

Article

Improving Robotic Bin-Picking Performances through Human–Robot Collaboration

Giovanni Boschetti  ^{1,2,*}, Teresa Sinico  ³ and Alberto Trevisani  ³¹ Department of Industrial Engineering (DII), University of Padova, 35131 Padova, Italy² Department of Information Engineering (DEI), University of Padova, 35131 Padova, Italy³ Department of Management and Engineering (DTG), University of Padova, 36100 Vicenza, Italy; teresa.sinico@phd.unipd.it (T.S.); alberto.trevisani@unipd.it (A.T.)

* Correspondence: giovanni.boschetti@unipd.it; Tel.: +39-049-827-6820

Abstract: The automation of bin-picking processes has been a research topic for almost two decades. General-purpose equipment, however, still does not show adequate success rates to find application in most industrial tasks. Human–robot collaboration in bin-picking tasks can increase the success rate by exploiting human perception and handling skills and the robot ability to perform repetitive tasks. The aim of this paper, starting from a general-purpose industrial bin picking equipment comprising a 3D-structured light vision system and a collaborative robot, consists in enhancing its performance and possible applications through human–robot collaboration. To achieve successful and fluent human–robot collaboration, the robotic workcell must meet some hardware and software requirements that are defined below. The proposed strategy is tested in some sample tests: the results of the experimental tests show that collaborative functions can be particularly useful to overcome typical bin picking failures and to improve the fault tolerance of the system, increasing its flexibility and reducing downtimes.

Keywords: bin picking; collaborative robotics; human–robot interaction; Industry 4.0/5.0



Citation: Boschetti, G.; Sinico, T.; Trevisani, A. Improving Robotic Bin-Picking Performances through Human–Robot Collaboration. *Appl. Sci.* **2023**, *13*, 5429. <https://doi.org/10.3390/app13095429>

Academic Editors: Lorenzo Scalera, Andrea Giusti and Renato Vidoni

Received: 30 March 2023

Revised: 24 April 2023

Accepted: 24 April 2023

Published: 26 April 2023



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

1. Introduction

Since the introduction of robots to industrial processes, researchers have been increasingly interested in fully automating as many tasks as possible. One of them is the task of grasping objects randomly placed in a bin for subsequent manipulation or positioning in a different place and pose, and it is usually referred to as bin-picking. This task is challenging because it involves 3D-object recognition, grasping strategies, and path planning. In many industrial processes, such as assembly operations or kitting, most of the parts come in boxes or bins, where the parts are randomly located and potentially interlocked. As a regular practice, parts are unloaded either manually or automatically using part feeders, such as long conveyor belts. However, manual bin-picking has several drawbacks, such as the possibility of causing health problems due to the weight of parts and the typically limited task execution throughput [1]. Conversely, the use of part feeders is often expensive, is not flexible, and requires large spaces. For these reasons, many research groups have addressed the problem of enabling robots to perform bin-picking tasks, guided by machine vision and other sensors [2–4]. Full automation of bin-picking tasks is challenging because it involves 3D object recognition and pose estimation, grasp planning, path planning, and collision avoidance. Each of these problems represents a vast area of research in itself and is usually treated separately. Despite technical developments in all these areas, robotic arms still cannot surpass human hands in terms of speed, adaptability, and dexterity [5]. As a consequence, applications of robotic bin-picking are often difficult to implement and lack standardization. However, some general-purpose equipment is available on the market, which can be used just in basic bin-picking scenarios, especially for bins containing a single type of part with a relatively simple shape. On the contrary, handling mixed bins that

contain multiple different types of parts with complex or unknown shapes remains an open and challenging task.

