applied sciences

MDPI

*Article*

# Vision- and Lidar-Based Autonomous Docking and Recharging of a Mobile Robot for Machine Tending in Autonomous Manufacturing Environments

**Feiyu Jia [1], Misha Afaq [1], Ben Ripka [2], Quamrul Huda [2] and Rafiq Ahmad [1,*]**

[1] Smart & Sustainable Manufacturing Systems Laboratory (SMART LAB), Department of Mechanical Engineering, University of Alberta, 9211 116 Street NW, Edmonton, AB T6G 1H9, Canada; feiyu1@ualberta.ca (F.J.); mafaq@ualberta.ca (M.A.)

[2] Centre for Sensors and System Integration (CSSI), Northern Alberta Institute of Technology (NAIT), 11762 106 Street, Edmonton, AB T5G OY2, Canada; bripka@nait.ca (B.R.); qhuda@nait.ca (Q.H.)

* Correspondence: rafiq.ahmad@ualberta.ca

**Abstract:** Autonomous docking and recharging are among the critical tasks for autonomous mobile robots that work continuously in manufacturing environments. This requires robots to demonstrate the following abilities: (i) detecting the charging station, typically in an unstructured environment and (ii) autonomously docking to the charging station. However, the existing research, such as that on infrared range (IR) sensor-based, vision-based, and laser-based methods, identifies many difficulties and challenges, including lighting conditions, severe weather, and the need for time-consuming computation. With the development of deep learning techniques, real-time object detection methods have been widely applied in the manufacturing field for the recognition and localization of target objects. Nevertheless, those methods require a large amount of proper and high-quality data to achieve a good performance. In this study, a Hikvision camera was used to collect data from a charging station in a manufacturing environment; then, a dataset for the wireless charger was built. In addition, the authors of this paper propose an autonomous docking and recharging method based on the deep learning model and the Lidar sensor for a mobile robot operating in a manufacturing environment. In the proposed method, a YOLOv7-based object detection method was developed, trained, and evaluated to enable the robot to quickly and accurately recognize the charging station. Mobile robots can achieve autonomous docking to the charging station using the proposed Lidar-based approach. Compared to other methods, the proposed method has the potential to improve recognition accuracy and efficiency and reduce the computation costs for the mobile robot system in various manufacturing environments. The developed method was tested in real-world scenarios and achieved an average accuracy of 95% in recognizing the target charging station. This vision-based charger detection method, if fused with the proposed Lidar-based docking method, can improve the overall accuracy of the docking alignment process.

**Keywords:** mobile robots; autonomous recharging; autonomous docking; manufacturing environments; 3D Lidar; computer vision



**Citation:** Jia, F.; Afaq, M.; Ripka, B.; Huda, Q.; Ahmad, R. Vision- and Lidar-Based Autonomous Docking and Recharging of a Mobile Robot for Machine Tending in Autonomous Manufacturing Environments. *Appl. Sci.* **2023**, *13*, 10675. https://doi.org/10.3390/app131910675

Academic Editor: Yutaka Ishibashi

Received: 3 August 2023
Revised: 7 September 2023
Accepted: 16 September 2023
Published: 26 September 2023



## 1. Introduction

The autonomous recharging process is an important part of a mobile robot's autonomous operations, enabling it to work continuously without any human intervention. Docking [1] can be understood as the navigation and localization of a robot toward a desired location. Docking requires an accurate estimation of the robot's pose, often from a position close to the docking station, through path planning [2]. Mobile robots are used across various fields [3–8], including surveillance, planetary exploration, dangerous environments, factory automation, search and rescue operations, and indoor manufacturing environments. The role of mobile robots has become increasingly important for present and

future applications. Thus, independent autonomous recharging has become a fundamental requirement through which to ensure the autonomous operation of mobile robots in various conditions. For a mobile robot to initiate the docking and recharging process, it first needs to identify the charging station and then align itself with the charger autonomously by following a series of rotational and translational steps.

The location of the charging station (e.g., indoor or outdoor) plays an important role in the selection of sensors for a docking procedure. Outdoor environments are more complex, unpredictable, and dynamic due to the presence of moving objects and obstacles. Moreover, the performance of non-visual sensors, such as Lidar sensors, which are used for docking, can depreciate in outdoor weather conditions, such as snow, dust, and fog [9]. Based on the sensor implemented, the autonomous docking techniques described in the literature are divided into the three following categories: (i) infrared (IR) sensor-based methods [10], (ii) computer vision [11], and (iii) laser-based approaches [12]. To receive IR signals properly, the IR receiver needs to be implemented in a specific location on the mobile platform, which limits the mechanical design of mobile robots [10]. Computer vision and laser-based techniques, such as object detection [13] and Lidar-based approaches, are the most commonly used methods through which to solve odometry-related problems. However, both techniques have their respective limitations and benefits. Although Lidar methods [14] can extract different features from the environment, without being affected by changes in lighting conditions, and can obtain more accurate range measurements than cameras, Lidar data are sparsely distributed and have limited visibility. Furthermore, this technique's operation is based on collecting large amounts of data, which requires more computational power than a camera. In contrast, a camera provides rich and thick data that do not have limited visibility. However, a standard camera without a 360-degree view has a limited visibility angle, resulting in a blind spot [11].

To overcome the challenges of conventional methods, the combination of different sensors has been investigated for years, and recent research has proven that the fusion approach yields better performance than a single-sensor method [14]; the limitations of IR-based, laser-based, and vision-based methods for autonomous docking and recharging are overcome by combining multiple sensors. The authors of [15] attempted to integrate a camera and IR sensor with laser range finders in order to improve the reliability of the autonomous docking process. In [16], a vision-based autonomous docking and recharging approach was applied to a security robot. An artificial landmark was installed on top of a charging station, at the same height as the camera, to assist the robot in detecting and locating the charging station area. The rotational and translational errors were counteracted using a virtual spring model motion control approach. The model presented in [16] assumed that the robot and the charger could be connected with a virtual spring, and the compliant forces in the direction of the translation deformation and bending determined the motion control. However, the vision-based docking approach is prone to calibration errors, as demonstrated in [17], where a Faster R-CNN algorithm was used to detect arbitrary visual markers. The pose of the mobile robot was estimated using the solvePnP algorithm, which related the 2D–3D point pairs. However, the solvePnP algorithm gave systematically inaccurate pose estimates in the x-direction and, hence, proved to be ineffective for docking. Laser range finder techniques usually detect the charger based on the uniquely manufactured shape of the charging station to distinguish it from surrounding objects. One such example is the V-shaped recess on the MiR (mobile industrial robot) [18] made by Fetch Robotics, which required the charger to be placed separately from any laser-height obstacles to enable the successful detection of the contour of the charger using the laser range finder. However, the requirement of a special shape adds to a charger station's fabrication costs and limits mobile robots' practical applications in unstructured environments. To solve this problem, a self-adhesive reflective tape can be used to help the robot identify the charger, as reported in [19]. When using this reflection detection technique, the charger was easily distinguished from other similar objects in an unstructured environment, as verified by extensive experiments. Moreover, Lidar can be used for obstacle detection and avoidance,

