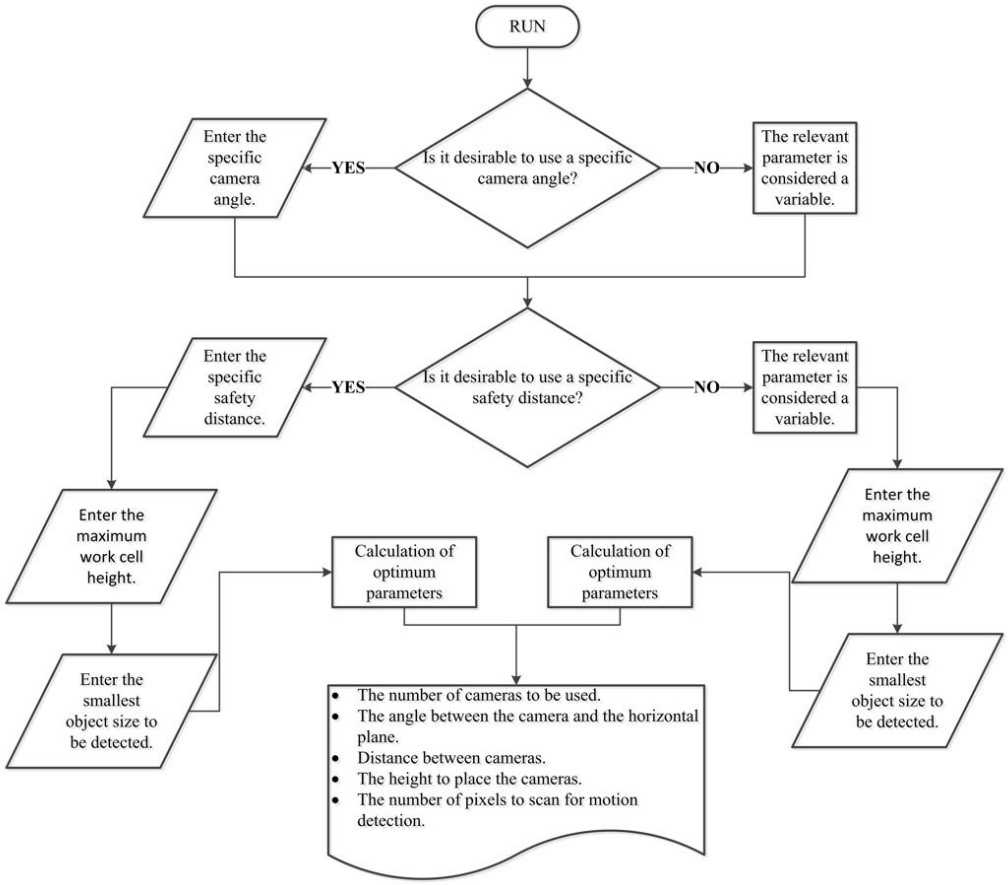



**Figure 6.** Flowchart of the algorithm to optimize the system design.

The image resolution represents the details that the image contains. This term applies to digital images, motion images, and other image types. Higher resolution means more detailed images. Image resolution can be measured in several ways. A resolution unit depends on a physical dimension, in other words, the overall size of an image. The resolution of digital cameras can be defined in various ways, such as pixel, spatial, spectral, and temporal resolutions. The pixel resolution is an important factor in the safety system because the actual portion of the image corresponding to each pixel varies with the camera height. According to international standards on digital cameras, "total pixel count" is the major specification concerning image sensors instead of the resolution. The term "registered pixels" is usually equal to the number of pixels [27,28].

The camera's projection forms an exact circle when it is placed vertically. However, since the camera is placed at a certain angle, the projection is an ellipse, and the optimization is performed based on the equation of the ellipse. The size of the detected object is critical in the safety system. Camera height and camera resolution are proportional to the number of pixels in the scanned area. Even if the dimensions of the detected object remain constant, the number of pixels that need to be scanned vary based on the camera height and resolution. Thus, the proposed mathematical model is a dynamic one that can adapt to these changes.

A generic and dynamic mathematical model including all the critical variables is needed for the image-based safety system. This system model is expected to practically adapt to various cases and constraints in the industry. The camera lens angle and the minimum object size to detect are the main input variables, the safety distance may or may not be an independent variable based on dimensions of the work cell, and the maximum camera height is the main constraint. The flowchart of the dynamic model of the proposed safety system is given in Figure 6. This relatively simple but effective algorithm produces a list of system design parameters that include the number of cameras, placement (heights

and relative distances) and angles of cameras, and the number of pixels to be scanned in the work cell.

Automated inspection, which is one of the main uses of the MV technology, was used in the study. The first step for automatic inspection is to acquire images using cameras, lenses, and lighting [29]. Then, a computer program is used to extract the necessary information. In this project, the MV algorithm was implemented in the Python programming language using the open-source "OpenCV" and "NumPy" libraries. An algorithm based on the principle of detecting the difference between sequential and real-time images taken from the camera was established. Differences between histograms of subsequent images provide information that there is movement in the projection area. Based on this information and constraints (such as the object size), the program transmits the slow-down and halt commands to the robot controller. Minimum two cameras were used in the system. Therefore, a novel image stitching method, which is another image processing operation of MV, was developed.

### 3.2. Combining Images from Multiple Cameras

In safety systems created with a single camera, it is predicted that blind areas may occur in the areas overshadowed by the robot arm. This situation can be seen in Figure 7. Therefore, it is recommended to use at least two cameras symmetrically to each other in all systems installed. Thus, blind areas are prevented.

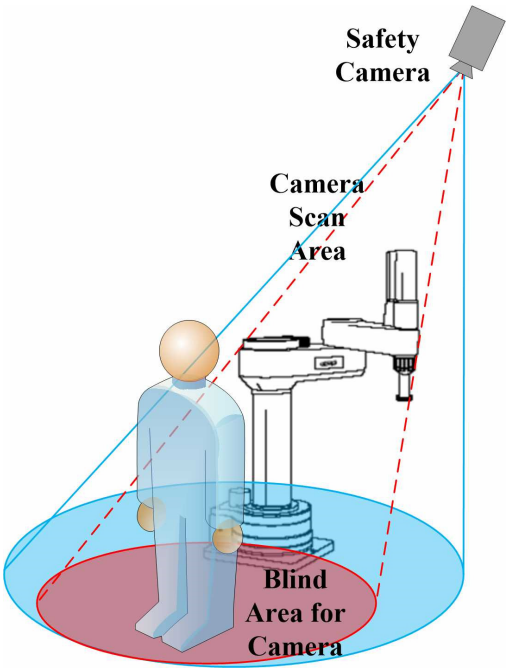



**Figure 7.** A blind area that may form when using a single camera.

One of the problems encountered in the safety system with multiple cameras is a combination of multiple images. A method of combining two images from two cameras at a distance from each other is stereo imaging, which is a relatively mature technology. Stereo imaging brings the advantage of three-dimensional perception at a significant computational cost, yet blind spots still exist. However, the proposed 2-D image stitching is computationally more economical and does not leave any blind spots. Thus, experiments were conducted to prove the elimination of blind spots in the work cell.

