ISSN: 3005-9852 *Journal of Robotics Spectrum 2(2024)*

Robot Docking and Charging Techniques in Real Time Deep Learning Model

[1]Michel Alain and [2]Amado Paul
[1, 2] School of Data science, PSL Research University, Rue Mazarine, Paris, France.
[2]paulmado@psl.eu

Correspondence should be addressed to Amado Paul: paulmado@psl.eu

Article Info
Journal of Robotics Spectrum (https://anapub.co.ke/journals/jrs/jrs.html)
Doi: https://doi.org/10.53759/9852/JRS202402002
Received 06 November 2023; Revised from 18 December 2023; Accepted 26 January 2024.
Available online 30 January 2024.

Abstract – This article describes various approaches that utilize computer vision and Lidar technology. These approaches include, but not limited to, vision-based algorithms such as the Faster RCNN model and AprilTag; and single shot detectors (SSD). In carrying out docking and recharging operations, the aforementioned approaches have shown varying degrees of success and accuracy. In order to make it easier for mobile robot systems to perform autonomous docking and recharging (ADaR) in industrial settings, this study presents a new method that employs vision and Lidar technology. In this study, we propose the YOLOv7 deep learning model to find charging stations. To further simplify docking with the specified wireless charging station, a Lidar-based approach is used to precisely modify the robot's position. An account of the assessment standards and training procedure used for the adjusted YOLOv7 model is provided in the results and discussion section. In this research, it was found that the model's 86.5% mean Average Precision (mAP) within the IoU range of 0.5 to 0.9 is evidence of its efficacy. In addition, the detection and identification of charging stations had an average accuracy rate of 95% in the studies conducted in real-world settings.

Keywords – Autonomous Docking and Recharging, Single Shot Detectors, Automated Driving Systems, Mobile Robots, Computer Vision, Mean Average Precision.

I. INTRODUCTION

There is a distinctive ability in mobile robots to explore vacant points within a given location. By using enhanced navigation systems and robust sensors, more adaptability to explore these points may be achieved. Docking is a common issue with mobile robots since it requires controlling three variables in a two-dimensional space: location, orientation, and velocity. An essential step in attaining long-term autonomy, the primary function of the docking station is the autonomous charging process. As discussed in this article, there have been extensive reviews of several docking approaches that make use of various sensors and algorithms in literature. To ensure module alignment during docking, previous studies have used pairs of infrared (IR) emitter receivers. In addition, the technology has been used due to its simplicity, portability, and low power consumption.

In this paper, "docking" refers to the process in which robots efficiency establish its orientation or position, and navigate to a predefined destination. Docking is different from long-distance route planning in that it does not need the employment of strategies, and focus to avoid obstacles. However, more focus should be dedicated to getting precise location and orientation estimates. According to Bakdi [1], once the robot's orientation and the intended destination's location are known inside a reference coordinate system, path planning algorithms may easily create control instructions. In the same study, the x-y pose (represented as $\vec{x}$) incorporates three independent variables (x, y, and θ), which stand for the z-axis rotation, inside the two-dimensional complex plane. What determines this posture is the system's condition at time t.

$$\left(x\dot{x}y\dot{y}\theta\dot{\theta}\right)_t^T = \vec{x}_t \tag{1}$$

The x-direction, y-direction, and rotational velocities of the robot are denoted by the variables $\dot{x}$, $\dot{y}$, and $\dot{\theta}$, respectively. According to Matthies and Shafer [2], there can be a lot of trouble figuring out the object's posture in relation to the measurement model and the motion model if a number of control steps are added without further measurements or observations. In order to fix this, the authors propose adding a measurement step after every control step, which would restore faith in the belief bel($\vec{x}$). The research conducted by the authors highlights the increasing popularity of computer vision (CV) based approaches for navigation in the context of automated driving systems (ADS). Recognition of artificial landmarks has been studied by Müller, Casser, Lahoud, Smith, and Ghanem [3], who found that it may be compared to current non-vision approaches such gradient-based optical flow. Conventional non-visual systems often use wireless fingerprinting methods, indoor GPS, or light detection and Ranging (Li-DAR) technologies. By deploying MiRs and Robotinos, LiDARs have shown to be extensively used in corporate settings. However, the scientific world is becoming more and more intrigued by deep learning's successes and their applications in ADS.

Vaz, Ferreira, Grossmann, and Ribeiro [4] detected passive reflectors connected to the docking station using a mobile platform equipped with two infrared scanners. Using triangulation, the position of the mobile robot with respect to docking stations is continuously determined. Two infrared (IR) detectors connected to the docking station were suggested as part of a comparable technique in an earlier study by Kwak [5]. Once two infrared sensors are activated, the surveillance robot changes its orientation by 90 degrees to be in line with the docking station. Up until then, it moves in rhythm with the station.

According to Fan and Ishibashi [6], the docking station is typically equipped with a total of 5 IR-LEDs (infrared light-emitting diodes), thereby dividing the docking spaces into 9 regions. Furthermore, the docking station has two clasps and a toggle switch that let the robot stay in alignment by adjusting for angle and offset errors as it gets closer. Six infrared (IR) sensors, mounted atop the robot's receiver and distributed across its front and sides, make up the receiver's optical system. Kobuki uses three infrared (IR) sensors installed on the TurtleBot as his docking method. Three infrared emitters placed at key angles on the docking station complement these receivers [7]. The IR emitters mounted on the docking station effectively segment the docking field into three separate regions. The local and distant environs are defined to differentiate each geographical area. Sunlight interference and infrared beam reflection are two sources of inaccuracy that may affect inexpensive infrared sensors.

An alternative approach to this problem may be found in vision-based docking systems. The combination of object recognition algorithms with navigation control techniques was the main subject of a research conducted by Dimitrov, Wills, and Padır [8]. According to Dunbabin, Lang, and Wood [9], the docking system employs a vision-based technology to detect and recognize a visual tag located on the docking station. The determination of the robot's trajectory during the docking approach is contingent upon obtaining the relative location and orientation of the tag. The research modelled the robot's path using a cubic Bezier curve. Where the robot was when it detected the tag is represented by the beginning point of the curve, and where the tag was at the same time is represented by the ending point. The aforementioned study included a comparative assessment of the docking technique's performance utilizing various tags. Having enough light is a need for conventional vision-based systems.