navigation, and pose estimation in a mobile robot without the use of additional hardware. In [20], a multi-sensor fusing method used intensity and range data fusion, with a covariance intersection approach, to estimate the robot pose during docking and recharging. Using the inverse perspective projection method, an artificial landmark was employed as a visual cue on the charging station to be identified by the robot. Then, based on the laser range data, the geometrical relationship between the robot and charger station was estimated precisely using the covariance intersection method. In [21], automated guided vehicle (AGV) autonomous docking was investigated in an unstructured environment with human presence. An autonomous docking technique was implemented with a non-visual sensor, such as Lidar or AprilTag, for charger detection. A deep learning network was used to detect and recognize humans and objects. Practical experiments verified that the AGV could co-exist with humans and perform autonomous docking in unstructured environments. With the development of deep learning techniques, deep-learning-based approaches perform better in autonomous docking applications. In [21], the MobileNetv2-SSDLite deep learning framework was adopted to detect and recognize a specific person in the human–robot collaborative environment. Once the particular human was identified, the robot system could achieve automatic docking to the target person based on LiDAR and a RGB-D camera. Given that high-resolution images from a camera can provide rich information, in [22], the authors proposed a fusion method to make use of images from a camera to enrich the raw 3D point clouds from LiDAR. The sparse convolutional neural network was adopted to predict the dense point clouds to enrich the raw point clouds and then employed to execute LiDAR SLAM. In [23], the Faster-RCNN model with a MobileNetv3-Large FPN backbone was used to detect arbitrary dynamic obstacles and identify the charging station. It was proven that it can distinguish the charging station from other surrounding objects in most scenarios.

Previous studies indicate that the autonomous docking and recharging process becomes more reliable and repeatable when using a multi-sensor fusion approach in both structured and unstructured environments. However, IR sensors require specific configurations, such as signal receivers, which are inconvenient and incur high costs [9]. Therefore, most existing fusion methods consider combining the Lidar sensor with computer vision techniques because of their low costs and non-destructive abilities. However, computer vision techniques, especially deep-learning-based object detection, require a large amount of proper task-oriented high-quality data for training and tuning to achieve the desired performance [24]. The changing lighting conditions and jerking of the camera on mobile robots can also affect the performance of deep-learning-based object detection models [25], which makes it difficult to implement solely computer-vision-based techniques in real-world manufacturing applications.

Considering the aforementioned challenges, this paper has the following aims:

- This paper aims to develop a vision–Lidar data fusion method for mobile robots to achieve accurate autonomous docking and recharging in a manufacturing environment.
- This paper contributes to the transition of state-of-the-art real-time object detection methods from general public datasets to real-world manufacturing tasks by combining deep-learning-based techniques to identify charging stations in a complex manufacturing environment; we then use a Lidar-based approach to localize the detected wireless charger and dock the mobile robot to it for recharging.
- An indoor manufacturing environment with an enclosed space where a wireless charging station is situated is considered for the implementation of the docking procedure. The proposed method is analyzed and discussed based on the autonomous docking and recharging of a Husky robot made by Clearpath Robotics.
- A YOLOv7-based method is used to detect the charging station for the robot to navigate to the desired location. The process of planning a path to the charger can be achieved with waypoints using the SLAM method, which is not discussed in this paper. Afterward, the Lidar sensor is used, along with the detected results from the camera,

to determine the distance from the charger and side wall to achieve an accurate pose estimation and then successfully dock the robot to the charging station. The proposed method can be easily adapted to different types and numbers of wireless chargers in a manufacturing environment. The distance data between the Lidar and the camera can be calibrated to achieve accurate alignment and pose estimation.

This paper is structured as follows: The related work is presented in Section 2; Section 3 explains the proposed method in detail; Section 4 contains the results; and Section 5 comprises the discussion and conclusions of this paper.

## 2. Related Work

In this section, we present the recent docking and recharging methods, based on Lidar and computer vision techniques, for mobile robot systems in the manufacturing field. Fan et al. [5] proposed a vision-based docking and recharging method that can be applied in a warehouse environment. This method used AprilTag for the detection and identification of the robot's pose. It achieved a docking success rate of approximately 97.33%. In [17], the authors proposed a Faster RCNN model to detect and localize the designed markers mounted on a docking station, combining it with the solvePnP algorithm to allow the mobile robot to navigate in a ROS simulation environment. This model achieved an accuracy of 96.3% based on thirteen testing images. The detector took around 35 ms to process each image. Song et al. [21] adopted a single-shot detector (SSD) to identify moving people and then dock to the target person for human–robot collaborative tasks in an unstructured environment. In [23], an SSD was developed to detect the charging stations in obstacle-free scenarios. This method could achieve a performance of 99.8% for successful docking to the charger. It took an average of 12 s to complete the docking procedures based on the designed scenarios.

Although these methods have made great contributions to autonomous docking and recharging applications, some limitations are observed. Most methods are evaluated in a simulation or laboratory environment instead of a manufacturing environment. In addition, two-stage deep learning models, such as Faster RCNN, are inefficient compared to one-stage real-time models. Considering these limitations, a state-of-the-art real-time deep-learning-based model, YOLOv7, is developed to distinguish and identify the target wireless charger from a complex manufacturing environment; it is integrated with the proposed Lidar-based approach to achieve efficient, low-cost, and robust docking and recharging.

## 3. System Overview

The autonomous mobile robot is shown in Figure 1. A Husky UGV field search robot made by Clearpath Robotics is used to implement the Lidar-vision-based docking method and conduct autonomous charging experiments in indoor manufacturing environments. Figure 1 shows the Husky robot installed with a Lidar sensor and a Hikvision camera. The ROS Melodic software development platform is used to program the docking process using the 3D Lidar sensor and to control the robot's motion through the docking steps.

The wireless charging station used in this study is presented in Figure 2; it is installed inside a custom-sized modular structure. A ramp door placed in the front allows the robot to come out of the docking station to run missions and return for recharging as necessary.