Another important criterion in combining multiple camera images is the formation of more than one image of an object entering the safety space when scanning the same regions. This creates complexity in determining the number of pixels to detect. The second criterion is that multiple scans of the same areas with image processing algorithms require an additional processing load for the control unit. Since the fastest possible reaction is

essential for the safety system, the additional load on the controller will be disadvantageous. Considering these criteria, partial masking of the image taken from more than one camera was proposed. For each camera, the region where there is no blind spot was processed, and the masking of the outlying zone where the blind spot could occur was performed, and thus, blind spots were avoided. As a result, processing the same zones more than once was prevented. Compared with the image-combining algorithms in the literature, a unique method that can be integrated into the safety system and work more efficiently was used. Figure 8 shows a schematic of a circular work cell placed on the image pixels of a two-camera system. Partial masking was applied to the upper half of the images taken from Camera 1 and Camera 2. A single image was produced by combining unmasked partial images with no blind spots. Figure 9 shows a more detailed 3-D perspective schematic of the image stitching proposed. The mono image developed from the unmasked portions of multiple images is naturally not photo realistic, but it is sufficient to detect a foreign object penetrating the safety zones of a work cell. When three or more cameras are used, pie slice images at 120°, 90°, 72°, etc. can also be stitched similarly.

Figure 10 shows a schematic of the monitoring system. The digital data from the cameras were processed in the image processor. As a result of the processed data, the reaction desired to be applied in the scenario was transmitted to the robot arm via the robot controller. Decelerate or stop command was sent back by the processor to the robot arm within a response time.