Because humans and robots share complementary strengths in performing tasks, a collaborative bin-picking workcell where robots and humans work in close proximity can enhance bin-picking success rates, increase system flexibility and productivity while improving the operators' working conditions by exploiting the strengths of both humans and robots. Indeed, robots can repeatedly perform pick-and-place operations without fatigue, and humans excel in their perception and handling ability in unstructured environments. Human–robot collaboration is made possible by collaborative robots (cobots), which are designed to work uncaged and interact with operators in the same working environment. Despite current interest and advancements in collaborative robotics, few studies have addressed the collaboration between robots and human operators in bin-picking tasks. In [6–8], a remotely located human assists the robot in critical situations by solving any automated perception problems encountered during bin-picking. In [9], a collaborative robotic cell for bin-picking was proposed, where the interaction between the human and robot is limited to a laser sensor that determines how far the operator is from the robot, and the controller adjusts the robot's velocity accordingly. In [10], another collaborative bin-picking cell was proposed, where the robot can hand over the picked item to the operator or take it back. However, in previous studies, robot collaborative functions such as manual hand guidance or collision detection have not been exploited to solve typical bin-picking failures.

This paper proposes a new framework for bin-picking, where a human operator working alongside the robot can help quickly and easily resolve faults when they inevitably occur. Starting from a general-purpose industrial bin-picking device composed of a 3D-structured light vision system and a collaborative robot, we show how its performance can be improved and its possible applications enhanced through human–robot collaboration. In particular, fault tolerance can be improved using collaborative functions such as robot manual guidance and the robot's ability to detect contacts. The hardware and software requirements that are needed to achieve successful and fluent human–robot collaboration when performing bin-picking tasks are defined, and the proposed strategy is tested in significant sample tests. Human–robot collaboration is proved to be particularly useful in resolving bin-picking faults and therefore increasing system fault tolerance.

The rest of this paper is organized as follows: in Section 2, the open technological issues are presented, and the complexity of fully automating bin-picking tasks as well as typical bin-picking failures are discussed. This is followed by Section 3, where improvement of system fault tolerance through human–robot collaboration is presented, as well as possible collaborative bin-picking applications. Section 4 outlines the hardware and software requirements for an effective collaborative bin-picking cell. In Section 5, the proposed strategy is tested in some illustrative experimental tests. Finally, in Section 6, the results are presented and discussed, and in Section 7, the conclusions are drawn.

2. Open Technological Issues

2.1. Vision Sensors and Computer Vision Algorithms

Three-dimensional visual data acquisition is the first challenge of the bin-picking problem chain. For this purpose, various technologies can be used: the most common are 3D laser scanners [11], structured light vision systems, consisting of a projector and one or two cameras [12] and stereo vision systems [13]. In bin-picking applications, the choice of the applied sensor for object registration is strongly correlated with failures, as discussed in [14]. When designing a robotic bin-picking workcell, a choice has to be made not only on the sensor's technology, but also on the vision sensor placement: it can be attached directly to the robot end effector or placed on a separate fixed camera stand.

Once the 3D point cloud map has been acquired, sophisticated computer vision algorithms are required for object detection and pose estimation, taking into account factors such as sensor noise, changing lightning conditions, shadows, and reflectance properties.

Most research on computer vision algorithms for bin-picking tasks has focused on Deep Learning Algorithms, in particular Convolutional Neural Networks and Deep Reinforcement Learning [15]. These approaches generally require 3D models of objects during testing or physical objects during training. Furthermore, they are very often supervised networks, and labeling is an indispensable stage for training the network and teaching where to grab an object based on its 3D model or its geometry. Therefore, these techniques do not scale easily for applications that frequently present objects that have never been seen before [16].

In the last decade, several manufacturers, such as Fanuc, Keyence, Cognex or Mech-Mind, developed their own ready-to-use 3D vision sensors with integrated computer vision algorithms for bin-picking. These systems are usually easy to install and set up, but it is not possible to customize the computer vision algorithms or access directly to the acquired 3D map.

2.2. End-Effector Design

The design of the end effector also plays a crucial role in bin-picking success rates, as the objects that need to be grasped from the bin are non-orientated, potentially interlocked, jumbled, and heavily occluding each other.

Gripper designs in bin-picking applications range from two-finger to multi-finger grippers, from suction or magnetic grippers [17] to soft grippers [18]. In the field of bin-picking, suction typically has an advantage over parallel-jaw or multi-finger grasping due to its ability to reach narrow spaces and pick up objects with a single point of contact [19]. However, the choice of the most suitable end effector for bin-picking tasks depends on a number of factors.