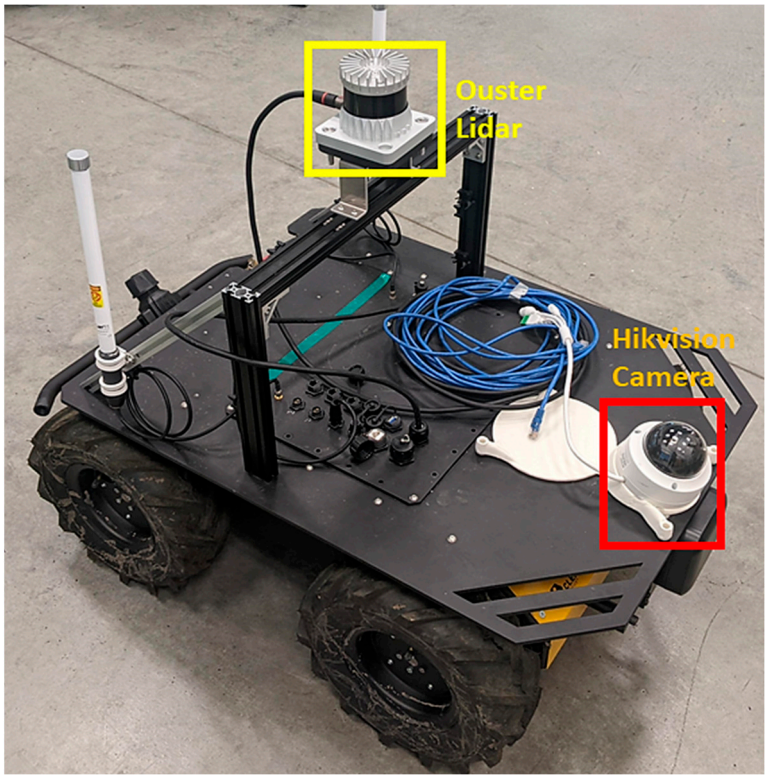


**Figure 1.** Husky robot setup with a Lidar sensor and Hikvision camera.

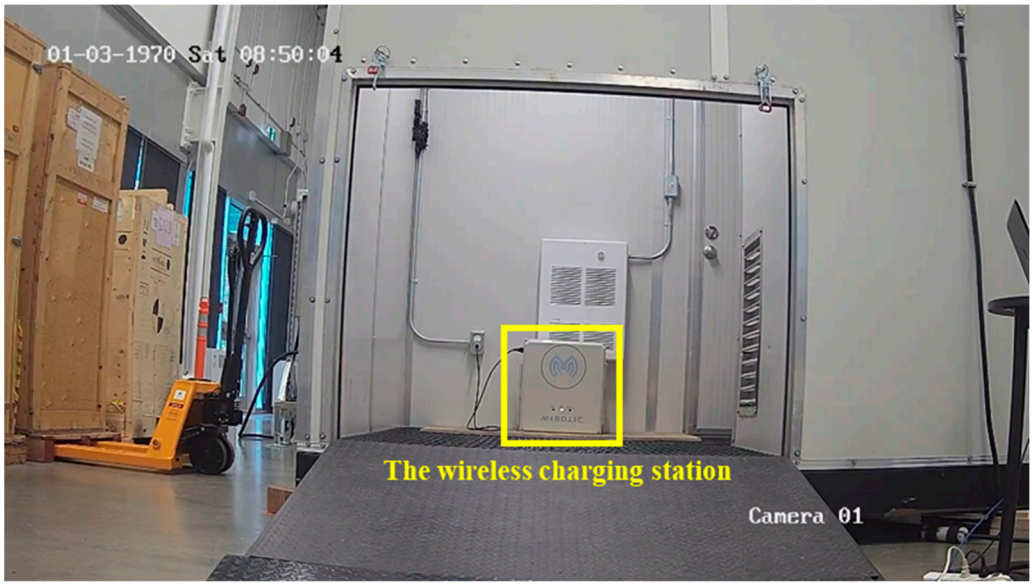


**Figure 2.** The charging station used in this study.

## 4. Proposed Method

This section proposes a vision- and Lidar-based autonomous docking and recharging approach. The proposed method consists of three main steps: (i) data collection, which is achieved by using a Hikvision camera and Ouster Lidar to capture the surrounding environments through RGB images and laser-based distance/depth information, respectively; (ii) a deep-learning-based object detection method, with the YOLOv7 model as the core architecture, which is used to recognize the charging station in the manufacturing environment; and (iii) a Lidar-based approach to adjust the pose of the mobile robot and

then dock it to the detected wireless charger. A flowchart of the proposed method is presented in Figure 3.

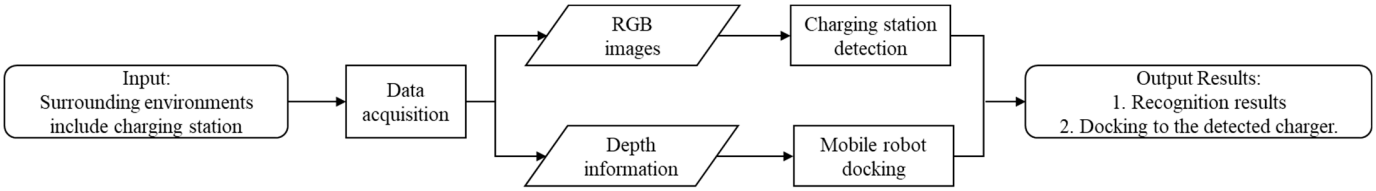


**Figure 3.** A block diagram of the proposed docking and recharging method.

### *4.1. YOLOv7 Architecture*

YOLOv7 is a one-stage model and the latest algorithm for real-time object detection, and it performs well in terms of both speed and accuracy [26]. The architecture of the proposed charging station detection method based on YOLOv7 is presented in Figure 4, and it is composed of three main components: a backbone, neck, and head. The convolutional backbone module adopts Darknet-53 [27] to extract image feature maps from the input image and transfer them to the neck layers. In the neck module, the Feature Pyramid Network (FPN) [28] is used to enhance the feature maps. These maps are then combined, fused, and passed to the subsequent layers. Finally, the head network predicts the bounding boxes and classes of the objects.

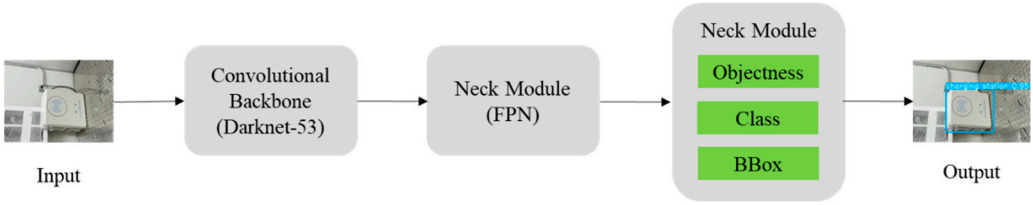


**Figure 4.** A flowchart of the charger detection method.

YOLOv7 adopts a developed extended efficient layer aggregation network to improve inference efficiency. This network can quicken the model's learning ability without disturbing or changing the original gradient propagation path. In addition, a novel scaling method, referred to as corresponding compound model scaling, is proposed to address the issue of a larger width output of the computational block by directly scaling the depth of the concatenation-based model. Moreover, several techniques are used to improve inference accuracy while keeping training costs low. These techniques, called Bags of Freebies (BoF), include planned re-parameterization, dynamic label assignment, and batch normalization. After thoroughly investigating the re-parametrized convolution, the authors demonstrate increased model accuracy when using RepConv without an identity connection. Furthermore, batch normalization integrates the mean and variance in the data to adjust the bias and weight of the convolutional layer, which can immediately impact the training process by utilizing a higher training rate and faster convergence.

According to [26], YOLOv7 optimizes the inference process and improves detection accuracy and speed compared with other existing real-time object detection methods because of its more advanced network structure and training strategies. However, it has not yet been used in the domain of autonomous docking and recharging. In this article, YOLOv7 is adopted as the backbone architecture to detect and recognize the charging station.

### *4.2. Lidar and Vision Data Fusion Method for Autonomous Docking*

In recent research, Lidar sensors and cameras have commonly been used together in autonomous driving applications, because a Lidar sensor can collect 3D spatial information. In contrast, a low-cost camera captures the appearance and texture of the corresponding area in 2D images. Therefore, the fusion of Lidar and the camera data can improve the object detection performance. Lidar–camera calibration estimates a transformation matrix that gives the relative rotation and translation between the 2D coordinates obtained from the Hikvision camera and the 3D spatial coordinates obtained from the Lidar, as demonstrated in Equation (1) [29]. The 3D coordinates of the charging station can be calculated using Equations (2)–(4) [29] based on the predicted bounding box in the image domain:

$$z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{f_x}{dx} & 0 & u_0 \\ 0 & \frac{f_y}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ z_c \\ 1 \end{bmatrix} \tag{1}$$

$$X = \frac{u - u_0 \cdot z_c \cdot dx}{f_x} \tag{2}$$

$$Y = \frac{v - v_0 \cdot z_c \cdot dy}{f_y} \tag{3}$$

$$Z = z_c \tag{4}$$

where $u$ and $v$ are the 2D coordinates from the camera; $u_0$ and $v_0$ are the origins of the coordinate system in the image domain; $f_x$ and $f_y$ are the focal lengths along the $x$ and $y$ directions, respectively; $X$, $Y$, and $Z$ are the 3D global coordinates from the Lidar; and $z_c$ is the distance between the detected object and the camera. An illustration of the transformation process is presented in Figure 5.