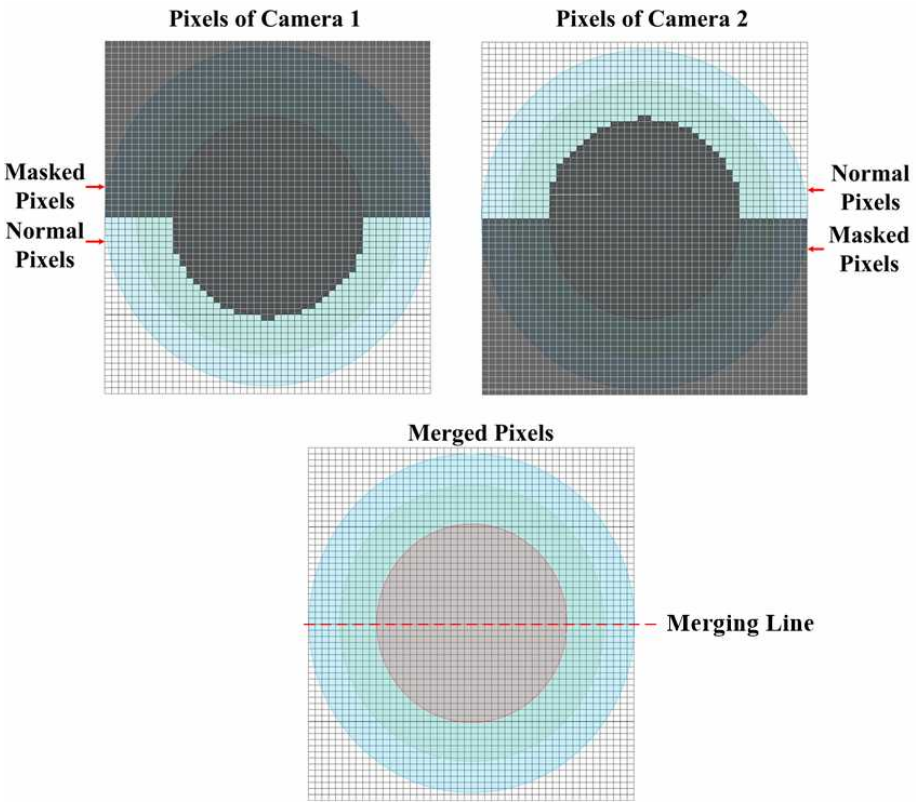



**Figure 8.** Combining partially masked images of two symmetrical cameras based on each pixel.

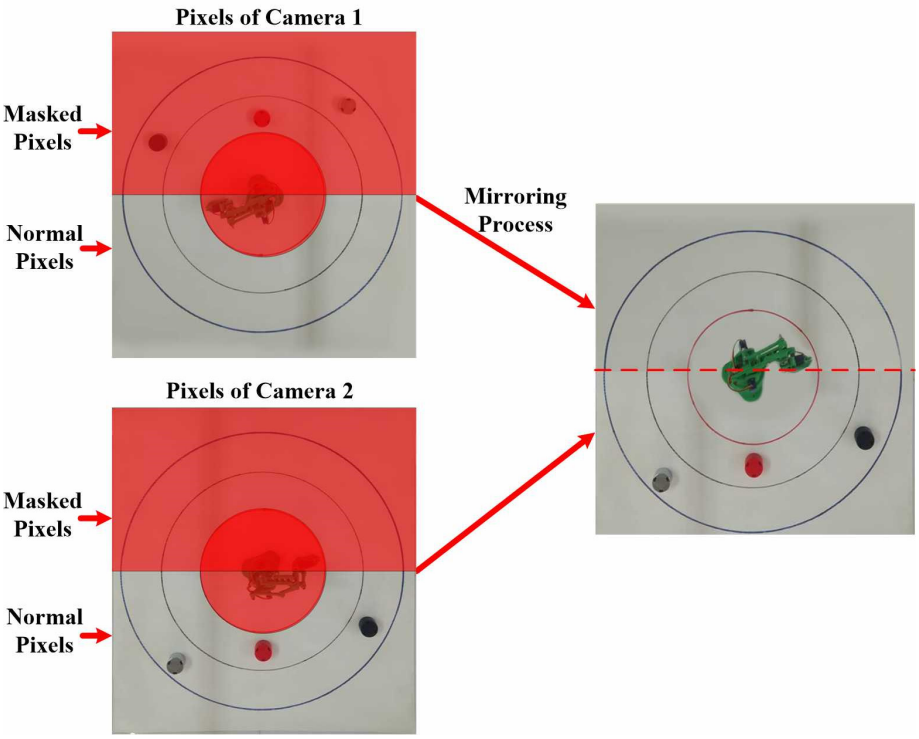



**Figure 9.** Combining different partially masked real-time camera images.

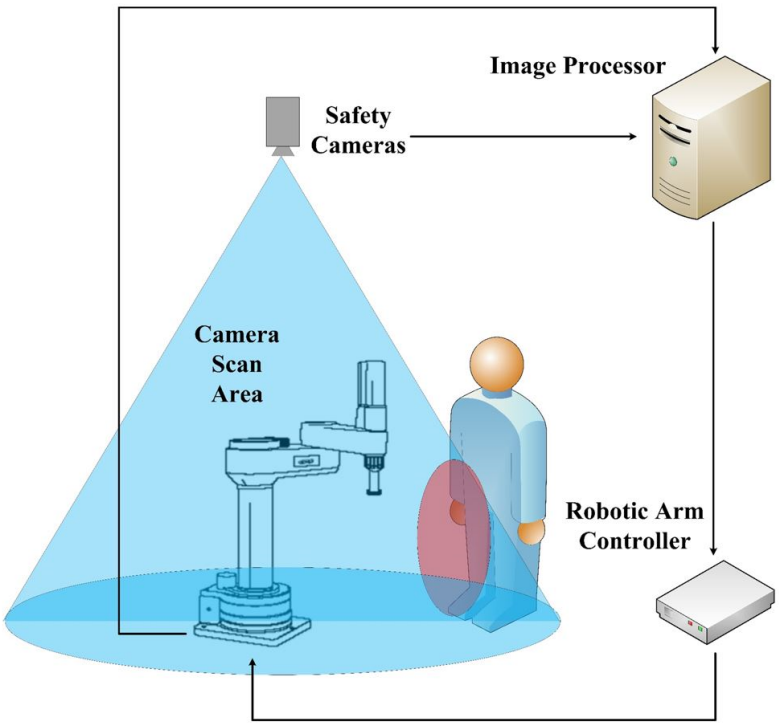



**Figure 10.** The control schematic of the safety system.

As a result of the optimization problem, a solution set or a solution point was obtained. In optimization problems involving multiple constraints, fractional programming is a beneficial method to find a practical solution set so that the decision maker can choose the most suitable one among the alternatives in the solution set.

Frame frequency (measured in frames per second, FPS) is the rate at which consecutive images, called frames, appear on display. Capturing and analyzing images from the camera

is directly related to frame frequency because it directly affects the operating time of the system. At 23 FPS, an image was taken from a camera in 0.0434 s, and the response time could not be shorter.

In industrial systems, a system with two cameras is sufficient if the maximum height at which the camera can be placed is not too low. In case of a ceiling height is not enough, the analysis suggests a three-camera system. It may not be desirable to use the optimal camera lens angle in every system. In such cases, the most suitable camera heights can be determined by analyzing the available camera angles in a mathematical model.

### *3.3. Image Characterization of the Camera(s) at Optimal and Nonoptimal Locations*

Before the validation experiments of the proposed design, the solution of the optimization problem needs to be verified. The leftmost image in Figure 11 shows the area scanned by a camera at optimal placement. The two images in the upper right show the red areas that occur with the positive and negative change of the camera orientation angle. The two images in the lower right show the red areas with the positive and negative variation of the camera height. As can be seen from the figure, there is a difference between the blue area to be scanned and the red areas formed as a result of the changes. To express this difference technically, an artificial intelligence algorithm is designed to give the similarity between the reference image and the altered images.

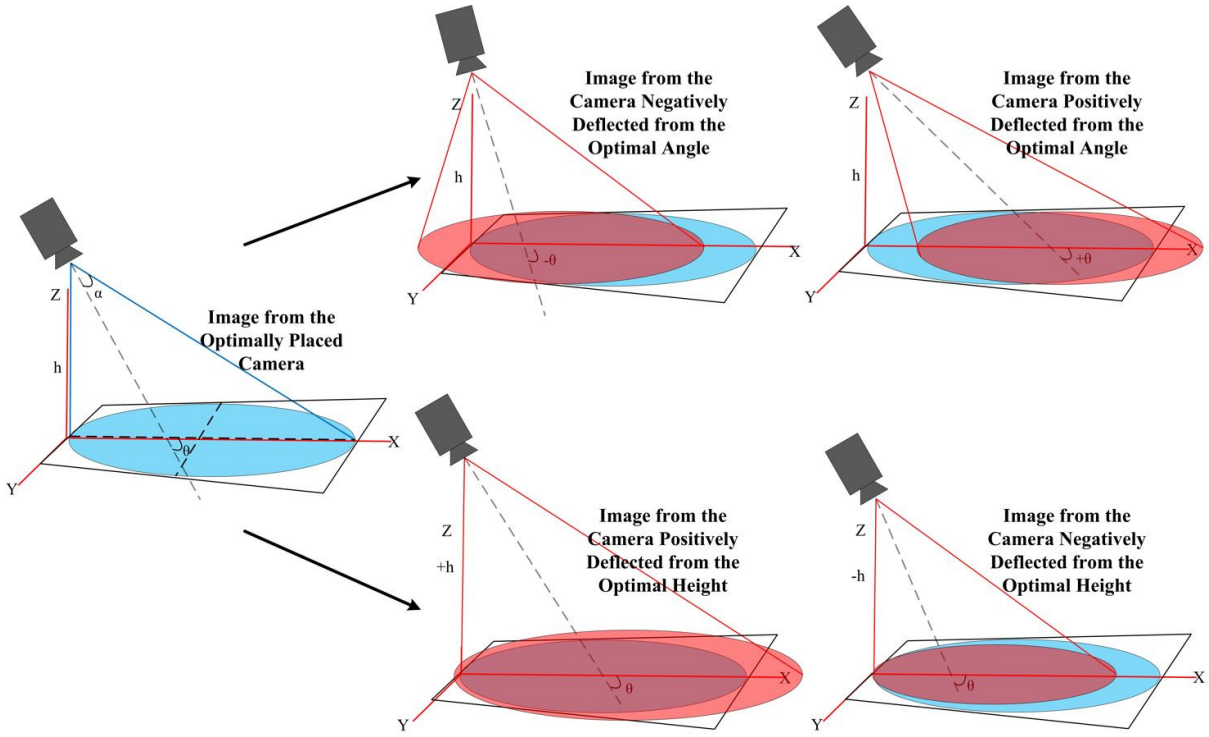



**Figure 11.** Difference of the areas scanned by the camera placed at nonoptimal orientation angle and height.

Eight different image classes were generated based on images taken from modified (increased and decreased) nonoptimal camera heights and orientation angles. About 500 images were taken from each of these 8 classes. These images were used to train the ResNet-152 deep learning algorithm.

ResNet-152 is a popular deep learning algorithm used to characterize optimal and nonoptimal image data so that the effects of the noise included in the nonoptimal data are eliminated. Thus, instead of using singular image data obtained at extremum locations and angles as a reference, the common characteristics of the images obtainable in the pre-determined range are extracted using a statistical approach. With this feature, which distinguishes the ResNet architecture from other architectures, as the number of

layers increases and the prediction accuracy increases, the optimization becomes more comfortable, and the disappearing gradient problem is solved. Therefore, this architecture is preferred for the comparison of the reference image with the trained algorithm. The aim here is to provide more realistic results by including the noisy images affected by vibration, light flashes, and pollution in the industrial environment instead of comparing a taken optimally placed camera image. Thus, it has been demonstrated that the optimal solution is the most accurate solution that completely covers the robotic work cell. The ResNet architecture provides a deeper link-hopping network added to the feed-forward network layers. Therefore, shortcut links and properties are learned by adding x (input value) to F(x) instead of learning directly from F(x) [30]. With the block approach, the output is not given directly to the block; instead, x + F(x) from the ReLu operation is feedback. The working principle of the classification and comparison algorithm is shown in Figure 12.