There are several open source implementations available for various learning challenges, which presents great opportunity to enhance robot systems' perceptual abilities. Convolutional Neural Networks (CNNs) can acquire many visual targets, which could vary based on the specific environment or application, thanks to their versatile feature extraction capabilities. This approach helps overcome the limitation of using visually designed markers, which is required by conventional computer vision techniques. This study has a well-defined and specific goal for the object detector. Recognizing and categorizing a subset of logos that stand out visually from the others inside the given space is what this task is all about (see **Fig 1**). In addition, a virtual target is included to allow the robot to precisely monitor its movements throughout the docking process, which helps with dynamic movement planning. So, the virtual target frame, $V = \{\overrightarrow{xV}, \overrightarrow{yV}\}$,, is also set up. **Fig 2** shows the links between the frames mentioned before.

Previous research has used the mobile robot LiDAR sensor, namely the robotino, in order to provide a thorough depiction of the immediate surroundings. Moreover, the attainment of localization was successfully executed by the use of the AMCL package. The LiDAR sensor is capable of operating in both lit and non-illuminated settings. In [10] employed the Censil Canonical Scan Matcher and LiDAR equipment in their experiment. The [11] introduced a novel approach to docking that integrates landmarks as key elements. To enhance the efficiency of dock extraction, the attributes linked to the dock are determined by a mix of Euclidean clustering and nearest neighbor analysis. The procedure of aligning the dock model with the specified lines necessitates doing a comprehensive examination across all clusters. The present methodology employs an iterative closest point strategy for ascertaining the transformations between robot and dock coordinate frames.

The subject matter pertaining to the mobile robot systems of autonomous docking and recharging within firm settings possesses considerable importance in enhancing productivity and diminishing the necessity for human intervention. The existing techniques provide several limitations in relation to the evaluation context and operational effectiveness. The objective of this research is to address the current limitations by developing an improved real-time deep learning model called YOLOv7, focusing specifically on the detection and identification of charging stations. Furthermore, the amalgamation of this model has been accomplished by the utilization of a Lidar-based approach. The primary aim of the proposed technique is to offer economical, efficient, and robust docking and recharging processes inside manufacturing environments. The subsequent sections of the article have been organized in the following format: The second section provides a comprehensive assessment of the existing literature pertaining to the autonomous docking and recharging (ADaR) of robotics. This article delineates the procedures used throughout the composition process in the third section. The fourth section of the study analyzes the results related to the dataset's development and training environment, together with the evaluation metrics used to evaluate the acquired results. The fifth section concludes the study comprehensively and delineates prospective directions for further research.

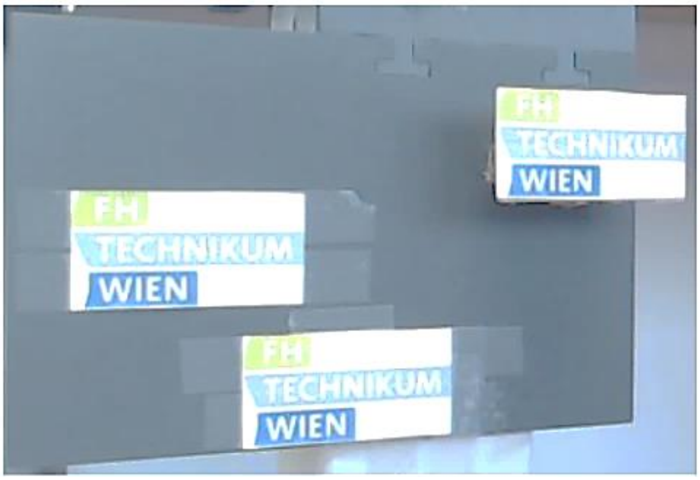

Fig 1. The Visual Target Used in The Process of Docking.

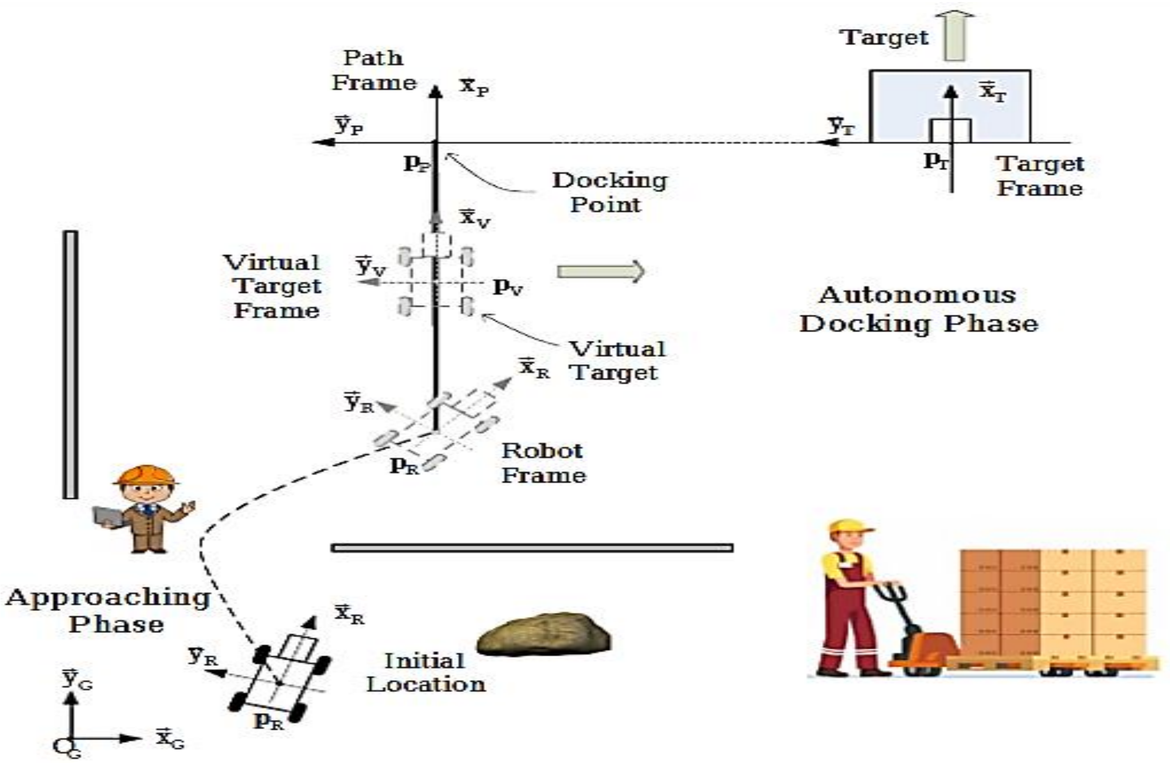

Fig 2. The Process of Target Docking.