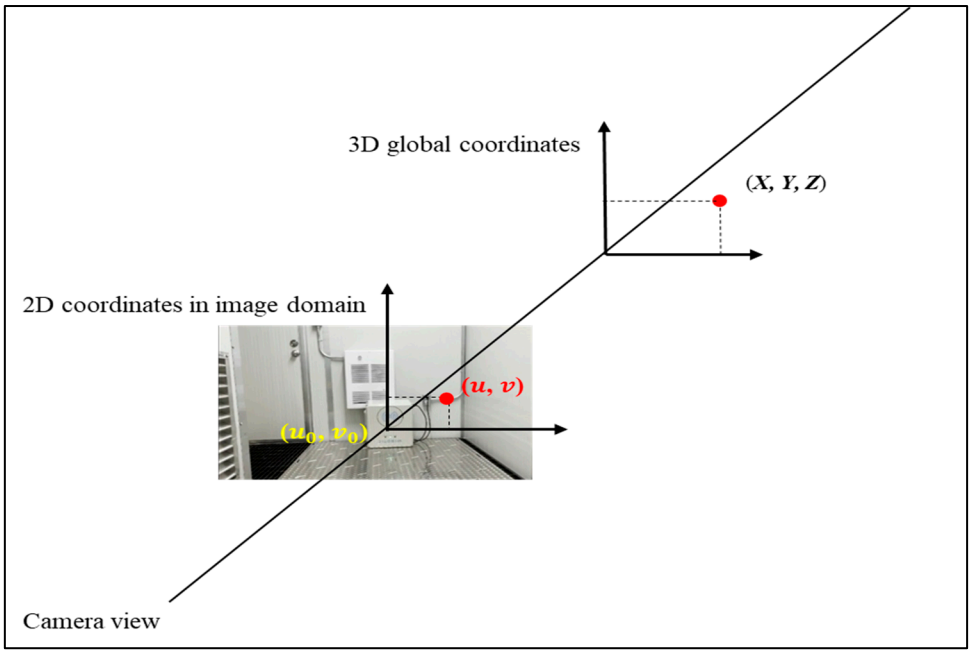


**Figure 5.** Illustration of the transformation process.

An Ouster Lidar sensor is utilized to calculate the distances from the robot frame of reference to the side wall and the depth or the distance to the charger. It is assumed that the charging station is enclosed within walls to simplify the pose estimation of the robot for the docking process. Two scenarios are considered for the implementation of the Lidar–vision docking method: docking in an environment with only one charger and in one with three chargers, as shown by the Gazebo virtual environment setups in Figure 6.

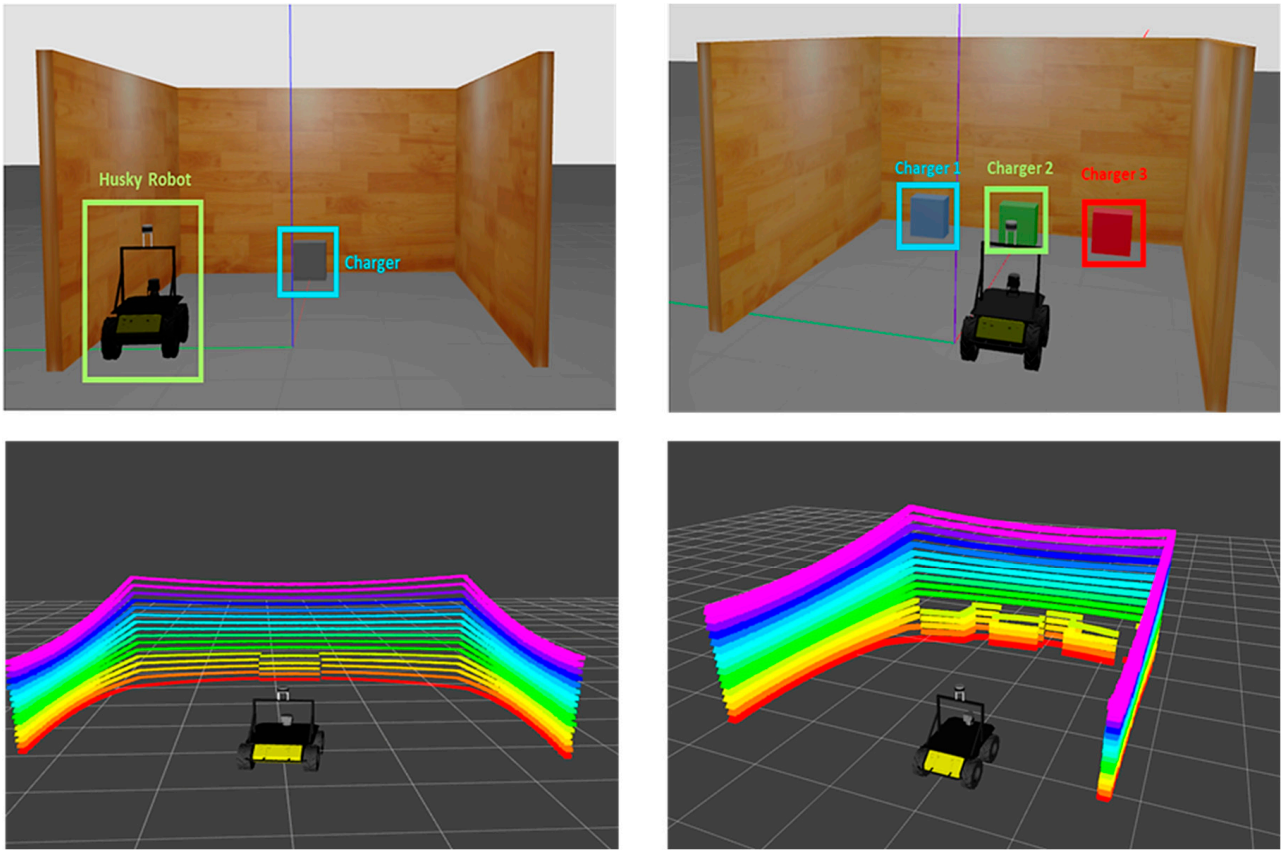


**Figure 6.** Docking station Gazebo virtual environment setup with one charger (**top left**) and three chargers (**top right**), and Rviz Lidar point cloud visualization for one charger (**bottom left**) and three chargers (**bottom right**).