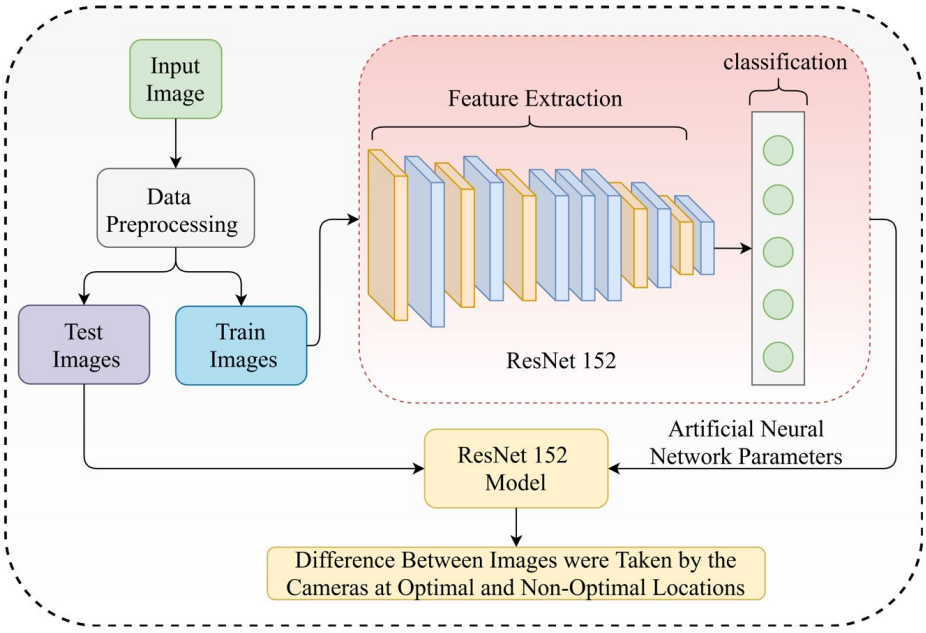



**Figure 12.** Flowchart of the deep learning algorithm.

Finally, the trained datasets in eight related classes were compared with the visuals taken at the optimal setup. As a result of this comparison, the difference between the images taken from the cameras placed at the optimal parameters is expressed as a percentage. These eight classes and differences are shown in Tables 1 and 2.

**Table 1.** Percentage of similarity between the images from the optimally placed camera with the image from the changed camera height.

| Camera Height Difference | Percentage of Similarity to Optimal Placed Camera Image |
|---|---|
| 10% Decreased | 58% |
| 5% Decreased | 60% |
| 5% Increased | 55% |
| 10% Increased | 41% |

**Table 2.** Percentage of similarity between the images from the optimally placed camera with the image from changed camera orientation angle.

| Camera Orientation Angle Deviation (Degree) | Percentage of Similarity to Optimal Placed Camera Image |
|---|---|
| 10° Decreased | 57% |
| 5° Decreased | 59% |
| 5° Increased | 51% |
| 10° Increased | 41% |

Table 1 shows that the deviation from the optimal image increases depending on the deviation of camera height. Table 2 shows that the difference from the optimal image increases depending on the increase and decrease in the camera placement angle. The referenced image is the one that excludes unwanted areas where the robot arm is visible, scanning all safety zones. Therefore, the increase in the difference to this reference means that the unwanted regions cannot be scanned, and the desired regions cannot be scanned partially. In comparing nonoptimal parameters using deep learning algorithms, the accuracy of the parameters obtained with the fundamental optimization methods was proven.

## 4. An Industrial Case Study and Experimental Verification

### *4.1. Design of the Experimental Setup*

All calculated values of $\alpha$ and h provided for design constraints are given in Figure 13, which shows the heights to be placed according to different camera lens angles graphically. The graph in the figure contains two curves. One is "Camera Height for Safety System for 2 Cameras", and the other one is "Camera Height for Safety System for 3 Cameras". In the previous calculations, it is seen that the optimal camera lens angle is 57.874°. Therefore, it was calculated that the required height for the using two cameras system is 6028 mm, and the required height for the using three cameras system is 5023 mm. For cases in which the optimal camera angle is not desired to be used, the camera heights suitable for the lens angle of the existing camera are also shown in graphs. Three cameras were needed when two cameras were not enough (i.e., the camera height is larger than 10 m). In case of the need to use three cameras, the cameras were positioned as circular and equally spaced. If three cameras were used, the total projection area and the previous solution's constraints would change. The new constraint (htan(2$\alpha$)) should be more than 6000 mm and less than 8000 mm.

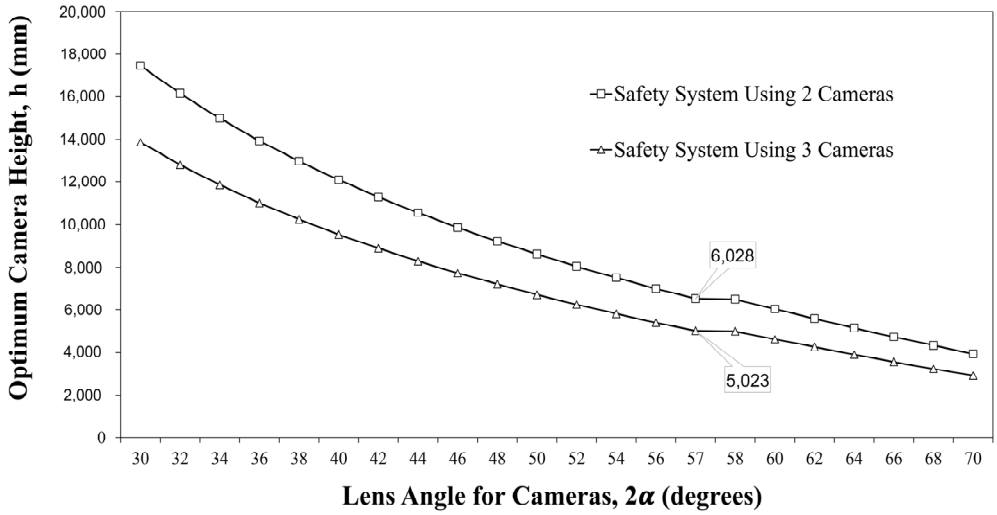



**Figure 13.** Optimum camera height versus camera lens angle for two- and three-camera systems.

A comparison between a commercial setup and the experimental one is given in Table 3. The manipulator reach, work envelope radius, maximum manipulator speed,

alarm zone radius, and warning zone radius were compared, as well as the number of cameras, camera resolution, camera angle, graphics cards, penetrating object radius, minimum illuminance range, and minimum frame per second.

**Table 3.** Specifications of a commercial example [31] and experiment setup.

| Case | Industrial | Experimental |
|---|---|---|
| No. of Manipulators | Variable | 1 |
| Manipulator Reach (mm) | 2800 | 270 |
| Work Envelope Radius (mm) | 3000 | 300 |
| Maximum Manipulator Speed (deg/s) | 260 | 500 |
| Alarm Zone Radius (mm) | 3500 | 300 |
| Warning Zone Radius (mm) | 4000 | 350 |
| No. of Cameras | 1 | 2 |
| Camera Resolution | Undisclosed | 1280 × 960 and 1920 × 1080 |
| Camera Angle (degree) | 80 | 50 and 70 |
| Graphics Cards | PSEN se AU AM3 | Nvidia Geforce GTX 1050 and GTX 850 M |
| Penetrating Object Radius (mm) | Human Body | 5 |
| Minimum Illuminance Range (lux) | 300 | 120 |
| Minimum Frame per Second | Undisclosed | 23 and 28 |

An optimized solution was presented considering the limit values and variables guided by the standards. A flexible system that can adapt to fenceless work cells at different sizes with multiple 2-D cameras was proposed. A comparative verification of the machine vision system was made. Consequently, the system operates stably at the optimal setting. Figure 14 shows the experimental setup built to test the system design. Two cameras were placed at optimal height and placement angle in the experimental setup.