II. LITERATURE REVIEW

According to Melenbrink, Werfel, and Menges [12], the investigation into the phenomenon of recharging mobile robots can be traced back to the 1950s, a significant period marked by the first advancements in the creation of self-recharging autonomous mobile robots. The robots were outfitted with a luminous beacon that functioned as a navigational aid, directing them towards a battery recharging station located inside a containment facility. Most of the currently available autonomous recharging systems have been designed based on similar principles, whereby sensors are used to detect a station of charging and enable the process of docking. In this paper, an analysis of scholarly literature on ADaR, focusing specifically on the wide range of sensor advancements used in this field of study, is provided. Prior studies have investigated the use of infrared (IR) methods for the analysis of ADaR in mobile robots. The present study employed a total of 3 IR receivers and 3 transmitters for the purpose of inquiry.

A detailed analysis was undertaken by Vaz, Ferreira, Grossmann, and Ribeiro [13] to examine the positions and orientations of three infrared (IR) receivers with the aim of optimizing the docking process through a reduction in time requirements. The algorithm constructed was subjected to an extensive verification process, which encompassed the execution of a series of rigorous experiments within a controlled laboratory environment. In their study, Roh et al. [14] established a setup wherein a collection of infrared (IR) emitters were securely attached to the docking station, while an IR receiver was positioned at the posterior end of the robot. The arrangement mentioned above was employed to enable the autonomous homing approach. Two novel docking methodologies were proposed as prospective remedies for mitigating lateral and directional docking errors. The efficacy of these activities is contingent upon the frictional forces that arise between the docking components of the docking station and the mobile robot.

The experimental validation presented empirical support for the effectiveness of the proposed technique. Within the domain of vision-based techniques, many scholars has presented a proposition for an ADaR system that is tailored to meet the demands of mobile robots functioning within warehouse environments. The integration of a fiduciary marker system, namely the AprilTag system, was used to augment the navigating capabilities of the mobile robot towards the charging station. The previously described equipment was strategically placed in close proximity to the charging station. The methodology employed in this technology entails the extraction of markers from the image obtained by the camera incorporated into the device.

The study conducted by Olatomiwa, Mekhilef, Ismail, and Moghavvemi [15] revealed that the algorithm devised by the researchers shown favorable performance inside warehouse environments. Nevertheless, it is important to acknowledge that discrepancies in lighting circumstances and the existence of tag shielding have the potential to impede the docking procedure. In their study, Kartoun et al. [16] designed an ADaR system that used vision-based technology. This system was specifically designed to be implemented on a surveillance robot. To enhance the mobile robot's ability to identify the station of charging, a circular marker of predetermined diameter and navy blue hue was deliberately positioned on the docking station. The main aim of installing the vision system was to identify and categorize man-made landmarks, with the eventual purpose of accurately determining the location of the charging station. This work introduces a proposed control mechanism with the objective of addressing horizontal and rotational errors that may arise during the docking procedure. The proposed technique suggests the existence of a virtual spring link between the station and the robot. The regulation of motion in the system is determined by the forces produced by compliant bending and deformation in the direction of translation.

The experimental validation conducted by Wang et al. [17] served to confirm the success of the recommended docking control technique. Within realm of laser-based techniques, the mobile robots employ an indentation that is V-shaped on 24V charger to enhance the perception of the charging ports through the integration of an intrinsic laser range finder. Similarly, Fetch Robotics deemed it essential to physically isolate their charging dock from any potential obstructions that may impede the laser range finder's capacity to precisely discern the charging station's configuration. Moreover, the recharging methodology suggested by Fetch robots imposes some limitations on the practical implementation of the robot in unstructured settings.

In the research conducted by Guangrui and Geng [18], a novel methodology was introduced that employs vision-based methods to enhance the efficiency of docking and recharging procedures. The approach used in this study was specifically designed to be utilized within the confines of a warehouse setting. The current study used the AprilTag methodology to effectively ascertain and perceive the orientation of the robot. The percentage of successful docking operations was around 97.33%. Liu, Özay, Okatani, Xu, Sun, and Yang [19] introduced a Faster RCNN model in their research, which aims to accurately detect and precisely localize markers specifically developed and installed on a docking station. The incorporation of this model into the solvePnP methodology enabled the mobile robot to navigate well inside a simulated habitat using the ROS framework. The model of accuracy was evaluated to be 96.3% after the analysis of thirteen testing images. The detector exhibited an average processing time of around 35 milliseconds per each image. In their study, Liu et al. [20] used an approach known as the single-shot detector (SSD) to recognize and track individuals in motion. The methods described above was used to improve the efficacy of human-robot collaboration in a context without a pre-established framework. The present work included the development and deployment of a novel solid-state detector (SSD) for the purpose of identifying outlets of charging in areas devoid of obstacles [21]. The approach demonstrates a high level of proficiency in establishing a connection with the charger, resulting in a success percentage of 99.8%. The average duration of the docking procedures, as calculated under the required conditions, was found to be 12 seconds.