In the case of the scenario with three different chargers, the vision-based method will aid the robot in identifying the correct charger and autonomously docking with it. Rviz software is used to visualize the Lidar point cloud data of the charging stations for both the one-charger and three-charger setups, as demonstrated in Figure 6. The pose estimation and navigation for docking are used with the Lidar sensor data based on the information given in Figure 7. After the correct charging station is identified using computer vision algorithms, the Lidar point cloud data are filtered to obtain two diagonal and two straight lines, called, respectively, Front_laser, Back_laser, Wall_laser, and Charger_laser. Based on this information, a series of rotations and linear motions can be applied to the robot to move it to the desired location in front of the charger. The pseudo-code Algorithm 1 used to carry out the Lidar-based docking procedure is described as follows:

**Algorithm 1.** Lidar-based docking.

***State 1: Robot straightening***
*Initialize Front_laser, Back_laser, Charger_laser, and Wall_laser*
*If (Front_laser − Back_laser) > 0 then rotate clockwise until Front_laser = Back_laser*
*elseif (Front_laser − Back_laser) < 0 then rotate anti-clockwise until Front_laser = Back_laser*
*If Wall_laser >known_distance*
*Change state to 3*
*esleif Wall_laser <known_distance*
*Change state to 2*
***State 2: Robot turning left if to the right of the charger***
*Turn the robot anti-clockwise until Back_laser = Wall_laser*
*Then change state to 4*
***State 3:Robot turning right if to the left of the charger***
*Turn the robot clockwise until Front_laser = Wall_laser*
*Then change state to 4*
***State 4:Robot's linear motion***
*Move robot in a linear motion until Wall_laser = known_distance*
*Then change state to 5*
***State 5: Robot straightening second time***
*If (Front_laser − Back_laser) > 0 then rotate clockwise until Front_laser = Back_laser*
*elseif (Front_laser − Back_laser) < 0 then rotate anti-clockwise until Front_laser = Back_laser*
*Then change state to 6*
***State 6:Robot moving towards the charger***
*Move robot in a linear motion until charger_laser within 2 to 3 cm's away from the charger*
*Then change state to 7*
***State 7:Robot docking with the charger***
*Stop the robot's motion and change status to docked*

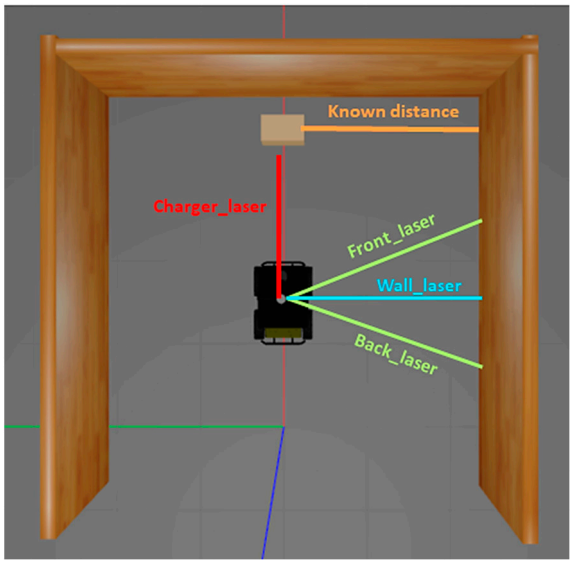


**Figure 7.** Lidar-based docking method visualization.

The above algorithm was tested on the Husky robot and gave a fairly accurate pose estimation and localization for docking, except for systematic errors based on the Lidar data readings. The algorithm was tested for various initial poses and locations from the charger; a few of the scenarios considered can be seen in Figure 8. The experiments conducted on the actual Husky robot are shown in Figure 9. The known distance of the charger from the side wall can be determined using the vision-based method and matched with the Wall_laser to fuse the Lidar–vision data. Moreover, once the robot is in the correct docking position for charging or close to the desired location, the Lidar point cloud data and the

camera-based 2D image can be calibrated to eliminate any errors and improve the pose estimation for the autonomous docking of the robot.

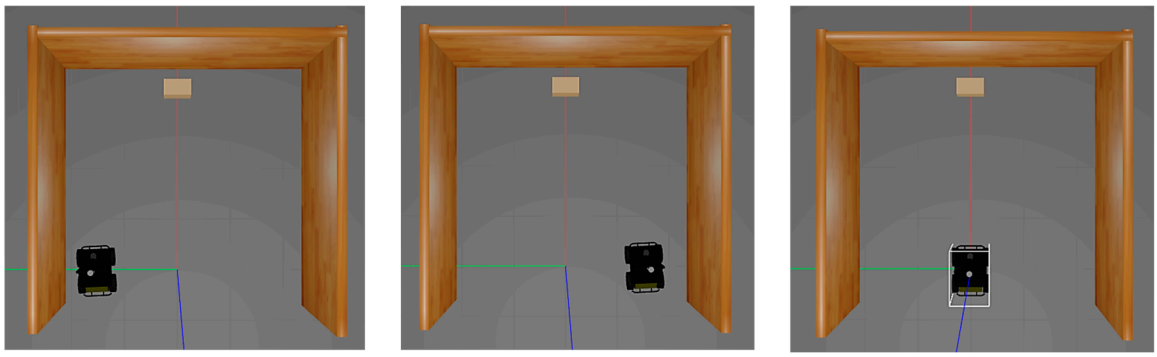

**Figure 8.** Robots in different locations and with different orientations from the charger in a Gazebo virtual environment setup.

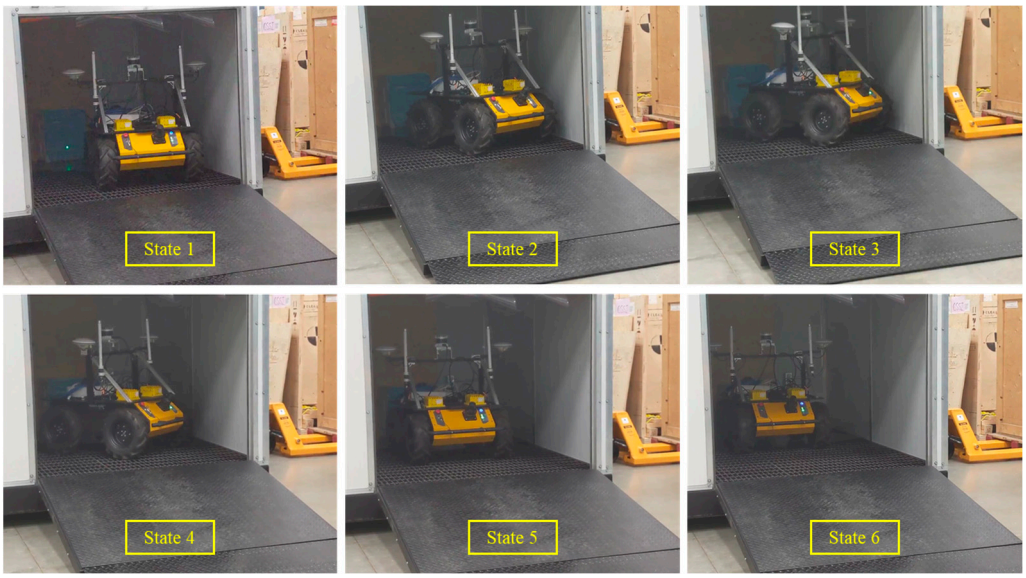


**Figure 9.** Example of actual docking.

## 5. Results

### *5.1. Transfer Learning and Data Augmentation*

Deep learning models frequently require extensive input images for the training process. However, gathering enough practical images for some applications can be difficult. Therefore, rather than building a model from scratch, transfer learning provides an alternative strategy for addressing this problem [30]. It uses a pre-trained deep learning model as a template for another training task. The modified YOLOv7 model is trained and tested on the Microsoft COCO dataset with the parameters used in this study, significantly improving training efficiency. Due to the limited number of charging stations, the images do not have extensive features. As a result, diversifying the training data is a common technique for improving generalization and reducing overfitting [31]. This study randomly introduces geometric distortions, such as rotation, translation, scaling, and vertical flipping, and image distortions, such as Gaussian blur and noise.