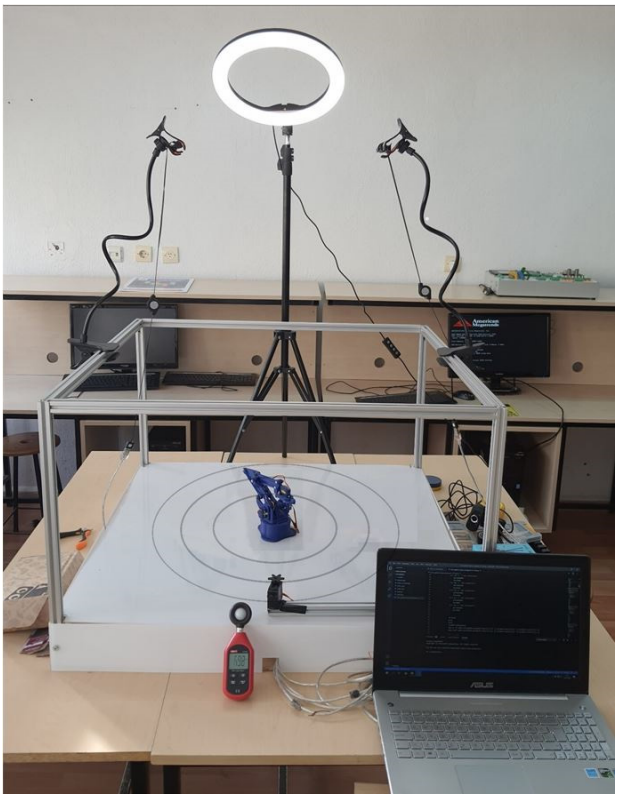


**Figure 14.** The experimental setup with two cameras and one robot.

Real-time operating conditions and the response of the system should be examined in a controlled manner. Therefore, a test system under controlled conditions was designed and installed. First, the parameters to be controlled were selected. One of the critical variables in the system is the reaction time of the robot arm for decelerating or stopping

operations if a foreign object enters the controlled area. However, the reaction time is directly related to the object entering the control area. Therefore, the speed of the entering object must be under control. For this system, a controllable speed foreign object design was made.

In this validation demo, a servo motor is preferred for controllable speed. During the driving of a servo motor performing the rotary motion, the software can determine how many degrees of rotation per second can be made. Using the motor driver software's temporal change, a controlled movement can be achieved for the object at the desired speed. Thus, it is necessary to place an object at a certain distance from the motor drive shaft center. When the robot arm is placed on the border of the control area for this object's movement, the foreign object is allowed to enter the control area at the desired speed. For this purpose, a servo motor driven foreign object system was established, as shown in Figure 15. The controlled objects entering the warning and alarm zones and their movement trajectories are shown in the left and right images of Figure 15, respectively.

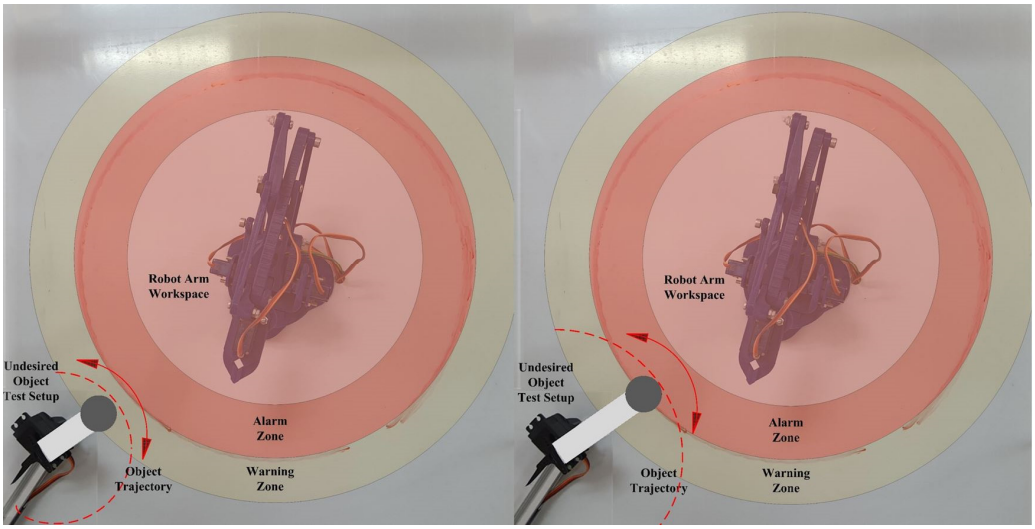



**Figure 15.** Speed controlled foreign object and controlled area in the experimental setup.

### *4.2. Test Results and Discussion*

The entire object in the images did not fall within the monitored areas. The desired linear velocity was reached at the exact endpoint of the object, and this end part was allowed to enter the control area. The response time of a mechanism whose speed can be controlled when the robot arm enters the camera-controlled working space was measured. The system responded at various times in response to a foreign object entering the area at various speeds. This controlled speed and the system response time are shown in the graphs given in Figure 16. On the curves in these graphs, there are two marked points, which are indicated for "Walking Speed" (2 m/s) and "Hand Speed" (1.6 m/s), as specified in the standards.

The plot above and below are derived from the safety systems using low and high-resolution cameras, respectively. In both plots, the performance of two different graphics cards, including NVIDIA Geforce GTX 850 M and a GTX 1050, were compared. GTX 1050 is reported to have approximately 60% higher performance than GTX 850 M according to benchmark tests. All experiments were performed at a constant illumination level of 300 lux.