While significant progress has been made in the sector of autonomous docking and recharging applications, some limits have been recognized. The assessment of the majority of processes commonly takes place within a controlled setting, such as a simulation, as opposed to an authentic production context. Furthermore, empirical research has demonstrated that one-stage real-time models outperform two-stage deep learning models, like Faster RCNN, in terms of efficiency. Given the aforementioned limitations, researchers have developed an improved DL model called YOLOv7 to accurately detect and categorize the specific wireless charger inside a complex manufacturing environment. The successful implementation of this model in conjunction with the suggested Lidar-based approach has led to a seamless completion of the docking and recharging procedure, notable for its cost-efficiency and durability.

Based on the findings of Wang, Bochkovskiy, and Liao [22], their study demonstrates that YOLOv7 has enhanced performance in terms of accuracy and velocity when compared to existing object recognition models. The findings suggest that YOLOv7 has a notable improvement in speed, exhibiting an estimated 1200% increase in efficiency relative to alternative models during evaluation on the MS COCO min-val dataset (see **Fig 3**). In addition, YOLOv7 exhibits a significant increase in computational efficiency, as seen by performance gains of about 115% and 145% observed in MSCOCO object identification tasks, as depicted in **Fig 4**. The YOLOv7 object identifier exhibits a significant balance between accuracy and efficiency across a wide range of frame rates, spanning from 5 to 160 frames per second (FPS).

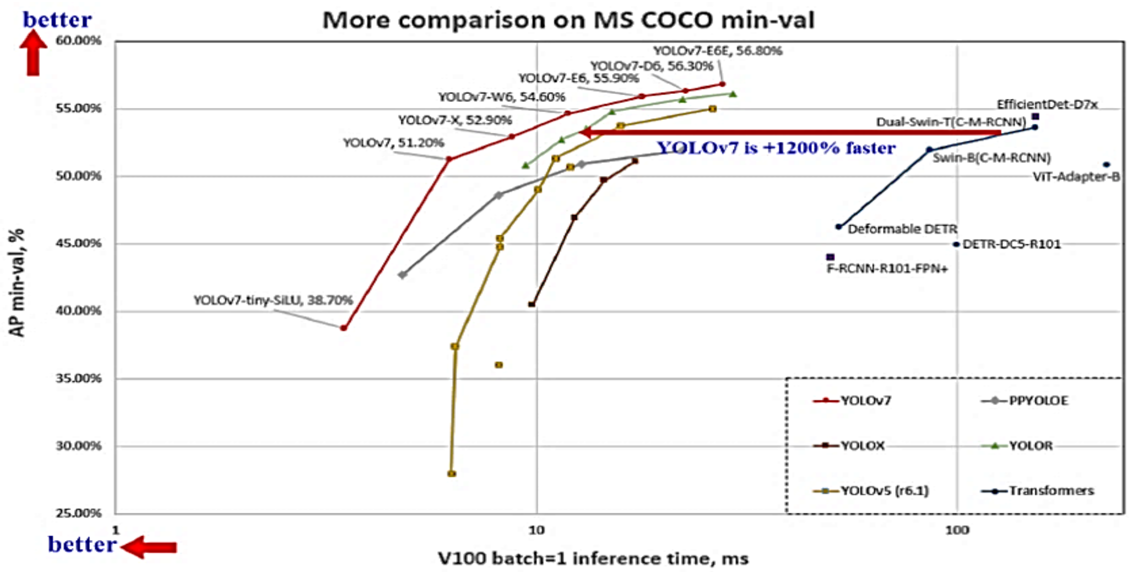

Fig 3. Comparison of YOLOv7 with Other Detectors.

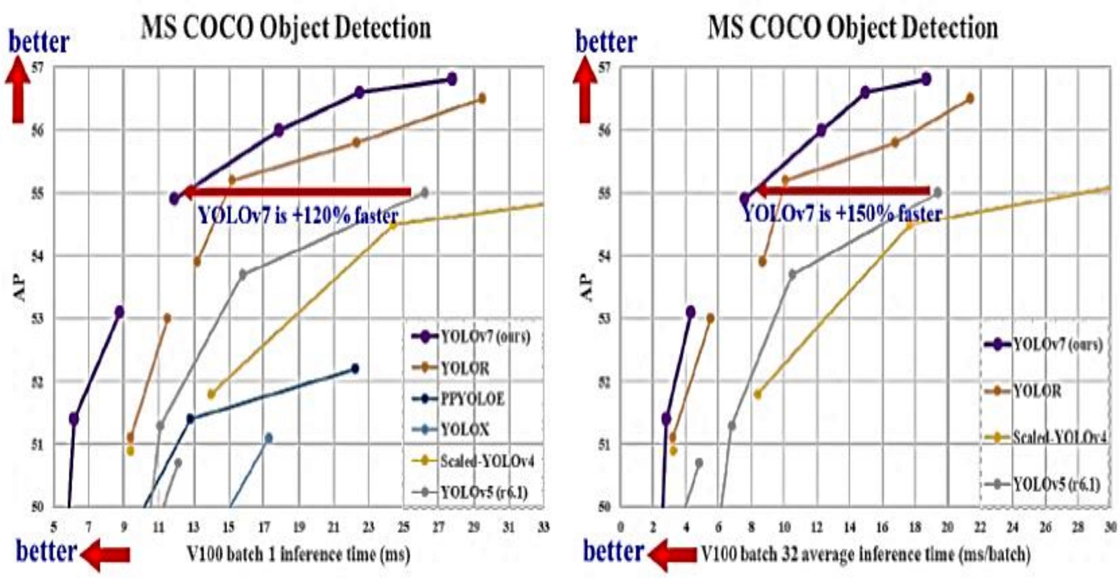

Fig 4. Comparison of YOLOv7 Object Detectors in Real-Time.