### *5.2. Datasets Building*

Since there are no public datasets for the charging stations used in this study, a specific dataset was built for the experiments. The images of charging stations were collected

through the Hikvision camera mounted on the mobile robot. The generated dataset has 240 images with a resolution of 1920 × 1018 pixels, shot from different angles and split into three sub-datasets: 160 training images, 40 validation images, and 40 testing images. The images in the dataset were annotated using LabelImg Software, which is an open-source annotation tool. The labelled images are shown in Figure 10.


**Figure 10.** Example of labelled images.

### *5.3. Training Environment and Parameters*

The model for detecting and recognizing the dock and charging stations was trained and tested on a local desktop with the specifications listed in Table 1. The pre-trained hyper-parameters for the dock and charging station detection are presented in Table 2.

**Table 1.** Training environment and specifications.

| Specifications | Value |
|---|---|
| Operating System | Windows Server 2019 |
| CPU | AMD Ryzen Threadripper 3970X 32-Core |
| GPU | NVIDIA GeForce RTX 3090 |
| RAM | 128 GB |
| CUDA Version | 11.1 |
| PyTorch Version | 1.10.1 |

**Table 2.** Training parameters.

| Parameters | Value |
|---|---|
| Learning Rate | 0.001 |
| Learning Momentum | 0.9 |
| Batch Size | 16 |
| Epochs | 100–300 |

### 5.4. Results and Analysis

#### 5.4.1. Evaluation Metrics

This paper adopts the mean average precision (mAP) as the evaluation metric. It is the area under the precision and recall (true-positive rate) curve, calculated according to Equation (5), at different intersection-over-union (*IoU*) thresholds. mAP_0.5, at a 0.5 intersection-over-union (*IoU*) threshold, is commonly used as the evaluation metric. In addition, mAP_0.5:0.95, which is the average mAP over multiple *IoU* thresholds, can affect the model, resulting in better performance. Therefore, both metrics are considered in the training and testing procedures to evaluate the performance of charging station detection.

$$IoU = \frac{Area\ of\ Overlap}{Area\ of\ Union} \tag{5}$$

$$AP = \int_0^1 (Precision \times Recall)\, d(Recall) \tag{6}$$

$$Precision = \frac{TP}{TP + FP} \tag{7}$$

$$Recall = \frac{TP}{TP + FN} \tag{8}$$

Here, *TP*, *FP*, and *FN* are the true-positive, false-positive, and false-negative results of the predicted bounding box, respectively.

#### 5.4.2. Results

Figure 11 depicts the training and validation loss for detecting the charging station. To optimize the proposed model, the loss function used in YOLOv7 [26] needs to be minimized. It is clear that, at around 300 epochs, the training and validation losses both decrease to a stable point, with a minimal gap between the two final values. Figure 12 displays the model performance based on both performance metrics of the model in the validation. It can achieve about 99.4% mAP_0.5 and 86.5% mAP_0.5:0.95. During the training and validation, various epochs were tested. It can be observed that, when the epoch is below 300, both the training loss and validation loss continue to decrease at the end of the curves, which indicates that the proposed model can be further improved through further learning. However, with the epochs surpassing 300, the validation loss begins to increase, which leads to an overfit. Therefore, in the training and validation, an epoch of 300 was chosen to obtain the optimized pre-trained model, which can achieve the best performance.

In addition, an evaluation is performed for real-time charger detection while the mobile robot is moving based on the proposed method. Figure 13 depicts an example of the recognized results. A metric for evaluating the method performance in a practical environment is adopted, as shown in Equation (9):

$$Accuracy = \frac{N}{T} \tag{9}$$

where $N$ is the number of correctly recognized images, and $T$ is the total number of images used in the evaluation process. It can be observed that, in real-time scenarios, the accuracy of the developed charging station detection method can achieve an average of 95%.