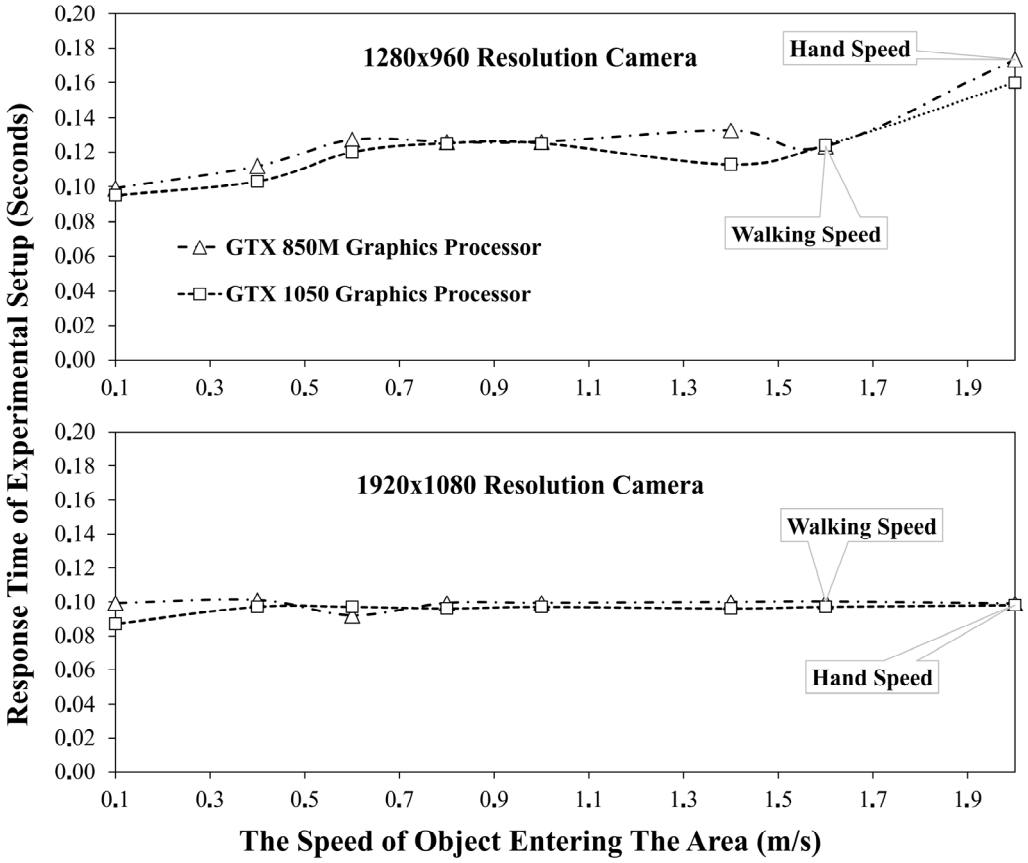



**Figure 16.** System response time versus foreign object speed as a function of camera resolution using two graphics processors.

It is seen that the effect of camera resolutions on the system response was high. When using a low-resolution camera, the system response was more unstable at different speed inputs, while it seemed to work more stable with a high-resolution camera. On the other hand, the graphics processor did not have a significant effect (less than 0.01 s) on the system response with the high-resolution cameras. In the system in which low-resolution cameras were used, it was observed that the high-performance graphics processor responded in an average of 10% shorter time than the lower performance one. Therefore, the proposed system yielded efficient results even in low-performance processors when 1080p resolution cameras were used.

Consequently, a camera with a resolution of at least 1080p should be used for the safety system. When a safety system with optimal values was installed, the system response time was between 0.08 and 0.1 s. These values were considered sufficient for an effective operation. When graphics processors are compared, it should not be ignored that the system memory (RAM) and the hard drive are also influential. In the comparisons in this study, both computers had 16 GB RAM and 256 GB SSD drive.

Another essential criterion in this study is the response times of the proposed system at various illuminance levels in the controlled area. The system was tested under various illuminance levels. illuminance is the total luminous flux incident on a surface per unit area and measured in lux. One lux is the illuminance created by a light source with an intensity of one candlelight in the center of the sphere with a radius of 1 m [32,33]. In the experimental system established with a sensor measuring illuminance, response time measurements were made at different illuminance, and these values are presented in Figure 17.

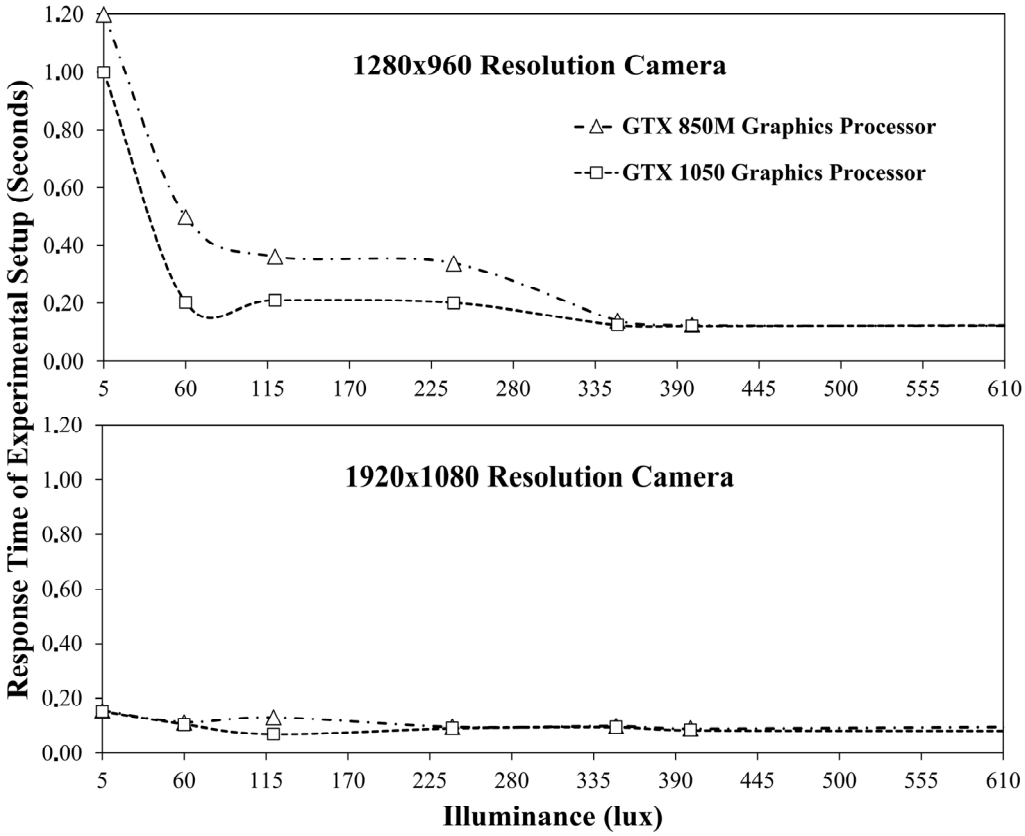



**Figure 17.** System response time versus illuminance as a function of camera resolution using two graphics processors.

The plots above and below in Figure 17 are derived from the safety systems using low and high-resolution cameras, respectively. In both plots, the performance of two different graphics cards, including NVIDIA Geforce GTX 850 M and a GTX 1050, were compared. GTX 1050 is reported to have approximately 60% higher performance than GTX 850 M according to benchmark tests. The graphics processor did not appear to impact system response times at high illumination levels significantly. At the same time, the same experiments were carried out with cameras with two different resolutions in the graphics shown in the figure. Camera resolution appeared to have a high impact on system response at different illuminance levels. As seen from this graph, the minimum brightness of the camera to detect a foreign object was 5 lux. At this level, the detection time was at the highest level. The low-resolution camera showed that the detection time became shorter, and the performance increased as the illumination rose to 120 lux. The fluctuation seen after this level was related to the stability of the system itself. In terms of illuminance, the lowest level available was 120 lux. The high-resolution camera observed that the system worked stably even at low illumination levels.

Commercial production areas are known to require at least 300–400 lux light [34]. To understand the illuminance levels more clearly, it is measured as 100 lux on a dark or cloudy day [35]. The illuminance of a residential living room at evening time is measured as 50 lux [36].