The real-time object identification system has a notable AP of 56.8% while functioning at a 25 FPS or above on GPUV100. The YOLOv7-E6 object detector is characterized with higher performance in comparison to SWIN-L CMR-CNN (Cascade-Mask R-CNN), a transformer-built detector and the ConvNeXt-XL CMR-CNN, a convolutional-based detector. When considering velocity, YOLOv7-E6 demonstrates much better performance in comparison to SWIN-L CMR-CNN by a margin of 508% and the ConvNeXt-XL CMR-CNN by 550%. Furthermore, YOLOv7-E6 exhibits a higher level of accuracy, surpassing SCMR-CNN by 2% and ConvNeXt-XL CMR-CNN by a little margin of 0.7%. Furthermore, YOLOv7-E6 demonstrates enhanced performance in terms of both speed and accuracy when compared to many other object detectors, including but not limited to ViT-Adapter-B, YOLOR, DINO-5scale-R50, YOLOX, Deformable DETR, YOLOv5, Scaled-YOLOv4, DETR, and numerous other models. In addition, our training protocol exclusively prioritizes the training of YOLOv7 using the MS COCO dataset.

The concept of ADaR in mobile robot systems within industrial environments has substantial importance in terms of minimizing human intervention and enhancing productivity. The existing techniques provide several limitations in terms of the evaluation framework and efficacy. The objective of this research is to address the current limitations by proposing the utilization of a sophisticated real-time DL model, specifically YOLOv7, for the effective identification of stations of charging. Furthermore, the integration of the model has been accomplished by the utilization of a Lidar-based methodology. The primary aim of the proposed methodology is to offer efficient, economical, and robust docking and recharging techniques inside industrial environments.

III. METHODOLOGY

This section presents a methodology for achieving autonomous docking and charging via the use of Lidar and vision technologies. The suggested methodology consists of three main steps. Firstly, data collection is performed using Ouster Lidar and a Hikvision camera. These devices capture laser distance measurements and RGB images of the surrounding environment. Secondly, a DL-based object detection technique is employed to locate the charging station within the manufacturing habitat. This technique utilizes the YOLOv7 model as its core architecture. Lastly, the mobile robot's orientation is adjusted using Lidar to establish a connection with the wireless charger. **Fig 5** illustrates a flowchart that represents the suggested methodology.

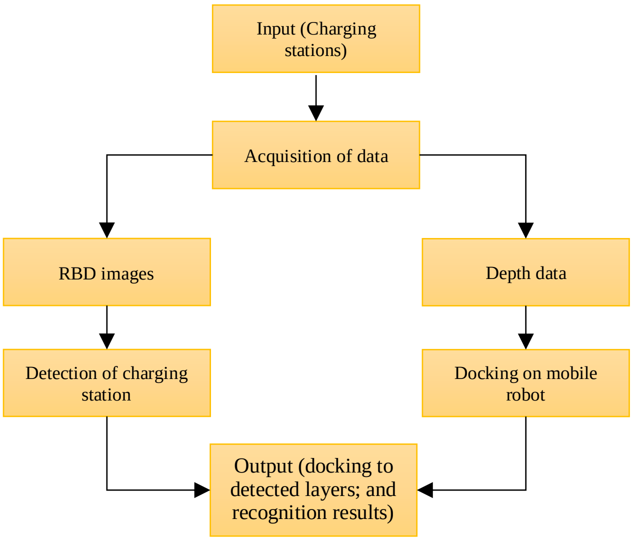

Fig 5. An Overview of The Recommended Docking and Recharging Technique.

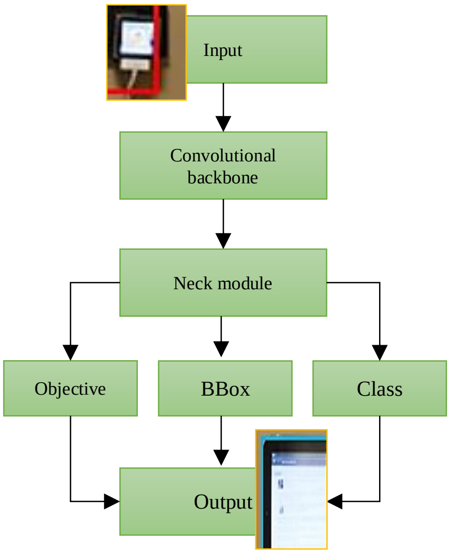

Fig 6. An Overview of The Charger Detection Process.

The most recent technique devised for real-time object identification, YOLOv7, is a contemporary one-stage methodology with high precision and pace [23]. The architectural design of the charging station's detecting mechanism, as seen in **Fig 6**, is based upon the YOLOv7 framework. The three primary components of the entity in question consist of the cranium, the cervical vertebrae, and the vertebral column. The feature maps collected from the input image are fed into the neck layers of the convolutional backbone model, using the Darknet-53 architecture. The neck module utilizes the Feature Pyramid Network (FPN) for the purpose of enhancing the feature maps. Subsequently, the maps are consolidated and sent to the subsequent hierarchical levels. The core neural network is ultimately tasked with making predictions regarding the bounding boxes and classifications of the observed items.

To optimize the performance of inference, the YOLOv7 model incorporates an improved extended efficient layer aggregation network. Without altering or disrupting the original gradient propagation pathway, this network has the potential to enhance the model's ability to learn. The current study introduces a new approach, referred to as Compound Model Scaling (CMS), to mitigate the issue of expanded width of output in the computation block. The primary focus of the suggested methodology is in the modification of the depth of the concatenation-based model. Furthermore, several approaches are employed to not only reduce training costs but also improve the accuracy of inference. Batch normalization, dynamic label assignment, and deliberate re-parameterization are three techniques that are together referred to as “Bags of Freebies” (BoF). The authors provide empirical evidence that supports the enhanced accuracy of the model when using RepConv without an identity link. This finding is derived from a comprehensive analysis of the re-parametrized convolution technique. Moreover, a technique of batch normalization that alters the bias and weight of the convolutional layer by including the variance and mean of the data. The potential effect of this modification on the training process is significant due to its ability to enhance the training rate and expedite convergence.

In contrast to other contemporary approaches, YOLOv7 demonstrates enhanced efficacy in the inference procedure, resulting in expedited real-time object recognition with heightened precision. The improvement may be attributed to the advanced training approach and network architecture used. The prospective applications of ADaR, however, remain unexplored. This article demonstrates the methodology for locating and classifying charging stations by using the YOLOv7 framework as the underlying architecture.