First, the optimal end effector design depends on the geometric shape of the parts that need to be picked from the bin. For example, in a crowded bin filled with adjacent cuboid objects, there may be no gaps to insert fingers for picking. In this situation, the objects can only be grasped from above, and a suction gripper is more effective than a two-fingered gripper. Additionally, when the bin is filled with items with various shapes, textures, and materials, a single end effector may not be able to effectively grasp all of them. Tool chargers, commonly used in bin-picking systems, can allow robots to switch grippers to attempt different types of grasp on the object [20]. In [9], a multi-gripper strategy has been implemented in order to enable more robust ways of grasping objects using suction and multiple fingertips. Additionally, in [16], the possibility of using a dual-arm cobot with a different end effector in each arm has been explored. Moreover, the most appropriate end effector also depends on the state of the bin, which changes while robots pick items from it: a multi-gripper switching strategy based on object sparseness was proposed in [21].

Finally, end-effectors used in bin-picking applications may also include force sensors to detect collisions with items and perform force control to avoid damaging items, as well as to ensure that the objects have been correctly grasped.

2.3. Bin-Picking Failures

Because robotic bin-picking involves picking of overlapping complex objects with an undefined position and subsequent grasping and placing, several types of failures can occur. Bin-picking failures can be classified as follows:

- Object recognition failures (perception failures): this type of failure occurs when the implemented computer vision algorithm is not able to recognize one or more items inside the bin. This happens frequently with reflective, shiny, or transparent objects or also with changes in lighting conditions. It is strongly correlated to the chosen vision sensor technology and computer vision algorithm, as well as with the resolution of the 3D map acquired;
- Pose estimation failures: this kind of failure happens when an object inside the bin is successfully recognized, but its calculated 3D pose is incorrect, and therefore the subsequent grasp fails. This may be due to vision sensor calibration errors or, more often, to an inaccurate pose estimation algorithm;

- Failures because of constraints on robot motion: this type of failure happens when it is not possible to reach the calculated object pose without collisions of the end effector with the bin edges and bin corners or with other objects inside the bin. These failures are strongly correlated to the shape and height of the bin, along with the shape of the end effector. For example, suction grippers are usually more compact and can reach the edges of the bin more easily than parallel grippers. In [22], an eccentrically mounted gripper and chamfered bin edges have been exploited in order to decrease the rate of this type of failure;
- Grasping failures: even though object recognition and pose estimation are successful, grasping can still fail for a number of reasons such as a porous or uneven surface, insufficient friction, grasp interference, entangled objects or multiple grasping of neighboring objects. A very detailed investigation of grasping failures in bin-picking applications with different end effectors can be found in [20].

Bin-picking task execution failures typically require the assembly/manufacturing line to be paused if the fault is unrecoverable and human intervention to enter the robotic workcell (thus, entering safety fences), clear the fault (for example, pick the entangled objects and either separate or discard them), and restart the line. Failures negatively influence the average cycle time, which often has a large impact on the value of bin-picking.

3. Collaborative Human–Robot Bin-Picking

Bin-picking is a challenging task to fully automate, and failure rates are still high for practical industrial applications. Rather than focusing on decreasing the rate of the typical bin-picking failures described in Section 2.3, we focused on improving the fault tolerance of an already existing workcell through human–robot collaboration, combining the best features of both operators and robots.

The diagram representing human–robot collaboration when resolving bin-picking failures is depicted in Figure 1. The green box represents successful bin-picking task execution, which consists of cyclically repeating 3D map acquisition with a 3D vision sensor, object recognition, pose estimation, path planning, and subsequent pick and place of the objects until the bin is empty. The use of a collaborative robot ensures that this sequence of operations can be performed by the robot while a human operator works by its side sharing the same working environment. Safe and fluent human–robot interaction can also be enhanced by implementing collision avoidance algorithms.

Regular bin-picking task execution can be interrupted either due to one of the failures described in Section 2.3 or due to autonomous human intervention. In a non-collaborative bin-picking scenario, when a fault occurs, the line needs to be paused, and a human operator needs to enter safety cages, clear the fault, for example, by removing interlocked objects, and restart the line. Instead, in a collaborative bin-picking scenario, when one of the possible faults occurs (orange box in Figure 1), an operator is already working side by side with the cobot and can interrupt his or her work and quickly clear the fault (light blue box in Figure 1).