Considering the values obtained, the experimental setup was sufficient. The desired system can be installed by considering the camera's high aperture, which is expected to detect in systems where it is desired to operate at a lower light intensity. As a result, the hypothesis outlined in the experimental setup carried out was validated. Improvement suggestions are presented for large-scale industrial applications. In addition, cost comparisons for the commercial product and the designed safety system are shown in Table 4.

**Table 4.** Cost comparison of the proposed and commercial systems (Approximate prices, 2021).

| | Commercial System (USD) | Proposed System with 2 Cameras (USD) | Proposed System with 3 Cameras (USD) |
|---|---|---|---|
| 3D Camera | 9500 | - | - |
| 2D Cameras | - | 120 | 180 |
| Analysis Unit | 14,000 | 360 | 540 |
| Control System | 12,000 | 1700 | 2550 |
| Total | 35,500 | 2180 | 3270 |

Another comparison between the experimental setup and the commercial product is how the system set up in different images will react. As a result of these comparisons, the effects of the following features were sought:

- Object size;
- Object brightness;
- Object geometry;
- Defective object corners;
- Camera resolution;
- The closeness of object and background colors to each other;
- Shadows on the object;
- Stray light reflections from the object.

To compare the results of these factors, the aim was to learn the effect of the size, color, geometry, brightness, and shadings while classifying. As a result of the comparisons, it was seen that the response times of the proposed safety system were close to each other. Thus, by this experimental setup, it was proven that a realistic and cost-effective alternative option to commercial ones was assured.

To validate the proposed safety system in the work cell with an industrial quality robot arm, the safety system was rescaled for a work cell with an Acrome-Acrobot experimental robot, and two 2-D cameras were used (Figure 18). The yellow-black safety strip indicates the safety zone. When this system was tested, very similar response time results were obtained. Hence, it was proven that the proposed design is also applicable on a larger scale.

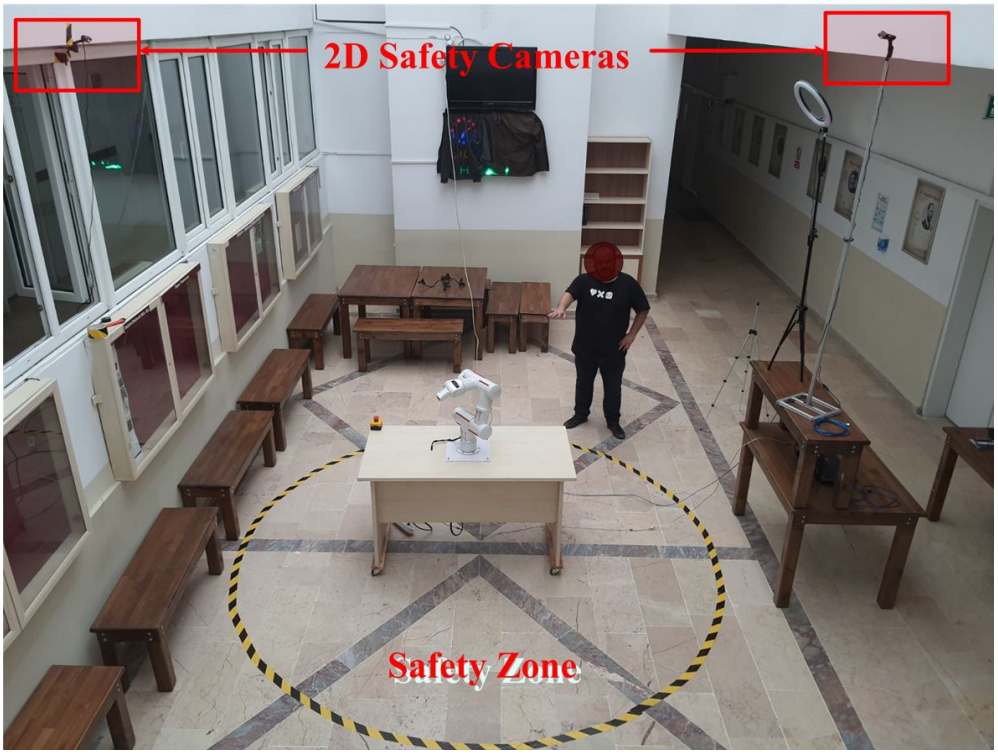



**Figure 18.** Implementing the optimally designed safety system at a larger scale robot work cell.

## 5. Concluding Remarks

This paper proposed an optimized safety system with 2-D cameras for fenceless robotic systems. Camera lens angle, camera orientation angle, and camera height were considered as variables of the optimization problem. To eliminate blind spots, a simple but novel 2-D image merging technique was proposed as an alternative to commonly employed stereo imaging methods. The designs at optimal and nonoptimal parameters were compared using a deep learning algorithm ResNet-152.

Validation experiments were conducted by using two camera resolutions as well as two graphic processors. In these tests, the graphic processor performance did not significantly affect the system response time, while the camera resolution was very effective. Hence, the most suitable 2-D camera resolution for a safety system should be at least 1080p at 23 FPS camera frame frequency. When the cameras were placed optimally, the system was proven to work effectively at a minimum illuminance of 120 lux, while commercial systems cannot be operated under 400 lux. Finally, the two-camera system was tested at a larger scale and a similar performance was observed. The experimental safety system ran on a PC with MS Windows, but it can be adapted to a commercial robot controller via two digital input channels. Consequently, a relatively simple 2-D safety camera system, which is more effective in terms of blind spots and with a 15 times better price–performance ratio, as compared with commercial alternatives, was proposed. In the next phase of research, improvement of the foreign object detection algorithm using deep and machine learning will be studied, and validation in a manufacturing factory is planned.

**Author Contributions:** Data curation, M.O.; formal analysis, M.O.; investigation, M.O.; methodology, C.Y. and H.L.; project administration, C.Y. and H.L.; software, M.O.; supervision, C.Y. and H.L.; validation, M.O.; writing—original draft preparation, M.O.; writing—review and editing, C.Y. and H.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The work presented in this paper is a part of M. Ozkahraman's doctoral thesis. Gratitude goes to Z. Yagiz Bayraktaroglu of Istanbul Technical University for his support and guidance.