IV. RESULTS AND DISCUSSION

The training procedure for deep learning models sometimes necessitates a significant quantity of input images. However, the procurement of a sufficient number of appropriate images for some purposes might be a significant challenge. Therefore, rather than building a model from scratch, transfer learning presents an alternative technique to address this problem. The use of an established DL model as a fundamental framework for a potential training endeavor is seen. The model of YOLOv7 was employed and applied to the task of evaluating and training on the Microsoft COCO dataset, utilizing the settings outlined in this study. The aforementioned alteration has led to a significant improvement in the efficacy of the training procedure. The limited availability of charging stations imposes limitations on the characteristics of the images. The use of different training data is a frequently utilized strategy to mitigate the problem of overfitting and enhance the generalization performance, as shown by the investigation done by Seyfioğlu, Erol, Gürbüz, and Amin [24]. The current investigation employs a randomized implementation of geometric distortions, including vertical flips, scaling, translation, and rotation, in conjunction with image distortions, namely Gaussian blur and noise.

Training Environment and Datasets Building

Due to the limited availability of publicly accessible data on the charging stations being reviewed, a separate set of data was specifically created for performing the research. The procurement of images depicting a charging stations was done using the Hikvision camera, which was firmly affixed onto a robot. The set of data comprised a total of 230 images. The assortment consists of images that possess a resolution of 1920 × 1018 pixels, which have been acquired from several viewpoints. The dataset has been partitioned into 3 distinct sub-sets (totaling 230 images); training (160); testing (40); and validation (30). The images contained within the collection were annotated utilizing the open-source annotation software referred to as LabelImg Software. The images that have been provided with detailed descriptions are shown in **Fig 7**.

Fig 7. Example of Labelled Images.

An on-site desktop computer that met the requirements listed in **Table 1** was used for both the model's assessment and training in order to identify and recognize the dock and charging station. The pre-trained hyperparameters used for charging and docking station detection are shown in **Table 1**.

Table 1. Conditions & Requirements and Parameters for Training

Element	Specifications	Value
Conditions & Requirements for Training	PyTorch version	1.10.1
	CUDA Version	11.1
	RAM	128 GB
	GPU	NVIDIA G-Forece RTX-3090
	CPU	AMD Ryzen 3970X 32-core
	OS	Windows Server 2019
Parameters of Training	Epochs	100 to 300
	Batch size	16
	Training momentum	0.9
	Training rate	0.001

Results Metrics

The evaluation metric employed in this study is the mAP (mean Average Precision). This metric under consideration relate to the area computation of recall and accuracy. This work revolves around the TP rate, as described in (2), under different thresholds of IoU (intersection over union). The commonly employed assessment metric is mAP_0.5, which quantifies the mAP by the utilization of a threshold of 0.5 for the IoU. In addition, the mAP within a certain range of IoU values, (i.e., mAP_0.5:0.9), has the capability to iaffect the model. Hence, the integration of these signals is essential in the assessment and refinement of the testing and training procedures to ascertain the efficacy of charging station identification.

$$I_oU = \frac{A_oO}{A_oU} \tag{2}$$

where A_oO is the Area of Overlap and A_oO is the Area of Union.

$$AP = \int_0^1 (R.P)4(R) \tag{3}$$

where R is recall and P is the precision

$$P = \frac{TP}{FP + TP}; R = \frac{TP}{FN + TP} \tag{4}$$

Within this framework, the designations FP, TP, and FN correspond to the results of the anticipated bound box, precisely signifying False Positive, True Positive, and False Negative results.

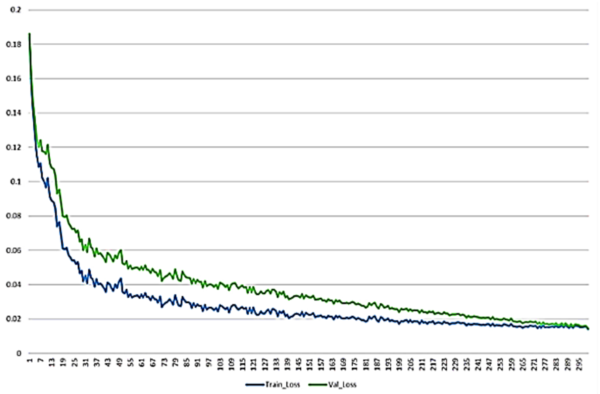

Fig 8. The Metrics Pertaining to The Charger Detection Model Include the Validation Loss and Training Loss.

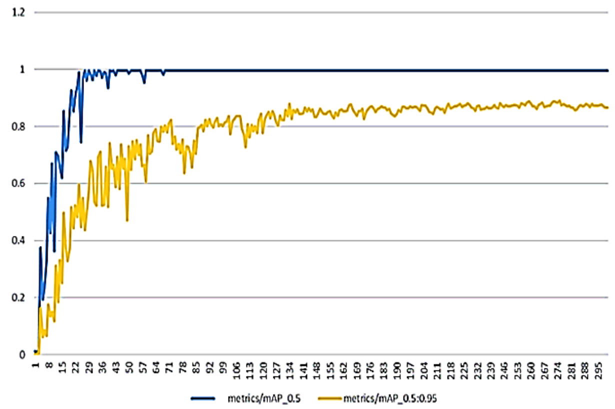

Fig 9. Performance Metrics Results.

Fig 8 depicts the validation and training loss pertaining to the identification of the charging points. To optimize the performance of the proposed model, it is imperative to minimize the loss functions in YOLOv7. The findings suggest that within approximately 299 epochs, the validation and training loss reach a point of convergence, displaying minimal disparity in their ultimate values. **Fig 9** depicts the assessment of the model's efficacy during the phase of validation, employing both performance metrics. The mAP attained at an IoU criterion of 0.5 is approximately 99%. However, when evaluating the mAP across IoU criteria ranging from 0.5 to 0.9, the resulting mAP score is around 86.5%. During the process of validation and training, several epochs were assessed. It is apparent that, for epochs less than 300, there is a constant decrease observed in both the validation loss and the training loss towards the respective curves end. This implies that the proposed model exhibits promise for future improvement through additional learning. However, as the value of epochs surpasses 300, the validation loss begins to increase, suggesting the occurrence of overfitting. Hence, to get an ideal pre-trained model with the highest degree of performance, a cumulative of 300 epochs were used for both the training and validation phases.