The use of a collaborative robot also enriches possible applications, as human intervention is not limited to the separation of entangled objects or to their being discarded. For example, when a perception failure occurs, the cobot can seek help from the human operator, who can intuitively hand guide the robot to the pick position. The cobot then proceeds to pick the part and restore bin-picking task execution.

Moreover, bin-picking task execution can also be interrupted by human intervention. The human operator working alongside the cobot can autonomously predict and prevent bin-picking failures: if the operator notices that two objects are entangled, he or she can intervene and disentangle them. In addition, if the operator notices that an object cannot be picked because of constraints in robot motion, he or she manually removes it from the bin and places it where it needs to be placed. Finally, the operator can deposit more items inside the bin when it is almost empty: in this case, the cobot must behave accordingly, for example, by moving away from the bin.

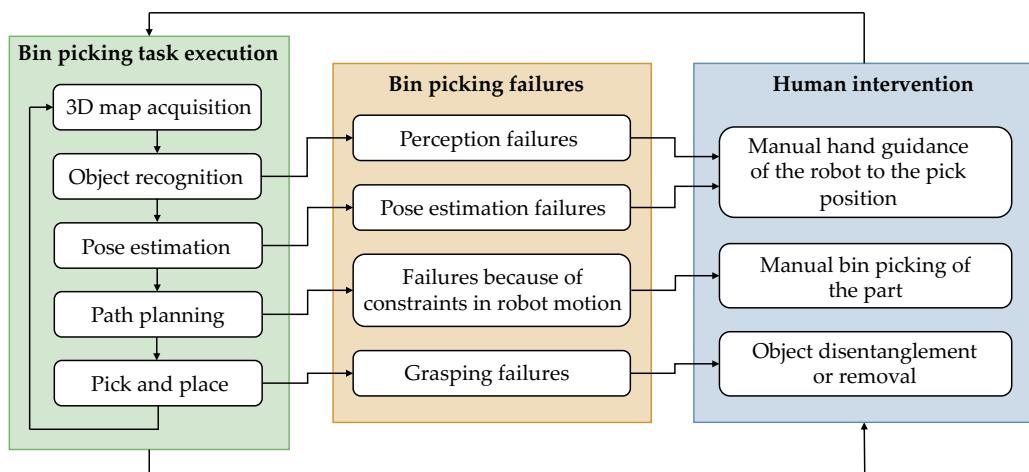


Figure 1. Human–robot collaboration when resolving bin-picking failures.

Possible Collaborative Bin-Picking Applications

Implementing a collaborative human–robot bin-picking cell allows a wide variety of possible applications, such as:

- When perception failures occur (i.e., the computer vision algorithm recognizes that the bin is not empty, but object recognition fails), the human operator can manually move the robot above the grasping pose by pushing it directly, exploiting zero gravity torque control [23]. Once the robot has been moved to the pick position, the operator notifies the cobot, for example, by touching it, exploiting collaborative (i.e., the robot ability to detect contacts) features, or by pushing a button. The robot then proceeds to approach and pick the part. The bin-picking execution task is restored and the fault is quickly and easily cleared;
- When bin-picking fails because of pose-estimation failures or because of constraints in robot motion, the cobot seeks external help from the operator. The human operator can manually perform bin-picking and subsequent placement of the part;
- When the computer vision algorithm recognizes that two or more objects are entangled with each other, the robot asks for external help, and the human operator can manually pick the entangled objects, separate them, and either put them back in the bin or place them where they need to be placed;
- The operator working alongside the robot can also autonomously predict possible bin-picking failures. For example, the operator can notice that two or more objects are entangled and disentangle them, without waiting for the cobot to ask for help. In addition, the operator can notice that an object near the edges of the bin cannot be picked and pick it manually. In these situations, the cobot must interrupt the execution of normal bin-picking tasks and behave accordingly: it might move away from the bin or decrease its speed;
- Human–robot interaction in collaborative bin-picking is not limited to fault handling and resolution. Since bin-picking is often the preliminary task in kitting or assembly applications, the robot can perform bin-picking and subsequently place heavier objects with a relatively simple shape, while the operator can perform bin-picking and subsequently place the objects with complex shapes that can easily become entangled and are difficult to separate;
- The workspace can be divided into two or more zones with bins or heaps of objects that need to be picked and placed, and the operator and the robot can work together, each on a different bin/heap of objects. To give an example, the operator might switch from an area to the area where a cobot is already working: in this case, the cobot moves away from that area and starts working in another one;
- The operator can also manually deposit more items inside the bin or replace an empty bin with a full one: in this scenario, the system recognizes the presence of the operator