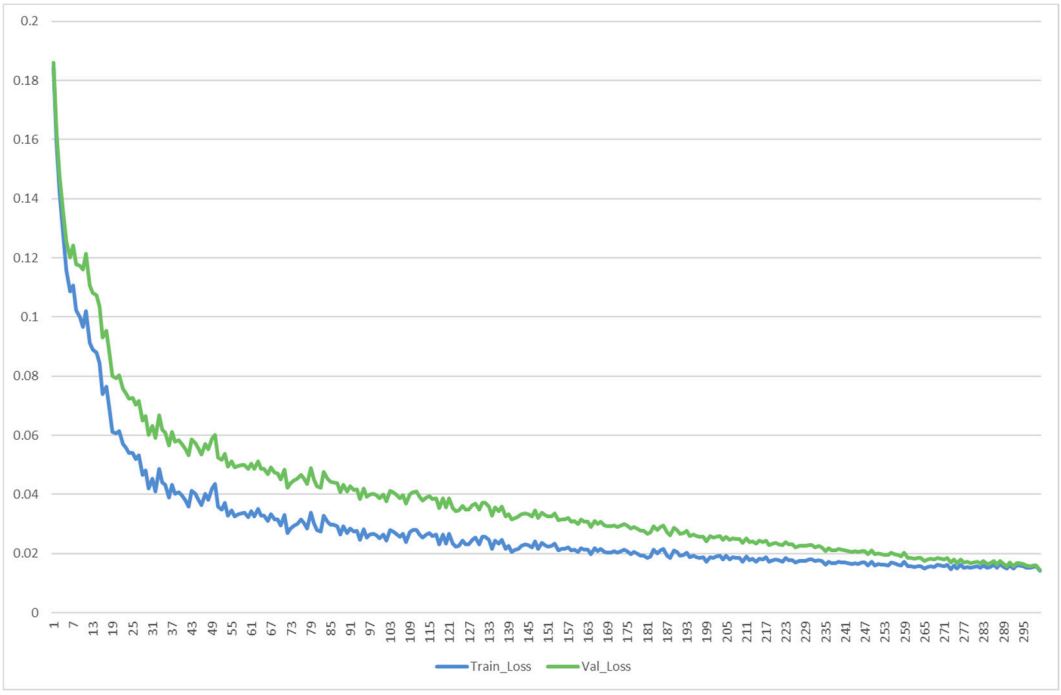


**Figure 11.** Training loss and validation loss of the charger detection model.

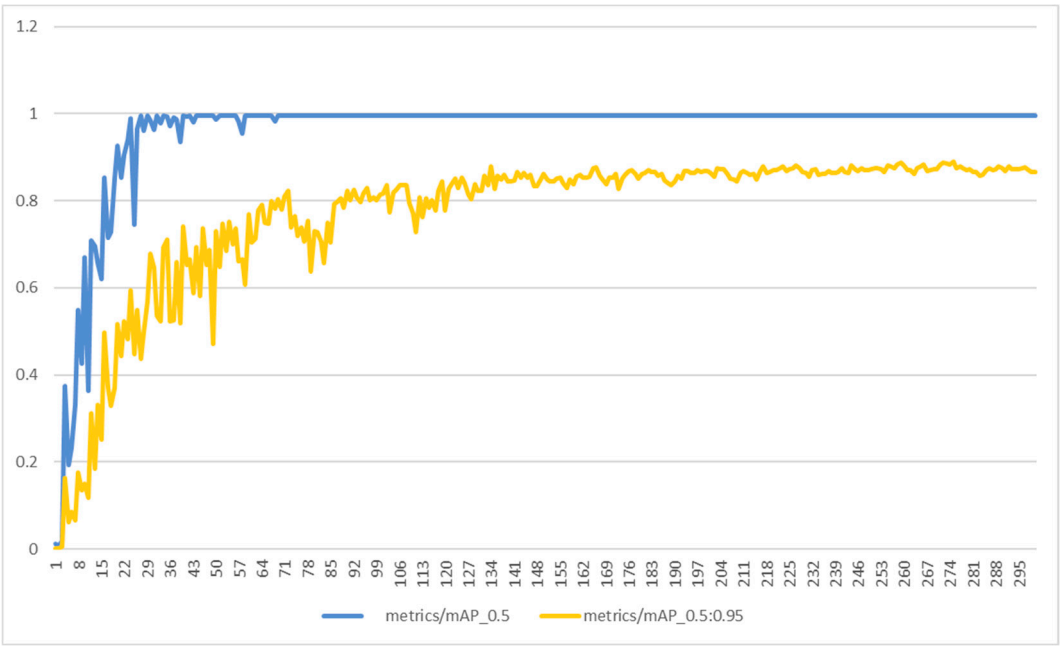


**Figure 12.** The results of both performance metrics.

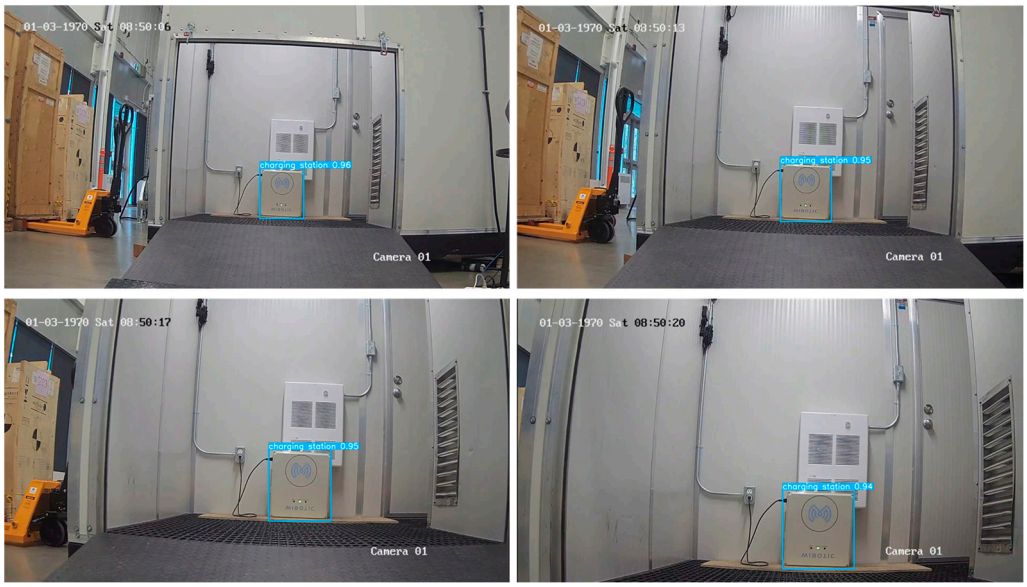


**Figure 13.** Example of real-time charging station detection.

## 6. Discussion and Conclusions

This paper discusses the challenges faced by current autonomous docking and recharging methods in the context of mobile robots in manufacturing environments. Current state-of-the-art methods heavily rely on Lidar, which makes it expensive and time-consuming for mobile robotic systems to achieve autonomous docking and recharging applications. Therefore, a Lidar and vision data fusion method, generated by combining deep learning object detection and Lidar-based docking approaches, was proposed to address the aforementioned problems. A YOLOv7-based real-time object detection model was developed to identify wireless chargers. To evaluate the developed detection method, a set of testing images and real-time video frames captured through a Hivision camera was used, and it achieved an average of 95% accuracy. The performance of the detection model for the charging station was compared with that of existing methods. According to the comparison results, the proposed method outperformed the other methods. A Lidar and vision data fusion approach was then developed to localize the wireless charger and then navigate the mobile robot to achieve docking to the charging station, reducing the computation costs of the system. Despite the advantages of the proposed method, it is limited by some challenges. For instance, the wireless charging station needs to be in an enclosed space, which can be used to calculate the wall_laser distance in the proposed method. Moreover, the developed charging station detection method can be affected by the low-illumination conditions in the manufacturing environment and by the blurring caused by the unstable movement of the mobile robot.

So far, this proposed Lidar–camera data fusion method for autonomous docking and recharging has only been validated on a 2D camera and a Lidar system. Future work will focus on the use of a stereo camera and Lidar system to improve the performance of the developed method in a practical autonomous manufacturing environment. Furthermore, for the docking procedure itself, to improve the pose estimation of the robot in relation to the charger, calibration between the vision and Lidar data needs to be implemented in future work.

**Author Contributions:** All authors contributed to the study's conception and design. Material preparation, data collection, and analysis were performed by F.J. The first draft of the manuscript was written by F.J. and M.A., and all authors commented on previous versions of the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Ministry of Economic Development, Trade, and Tourism of the Government of Alberta through the Autonomous Systems Initiative of the Major Innovation Funds, and the Go Productivity funding. The authors also would like to acknowledge the NSERC (Grant Nos. NSERC RGPIN-2017-04516 and NSERC CRDPJ 537378-18) for further funding this project.

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

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data are unavailable due to privacy concerns.

**Acknowledgments:** We express our gratitude to the Ministry of Economic Development, Trade, and Tourism of the Government of Alberta for funding this project through the Autonomous Systems Initiative of the Major Innovation Funds, and the Go Productivity funding. The authors also would like to acknowledge the NSERC (Grant Nos. NSERC RGPIN-2017-04516 and NSERC CRDPJ 537378-18) for further funding this project. The support of Anas Ahmed at Imperial Oil Limited in carrying out the work is gratefully acknowledged.