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

## References

1. Takata, S.; Hirano, T. Human and robot allocation method for hybrid assembly systems. *CIRP Ann.* **2011**, *60*, 9–12. [CrossRef]
2. Koren, Y.; Heisel, U.; Jovane, F.; Moriwaki, T.; Pritschow, G.; Ulsoy, G.; Van Brussel, H. Reconfigurable manufacturing systems. *CIRP Ann.* **1999**, *48*, 527–540. [CrossRef]
3. Krüger, J.; Lien, T.; Verl, A. Cooperation of human and machines in assembly lines. *CIRP Ann.* **2009**, *58*, 628–646. [CrossRef]
4. Krüger, J.; Nickolay, B.; Heyer, P.; Seliger, G. Image based 3D Surveillance for flexible Man-Robot-Cooperation. *CIRP Ann.* **2005**, *54*, 19–22. [CrossRef]
5. Weitschat, R.; Ehrensperger, J.; Maier, M.; Aschemann, H. Safe and efficient human-robot collaboration part I: Estimation of human arm motions. In Proceedings of the IEEE International Conference on Robotics and Automation, Brisbane, Australia, 21–25 May 2018.
6. Gecks, T.; Henrich, D. Human-robot cooperation: Safe pick-and-place operations. In Proceedings of the Robot and Human Interactive Communication, IEEE International Workshop, Nashville, TN, USA, 13–15 August 2005.
7. Bley, H.; Reinhart, G.; Seliger, G.; Bernardi, M.; Korne, T. Appropriate Human Involvement in Assembly and Disassembly. *CIRP Ann.* **2004**, *53*, 487–509. [CrossRef]
8. *ISO 10218-1: 2011: Robots and Robotic Devices-Safety Requirements for Industrial Robots-Part 1: Robots*; ISO: Geneva, Switzerland, 2011.
9. Oberer, S.; Schraft, R.D. Robot-dummy crash tests for robot safety assessment. In Proceedings of the Robotics and Automation, IEEE International Conference, Rome, Italy, 10–14 April 2007.
10. Saenz, J.; Behrens, R.; Schulenburg, E.; Petersen, H.; Gibaru, O.; Neto, P.; Elkmann, N. Methods for considering safety in design of robotics applications featuring human-robot collaboration. *Int. J. Adv. Manuf. Technol.* **2020**, *107*, 2313–2331. [CrossRef]

11. Tashtoush, T.; Garcia, L.; Landa, G.; Amor, F.; Nicolas, A.; Oliva, D.; Safar, F. Human-Robot Interaction and Collaboration (HRI-C) Utilizing Top-View RGB-D Camera System. *Int. J. Adv. Comput. Sci. Appl.* **2021**, *12*, 11–17. [CrossRef]
12. Kaonain, T.E.; Rahman, M.A.A.; Ariff, M.H.M.; Yahya, W.J.; Mondal, K. Collaborative Robot Safety for Human-Robot Interaction in Domestic Simulated Environments. *IOP Conf. Series: Mater. Sci. Eng.* **2021**, *1096*, 012029. [CrossRef]
13. *ISO 13855: 2010 Safety of Machinery–Positioning of Safeguards with Respect to the Approach Speeds of Parts of the Human Body*; ISO: Geneva, Switzerland, 2010.
14. Bessler, J.; Prange-Lasonder, G.B.; Schaake, L.; Saenz, J.F.; Bidard, C.; Fassi, I.; Valori, M.; Lassen, A.B.; Buurke, J.H. Safety Assessment of Rehabilitation Robots: A Review Identifying Safety Skills and Current Knowledge Gaps. *Front. Robot. AI* **2021**, *8*, 33. [CrossRef] [PubMed]
15. Hanh, L.D.; Lin, C.Y. Autonomous gluing based on image-based visual servoing. *Int. J. Robot. Autom.* **2021**, *36*. [CrossRef]
16. Yin, C.; Li, B.; Shi, W.; Sun, N. Monitoring-based visual servoing of wheeled mobile robots. *Int. J. Robot. Autom.* **2019**, *34*, 34. [CrossRef]
17. Tan, J.T.C.; Duan, F.; Zhang, Y.; Kato, R.; Arai, T. Safety design and development of human-robot collaboration in cellular manufacturing. In Proceedings of the IEEE International Conference on Automation Science and Engineering, St. Louis, MO, USA, 10–15 October 2009.
18. Himmelsbach, U.B.; Wendt, T.M.; Lai, M. Towards safe speed and separation monitoring in human-robot collaboration with 3-D-time-of-flight cameras. In Proceedings of the Second IEEE International Conference on Robotic Computing, Laguna Hills, CA, USA, 31 January–2 February 2018.
19. Michalos, G.; Makris, S.; Tsarouchi, P.; Guasch, T.; Kontovrakis, D.; Chryssolouris, G. Design considerations for safe human-robot collaborative workplaces. *Procedia CIRP* **2015**, *37*, 248–253. [CrossRef]
20. Schmidt, B.; Wang, L. Depth camera-based collision avoidance via active robot control. *J. Manuf. Syst.* **2014**, *33*, 711–718. [CrossRef]
21. Steger, C.; Ulrich, M.; Wiedemann, C. *Machine Vision Algorithms and Applications*; John Wiley & Sons: Hoboken, NJ, USA, 2018.
22. *ISO 13849-1: 2008. Safety of Machinery–Safety-Related Parts of Control Systems-Part 1*; ISO: Geneva, Switzerland, 2009.
23. Durgun, I.; Tokçalar, Ö.; Ilhan, R.; Şimşek, Ö. Human-robot interaction in hybrid assembly systems using medium payload robots. In Proceedings of the National Robotic Conference—ToRK 2016, İstanbul, Turkey, 2–3 November 2016.
24. *ISO 13849-2: Safety of Machinery-Safety-Related Parts of Control Systems-Part 2: Validation*; ISO: Geneva, Switzerland, 2003.
25. Metev, B.; Gueorguieva, D. A simple method for obtaining weakly efficient points in multiobjective linear fractional programming problems. *Eur. J. Oper. Res.* **2000**, *126*, 386–390. [CrossRef]
26. Rajpoot, M.H.C. "HCR's infinite-series" analysis of oblique frustum of a right circular cone. *Int. J. Math. Phys. Sci. Res.* **2014**, *2*, 1–17.
27. *CIPA DCG-001-2005. Guideline for Noting Digital Camera Specifications in Catalogues*; CIPA: Minato-ku, Tokyo, Japan, 2005.
28. *CIPA DC-003-2003. Resolution Measurement Methods for Digital Cameras*; CIPA: Minato-ku, Tokyo, Japan, 2003.
29. Hornberg, A. *Handbook of Machine Vision*; John Wiley & Sons: Hoboken, NJ, USA, 2006.
30. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016.
31. Pilz Safe Camera System Safety Eye. Available online: https://en.sslog.de/equipment/safety-sensors-psen/safe-camera-system-safetyeye (accessed on 3 December 2021).
32. Yusuf, M.; Azmi, M.S.B.M.; Ismail, A.; Ibrahim, I.I.; Hashim, M.; Kamarrudin, N. Webcam Based Lux Meter Using Grid-Based Image Processing. In Proceedings of the International Conference on Computational Approach in Smart Systems Design and Applications, Kuching, Malaysia, 15–17 August 2018.
33. Ismail, A.H.; Azmi, M.S.M.; Hashim, M.A.; Ayob, M.N.; Hashim, M.S.M.; Hassrizal, H.B. Development of a webcam-based lux meter. In Proceedings of the IEEE Symposium on Computers & Informatics, Langkawi, Malaysia, 7–9 April 2013.
34. He, Y.; Zhang, X.; Zhu, L.; Sun, G.; Qiao, J. Design and implementation of the visual detection system for amphibious robots. *Int. J. Robot. Autom.* **2019**, *34*. [CrossRef]
35. Schlyter, P. Radiometry and Photometry in Astronomy. Available online: Stjarnhimlen.se/comp/radfaq.html (accessed on 3 December 2021).
36. Pears, A. *Strategic Study of Household Energy and Greenhouse Issues*; Australian Environmental Report; Sustainable Solutions Australia: Brighton, Australia, 1998.