and the robot can move away from the bin and resume its work once the operator has finished the operation;

- When an human operator enters the workcell, the velocity of the robot can be adapted to its relative distance from the operator. In particular, the robot's speed can be reduced proportionally to its distance from the operator. The robot path can also be modified to prevent collisions by exploiting collision avoidance algorithms.

In particular, the great advantage of collaborative bin-picking applications is that operators can work alongside robots and intervene only when necessary, increasing system flexibility, productivity, and fault tolerance. Moreover, collaborative bin-picking allows handling a wide assortment of parts, without the need for any change in the hardware structure and design.

4. Requirements of a Collaborative Bin-Picking Robotic Cell

In order to effectively perform collaborative bin-picking tasks, the robotic workcell must satisfy some hardware and software requirements defined in this section.

4.1. Hardware Requirements

- The human operator and the robot must be able to work in close proximity without safety fences: therefore, a collaborative robot must be used, since it is designed to interact with operators in the same working environment [24]. Safety is guaranteed by torque measurements in each joint that allow the implementation of collaborative features, i.e., impedance control algorithms [25]. Many modern cobots also offer the possibility to manually hand guide the robot end effector to a certain position [26], which can be particularly useful for some collaborative bin-picking applications;
- A 3D vision sensor is needed for 3D visual data acquisition of the bin [27]. The vision sensor technology must be chosen in accordance with the objects that need to be picked and its 3D point cloud resolution must be sufficient to detect the objects inside the bin;
- The presence of a human worker inside the workcell must be detected, and his/her position must be tracked, for example, in order to avoid collisions. A vision sensor can be used for this purpose. This could be a 2D- or 3D-camera and could even be the same one used for object recognition, as experimented in Section 5 of this paper. It is worth noting that using a collaborative robot already ensures the operator's safety, since if collisions occur, they are not dangerous. However, a collision usually results in an undesirable fault state of the robot and for this reason should be avoided in advance;
- Robot end effectors must ensure the operator safety: collaborative or soft robotic end effectors must be used [18];
- The robot and the operator must be able to communicate: in particular, the robot must be able to ask for help when it detects a potential failure. This might be carried out, for example, by turning on one or more LEDs or by exploiting a human-machine interface (HMI). The operator must also be able to notice the robot when the fault has been cleared and bin-picking task execution can be restored. For this purpose, the operator can use some buttons on the robot arm or on the HMI panel. However, to inform the robot that the fault has been cleared, the operator can also simply touch the robot, since cobots are able to detect contacts.

4.2. Software Requirements

- Object recognition and pose estimation [28] must have a fairly high success rate, thus justifying the automation of manual bin-picking;
- The computer vision system needs to be able to predict if a potential failure may occur. For example, it needs to recognize whether an object cannot be grasped because it is entangled with a neighboring object, whether no objects are recognized but the bin is not empty, or whether a certain pose cannot be reached without collisions. In these situations, the robot will ask for external help;

- The system must perform collision avoidance [29]: in particular, collisions with the bin edges and corners must be predicted and, if possible, path planning must be modified accordingly. In addition, collisions with the operators must also be avoided [30];
- For some collaborative bin-picking applications, it might also be necessary for the computer vision system to recognize the position of the bin and, in particular, to recognize if the bin has been moved from its previous position (this can happen when removing an empty bin and replacing it with a new one and the parameters for collision avoidance must be updated);
- The presence of an operator inside the workcell must be detected promptly to ensure his/her safety.