Moreover, a thorough evaluation is performed to examine the impact of the proposed methodology in real-world situations for the purpose of charger detection, specifically during the movement of the mobile robot. **Fig 10** illustrates the recognized consequences. The evaluation of technique performance in a real-world context utilizes a statistical measure referred to as N/T. Here, N indicates the number of accurately identified images, while T reflects the total number of images employed in the assessment. The charging station detection system has a significant degree of accuracy, as shown by an average success rate of 95% in real-world circumstances.

This research paper introduces a novel methodology for implementing ADaR control in mobile robot systems within the industrial domain. The method being presented capitalizes on the convergence of Lidar and vision technology. The methodology comprises three primary stages: localization refinement of the robot to accomplish docking with the wireless charger by utilizing measurements of Lidar, object recognition employing the YOLOv7 model based on DL techniques and data acquisition employing technologies and camera and Lidar. The model of YOLOv7 is utilized for the aim of detecting stations of charging in real-time, owing to its capacity for object detection. The software completed a testing and training procedure utilizing a specifically selected dataset comprising images of charging stations. The evaluation metric employed in this research is the mAP that comprises both mAP_0.5 and _0.5: 0.9. The model has a mAP of approximately 99% when the threshold is set at 0.5. Furthermore, it attains a mAP of 86% within the threshold range of 0.5 to 0.95.

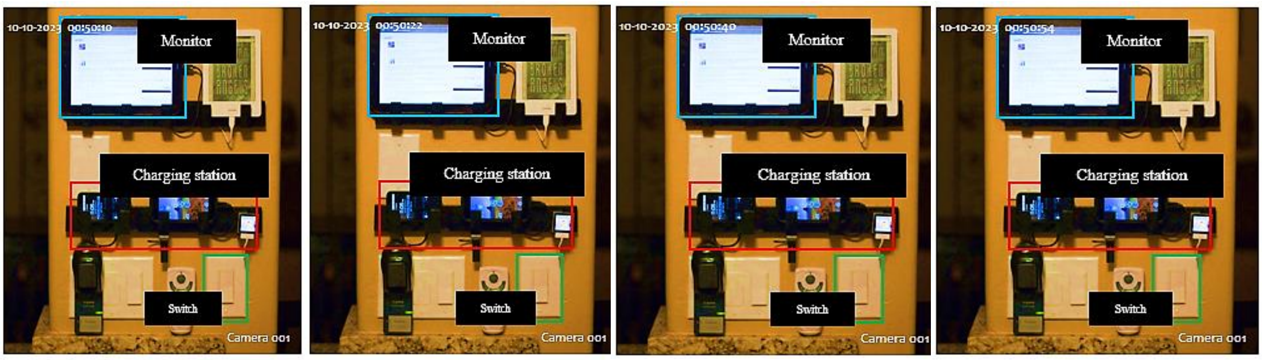

Fig 10. Real-Time Charging Stations.

The effectiveness of the suggested methodology is evaluated in practical scenarios, demonstrating a charging station identification accuracy of about 95%. The proposed methodology aims to combine Lidar and vision techniques with DL algorithms in order to deliver economically viable, dependable, and streamlined docking and recharging procedures for systems of mobile robot within the firm domain. The usage of the model of YOLOv7 in the detection of charging outlets exhibits promising results in terms of accuracy and operational efficiency in real-time scenarios.

V. CONCLUSIONS AND FUTURE SCOPE

This paper investigates the problems faced by present ADaR methods in the mobile robots context in industrial habitats. The implementation of Lidar in the context of autonomous docking and recharging poses significant issues for mobile robotic systems, primarily due to their heavy dependence on this specific technology. Consequently, the deployment of such technologies leads to heightened financial costs and temporal challenges. Therefore, a unique approach was employed to address the aforementioned challenges by merging deep learning approaches for object recognition with Lidar-based docking procedures. This integration resulted in the formulation of a fusion methodology that effectively combines Lidar and vision data. The YOLOv7 framework was employed to develop a real-time object recognition model, specifically focused on accurately detecting and classifying wireless chargers. In order to evaluate the impact of the proposed detection system, a dataset consisting of real-time video frames and testing images acquired from a Hivision camera was utilized. The technique exhibited a statistically significant mean accuracy rate of 95%. A comparative study was done to assess the effectiveness of the charging station detection model in comparison to known approaches.

The findings derived from the comparative analysis indicate that the proposed model exhibited superior performance when compared to the other alternatives. Following this, a technique was established to effectively combine optical and Lidar data in order to accurately determine the wireless charging system location and enable the mobile robot to navigate towards the docking station for recharging. The use of this method effectively mitigates the computational expenses associated with the system. Although the suggested technique offers certain benefits, it is important to acknowledge that it is also constrained by certain constraints. In this scenario, it is imperative that the wireless station of charging be positioned within a restricted enclosure. The enclosed space functions to determine the exact distance between the laser and the wall, as outlined in the recommended methodology.

At now, the evaluation of the proposed model for ADaR has been limited to the use of a Lidar system and a 2D camera. In future research initiatives, it is recommended that priority be given to the integration of a Lidar system and a stereo camera. This integration aims to advance the precision of the robot's location assessment in relation to the charger during the docking operation. It is advisable that next study places emphasis on prioritizing the investigation of calibration procedures aimed at synchronizing vision and Lidar data.

Data Availability
No data was used to support this study.

Conflicts of Interests
The author(s) declare(s) that they have no conflicts of interest.

Funding
No funding agency is associated with this research.

Competing Interests
There are no competing interests.

References
[1]. A. Bakdi, A. Hentout, H. Boutami, A. Maoudj, O. Hachour, and B. Bouzouia, "Optimal path planning and execution for mobile robots using genetic algorithm and adaptive fuzzy-logic control," Robotics and Autonomous Systems, vol. 89, pp. 95–109, Mar. 2017, doi: 10.1016/j.robot.2016.12.008.