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

## References

1. Vargas, J.; Alsweiss, S.; Toker, O.; Razdan, R.; Santos, J. An Overview of Autonomous Vehicles Sensors and Their Vulnerability to Weather Conditions. *Sensors* **2021**, *21*, 5397. [CrossRef] [PubMed]
2. Rashid, E.; Ansari, M.D.; Gunjan, V.K.; Ahmed, M. Improvement in extended object tracking with the vision-based algorithm. *Stud. Comput. Intell.* **2020**, *885*, 237–245. [CrossRef]
3. Jia, F.; Tzintzun, J.; Ahmad, R. An Improved Robot Path Planning Algorithm for a Novel Self-adapting Intelligent Machine Tending Robotic System. In *Mechanisms and Machine Science*; Springer International Publishing: Cham, Switzerland, 2020; Volume 86, pp. 53–64. ISBN 9783030454029.
4. Yao, C.; Li, Y.; Ansari, M.D.; Talab, M.A.; Verma, A. Optimization of industrial process parameter control using improved genetic algorithm for industrial robot. *Paladyn* **2022**, *13*, 67–75. [CrossRef]
5. Guangrui, F.; Geng, W. Vision-based autonomous docking and re-charging system for mobile robot in warehouse environment. In Proceedings of the 2017 2nd International Conference on Robotics and Automation Engineering (ICRAE), Shanghai, China, 29–31 December 2017; pp. 79–83. [CrossRef]
6. Rubio, F.; Valero, F.; Llopis-Albert, C. A review of mobile robots: Concepts, methods, theoretical framework, and applications. *Int. J. Adv. Robot. Syst.* **2019**, *16*, 1–22. [CrossRef]
7. Abbasi, R.; Martinez, P.; Ahmad, R. The digitization of agricultural industry—A systematic literature review on agriculture 4.0. *Smart Agric. Technol.* **2022**, *2*, 100042. [CrossRef]
8. Maddikunta, P.K.R.; Pham, Q.-V.; Prabadevi, B.; Deepa, N.; Dev, K.; Gadekallu, T.R.; Ruby, R.; Liyanage, M. Industry 5.0: A survey on enabling technologies and potential applications. *J. Ind. Inf. Integr.* **2022**, *26*, 100257. [CrossRef]
9. Liu, Y. A Laser Intensity Based Autonomous Docking Approach for Mobile Robot Recharging in Unstructured Environments. *IEEE Access* **2022**, *10*, 71165–71176. [CrossRef]
10. Doumbia, M.; Cheng, X.; Havyarimana, V. An Auto-Recharging System Design and Implementation Based on Infrared Signal for Autonomous Robots. In Proceedings of the 2019 5th International Conference on Control, Automation and Robotics (ICCAR), Beijing, China, 19–22 April 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 894–900.
11. Luo, R.C.; Liao, C.T.; Lin, S.C. Multi-sensor fusion for reduced uncertainty in autonomous mobile robot docking and recharging. In Proceedings of the 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, St. Louis, MO, USA, 10–15 October 2009; pp. 2203–2208. [CrossRef]
12. Khan, S.; Wollherr, D.; Buss, M. Modeling Laser Intensities for Simultaneous Localization and Mapping. *IEEE Robot. Autom. Lett.* **2016**, *1*, 692–699. [CrossRef]
13. Hadi, R.H.; Hady, H.N.; Hasan, A.M.; Al-Jodah, A.; Humaidi, A.J. Improved Fault Classification for Predictive Maintenance in Industrial IoT Based on AutoML: A Case Study of Ball-Bearing Faults. *Processes* **2023**, *11*, 1507. [CrossRef]
14. Tibebu, H.; De-Silva, V.; Artaud, C.; Pina, R.; Shi, X. Towards Interpretable Camera and LiDAR Data Fusion for Autonomous Ground Vehicles Localisation. *Sensors* **2022**, *22*, 8021. [CrossRef] [PubMed]
15. Rao, M.V.S.; Shivakumar, D.M. Sensor Guided Docking of Autonomous Mobile Robot for Battery Recharging. *Int. J. Recent Technol. Eng.* **2019**, *8*, 3812–3816. [CrossRef]
16. Luo, R.C.; Liao, C.T.; Lin, K.C. Vision-based docking for automatic security robot power recharging. In Proceedings of the IEEE Workshop on Advanced Robotics and its Social Impacts, Nagoya, Japan, 12–15 June 2005; IEEE: Piscataway, NJ, USA, 2017; pp. 214–219.
17. Kriegler, A.; Wöber, W. Vision-Based Docking of a Mobile Robot. *Proc. Jt. Austrian Comput. Vis. Robot. Workshop* **2020**, 6–12. [CrossRef]

18. Mobile Industrial Robots A/S, MiRCharge 24V. Available online: https://www.mobile-industrial-robots.com/solutions/mir-applications/mir-charge-24v (accessed on 3 July 2022).
19. Fetch Robotics, Tutorial: Auto Docking. Available online: https://docs.fetchrobotics.com/docking.html (accessed on 24 May 2022).
20. Kartoun, U.; Stern, H.; Edan, Y.; Feied, C.; Handler, J.; Smith, M.; Gillam, M. Vision-Based Autonomous Robot Self-Docking and Recharging. In Proceedings of the 2006 World Automation Congress, Budapest, Hungary, 24–26 July 2006; IEEE: Piscataway, NJ, USA, 2006; pp. 1–8.
21. Song, K.T.; Chiu, C.W.; Kang, L.R.; Sun, Y.X.; Meng, C.H. Autonomous Docking in a Human-Robot Collaborative Environment of Automated Guided Vehicles. In Proceedings of the 2020 International Automatic Control Conference (CACS), Hsinchu, Taiwan, 4–7 November 2020. [CrossRef]
22. Yue, J.; Wen, W.; Han, J.; Hsu, L.-T. LiDAR Data Enrichment Using Deep Learning Based on High-Resolution Image: An Approach to Achieve High-Performance LiDAR SLAM Using Low-cost LiDAR. *arXiv* **2020**, arXiv:2008.03694.
23. Burgueño-Romero, A.M.; Ruiz-Sarmiento, J.R.; Gonzalez-Jimenez, J. Autonomous Docking of Mobile Robots by Reinforcement Learning Tackling the Sparse Reward Problem. In *International Work-Conference on Artificial Neural Networks*; Springer International Publishing: Cham, Swizerland, 2021; pp. 392–403. [CrossRef]
24. Zhou, L.; Zhang, L.; Konz, N. Computer Vision Techniques in Manufacturing. *IEEE Trans. Syst. Man Cybern. Syst.* **2023**, *53*, 105–117. [CrossRef]
25. Smith, M.L.; Smith, L.N.; Hansen, M.F. The quiet revolution in machine vision—A state-of-the-art survey paper, including historical review, perspectives, and future directions. *Comput. Ind.* **2021**, *130*, 103472. [CrossRef]
26. Wang, C.-Y.; Bochkovskiy, A.; Liao, H.-Y.M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *arXiv* **2022**, arXiv:2207.02696.
27. Reddy, B.K.; Bano, S.; Reddy, G.G.; Kommineni, R.; Reddy, P.Y. Convolutional Network based Animal Recognition using YOLO and Darknet. In Proceedings of the 2021 6th International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 20–22 January 2021; pp. 1198–1203. [CrossRef]
28. Xiao, J.; Guo, H.; Zhou, J.; Zhao, T.; Yu, Q.; Chen, Y.; Wang, Z. Tiny object detection with context enhancement and feature purification. *Expert Syst. Appl.* **2023**, *211*, 118665. [CrossRef]
29. Zheng, Y.; Mamledesai, H.; Imam, H.; Ahmad, R. A novel deep learning-based automatic damage detection and localization method for remanufacturing/repair. *Comput. Aided. Des. Appl.* **2021**, *18*, 1359–1372. [CrossRef]
30. Jia, F.; Ma, Y.; Ahmad, R. Vision-Based Associative Robotic Recognition of Working Status in Autonomous Manufacturing Environment. *Procedia CIRP* **2021**, *104*, 1535–1540. [CrossRef]
31. Jia, F.; Jebelli, A.; Ma, Y.; Ahmad, R. An Intelligent Manufacturing Approach Based on a Novel Deep Learning Method for Automatic Machine and Working Status Recognition. *Appl. Sci.* **2022**, *12*, 5697. [CrossRef]