5. Improving a Bin-Picking Workcell through Human–Robot Collaboration

The proposed strategy was tested in a collaborative robotic workcell that meets the aforementioned requirements consisting of a general-purpose industrial bin-picking device. The bin-picking process taken into account consists of picking M30 hexagon nuts from a $30 \times 40 \times 15$ mm European standard plastic bin. When failures occur, a human operator intervenes to quickly and effectively resolve them.

5.1. Collaborative Bin-Picking Robotic Workcell

The collaborative bin-picking robotic workcell exploited during the experimental tests is depicted in Figure 2: it is composed of a robot arm and a 3D vision sensor. The robot arm is a 6-axis Fanuc CR-15iA collaborative robot, which has a payload of 15 kg, a reach of 1441 mm and its controller is the R-30iB Plus Fanuc controller. This cobot has an increasingly common feature: it can either work in collaborative mode (TCP speed under 250 mm/s and impedance control active) or standard mode (maximum TCP speed 800 mm/s and robot inside safety fences). During the experimental tests, the robot was always used in a collaborative mode. Collaborative robots are sometimes equipped with torque sensors in every joint which, together with an accurate dynamic model of the robot, enable detecting contacts and estimating contact forces along the entire robot structure. The tested robot is equipped with a single force sensor incorporated into its base that enables collaborative functions, i.e., can detect and estimate contacts along the structure. The robot arm also presents two LEDs (white and red) and a button.

On the robot flange, a force sensor and a collaborative electric gripper are fitted. The force sensor is an FS-15iA Fanuc force sensor: it detects collisions with items and avoids damaging objects when picking or placing them. The gripper is a Schunk Co-act EGP-C electric 2-finger parallel gripper certified for collaborative operation with an integrated LED strip light (which can assume three different colors: yellow, green, and red).

The 3D vision sensor is a Fanuc 3DA structured light sensor, and it is placed on a fixed camera stand. As can be seen in Figure 2, the vision sensor is composed of a projector unit and two camera units. The 3D Area Sensor obtains 3D information in the field of view by using the two camera units to capture multiple stripe pattern images as projected by the projector unit.

Fanuc provides its own computer vision system (iRVision) as an option that can be installed and integrated directly into the robot controller, which performs image acquisition and processing. The use of iRVision system enables the user to plug in a camera directly without any additional third-party hardware and software for image processing. This minimizes the time and number of activities performed during the implementation phase and eliminates the need to configure a communication between the robot controller and an external vision sensor. The iRVision system also provides its own computer vision algorithms (which Fanuc calls tools) for part recognition and localization, eliminating the need to develop complex image processing algorithms. However, this can also be a limitation as users cannot implement personalized computer vision algorithms. The robot controller is connected to a PC through an Ethernet network: on the PC, the setup of vision processes can be carried out by exploiting one or more predefined iRVision tools.

Alternatively, the configuration of vision processes can also be performed directly using the robot teaching pendant.