[2]. L. Matthies and S. A. Shafer, "Error modeling in stereo navigation," IEEE Journal on Robotics and Automation, vol. 3, no. 3, pp. 239–248, Jun. 1987, doi: 10.1109/jra.1987.1087097.

[3]. M. Müller, V. Casser, J. Lahoud, N. Smith, and B. Ghanem, "SIM4CV: a Photo-Realistic Simulator for Computer vision applications," International Journal of Computer Vision, vol. 126, no. 9, pp. 902–919, Mar. 2018, doi: 10.1007/s11263-018-1073-7.

[4]. P. M. Vaz, R. Ferreira, V. Grossmann, and M. Ribeiro, "Docking of a mobile platform based on infrared sensors," ISIE '97 Proceeding of the IEEE International Symposium on Industrial Electronics, Nov. 2002, doi: 10.1109/isie.1997.649089.

[5]. Y. K. Kwak, "Development of docking system for mobile robots using cheap infrared sensors," 1st International Conference on Sensing Technology, p. 0, Nov. 2005, [Online]. Available: https://koasas.kaist.ac.kr/handle/10203/6285

[6]. F. Fan and S. Ishibashi, "Underwater applications of light emitting diodes," 2015 IEEE Underwater Technology (UT), Feb. 2015, doi: 10.1109/ut.2015.7108259.

[7]. M. L. Doumbia, C. Xu, and V. Havyarimana, "An Auto-Recharging System Design and Implementation Based on Infrared Signal for Autonomous Robots," 2019 5th International Conference on Control, Automation and Robotics (ICCAR), Apr. 2019, doi: 10.1109/iccar.2019.8813317.

[8]. V. Dimitrov, M. Wills, and T. Padır, "Realization of vision-based navigation and object recognition algorithms for the sample return challenge," 2015 IEEE Aerospace Conference, Mar. 2015, doi: 10.1109/aero.2015.7119271.

[9]. M. Dunbabin, B. Lang, and B. Wood, "Vision-based docking using an autonomous surface vehicle," 2008 IEEE International Conference on Robotics and Automation, May 2008, doi: 10.1109/robot.2008.4543182.

[10]. Haldorai, B. L. R, S. Murugan, and M. Balakrishnan, "Harnessing Intelligent AI to Elevate Business Modeling: A Perspective," EAI/Springer Innovations in Communication and Computing, pp. 429–440, 2024, doi: 10.1007/978-3-031-53972-5_22.

[11]. Haldorai, B. L. R, S. Murugan, and M. Balakrishnan, "Nature Inspired Optimizers and Their Importance for AI: An Inclusive Analysis," EAI/Springer Innovations in Communication and Computing, pp. 407–427, 2024, doi: 10.1007/978-3-031-53972-5_21.

[12]. N. Melenbrink, J. Werfel, and A. Menges, "On-site autonomous construction robots: Towards unsupervised building," Automation in Construction, vol. 119, p. 103312, Nov. 2020, doi: 10.1016/j.autcon.2020.103312.

[13]. P. M. Vaz, R. Ferreira, V. Grossmann, and M. Ribeiro, "Docking of a mobile platform based on infrared sensors," ISIE '97 Proceeding of the IEEE International Symposium on Industrial Electronics, Nov. 2002, doi: 10.1109/isie.1997.649089.

[14]. S. Roh et al., "Flexible docking mechanism using combination of magnetic force with error-compensation capability," 2008 IEEE International Conference on Automation Science and Engineering, Aug. 2008, doi: 10.1109/coase.2008.4626554.

[15]. L. Olatomiwa, S. Mekhilef, M. I. Ismail, and M. Moghavvemi, "Energy management strategies in hybrid renewable energy systems: A review," Renewable & Sustainable Energy Reviews, vol. 62, pp. 821–835, Sep. 2016, doi: 10.1016/j.rser.2016.05.040.

[16]. U. Kartoun et al., "Vision-Based Autonomous Robot Self-Docking and Recharging," 2006 World Automation Congress, Jul. 2006, doi: 10.1109/wac.2006.375987.

[17]. S.-J. Wang et al., "Effects and mechanisms of Xiaochaihu Tang against liver fibrosis: An integration of network pharmacology, molecular docking and experimental validation," Journal of Ethnopharmacology, vol. 303, p. 116053, Mar. 2023, doi: 10.1016/j.jep.2022.116053.

[18]. F. Guangrui and W. Geng, "Vision-based autonomous docking and re-charging system for mobile robot in warehouse environment," 2017 2nd International Conference on Robotics and Automation Engineering (ICRAE), Dec. 2017, doi: 10.1109/icrae.2017.8291357.

[19]. S. Liu, M. Özay, T. Okatani, H. Xu, K. Sun, and L. Yang, "Detection and pose estimation for Short-Range Vision-Based underwater docking," IEEE Access, vol. 7, pp. 2720–2749, Jan. 2019, doi: 10.1109/access.2018.2885537.

[20]. W. Liu et al., "SSD: Single Shot MultiBox Detector," in Lecture Notes in Computer Science, 2016, pp. 21–37. doi: 10.1007/978-3-319-46448-0_2.

[21]. X. Long, W. Cui, and Z. Zheng, "Safety Helmet Wearing Detection Based On Deep Learning," 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), Mar. 2019, doi: 10.1109/itnec.2019.8729039.

[22]. C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," arXiv (Cornell University), Jul. 2022, doi: 10.48550/arxiv.2207.02696.

[23]. I. Gallo, A. U. Rehman, R. H. Dehkordi, N. Landro, R. La Grassa, and M. Boschetti, "Deep Object Detection of Crop Weeds: Performance of YOLOv7 on a Real Case Dataset from UAV Images," Remote Sensing, vol. 15, no. 2, p. 539, Jan. 2023, doi: 10.3390/rs15020539.

[24]. M. S. Seyfioğlu, B. Erol, S. Z. Gürbüz, and M. G. Amin, "Diversified rADaR micro-Doppler simulations as training data for deep residual neural networks," 2018 IEEE RADaR Conference (RADaRConf18), Apr. 2018, doi: 10.1109/rADaR.2018.8378629.