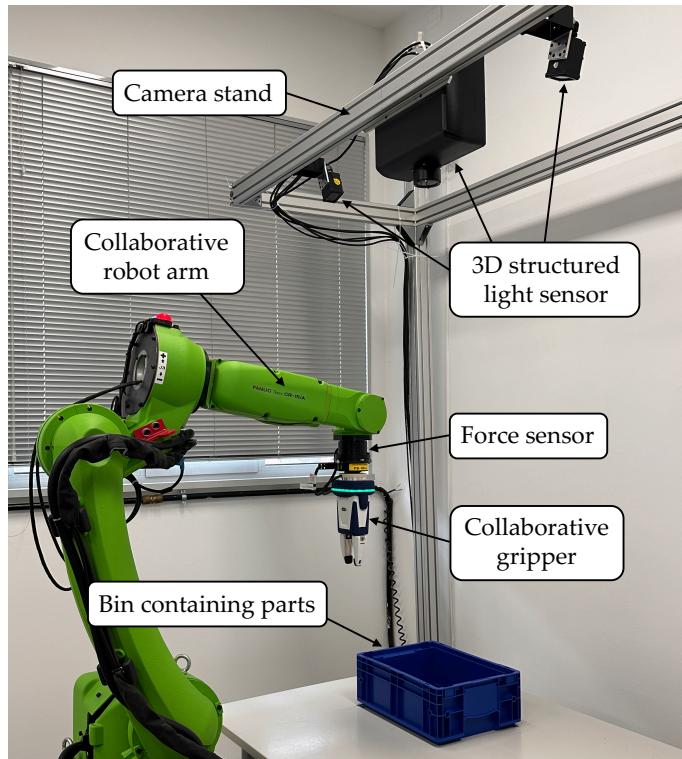


Figure 2. Collaborative bin-picking robotic workcell used during the tests.

5.2. Object and Human Detection

5.2.1. Object Detection

The objects picked from the bin during the experimental tests are M30 hexagon nuts. Hexagon nuts were chosen as objects to pick because they can be difficult objects to detect because of their reflectance properties.

Object recognition is performed by exploiting *3D One Sight Model Locator Tool*, an iRVision tool that detects a pre-trained 3D model in the acquired 3D data and calculates the 3D position and posture. The 3D model is created from 3D data obtained by measuring a workpiece from a certain point of view or can be a 3D CAD model. In this case, a 3D CAD model of the hexagonal nut was used to teach the tool which shape to find in the acquired 3D map. Figure 3 represents a successful object recognition of hexagon nuts exploiting the *3D One Sight Model Locator Tool*.

Once recognized, the detected objects are also numbered in the order in which they will be picked up from the bin. The order depends on the criteria chosen when setting up the vision process: in general, the first objects to be picked are the ones with the highest z quote. For example, the red nut in Figure 3 is labeled with number 1 and it will be picked before the green one underneath, labeled with number 6. Moreover, the nut under the orange and yellow one (labeled with numbers 2 and 3 respectively) is not recognized: this is not a problem because after each pick, the 3D map is acquired again, and therefore that nut will be recognized in the following scans, when the overlapping parts will be removed.

Due to the reflectance properties of the nuts and sensor technology, object recognition failures may occur. In particular, they may occur with the change in lightening conditions.

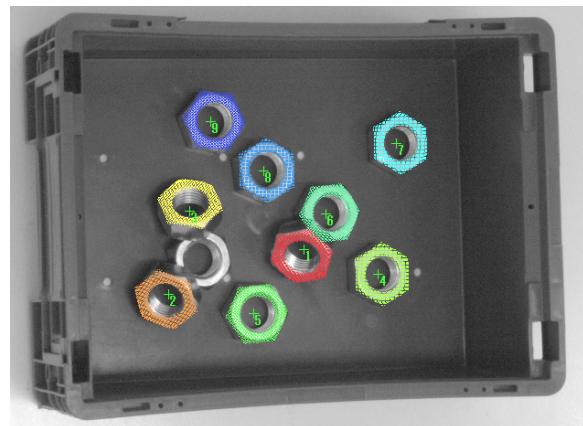


Figure 3. Three-dimensional object recognition of hexagon nuts through exploiting *3D One Sight Model Locator Tool*.

5.2.2. Human Detection

Because we were interested in improving bin-picking performance through human–robot collaboration using only off-the-shelf components, the presence of humans is detected using the same vision sensor used for 3D data acquisition.

As mentioned above, the Fanuc 3D sensor is composed of a projector and two camera units. One of the cameras was also used for human detection, exploiting the pre-defined 2D iRVision tools. However, the use of Fanuc iRVision tools does not easily allow the detection of complex shapes, such as human hands. Moreover, the two camera units are monochrome cameras and, therefore, it is not possible to detect a specific color in the captured image. Taking into account these limitations, we assumed that the human operator working alongside the cobot wears a special bracelet with geometric shapes that are easy to detect (rhombuses), as can be seen from Figure 4. Geometric shapes are placed all around the bracelet, making it easy to detect the presence of the operator with the arm in different positions and orientations.

The geometric target is recognized using the *Geometric Pattern Matching (GPM) Locator Tool*, which detects and locates a previously trained image pattern. This tool is based on 2D vision, and to make recognition even faster, the resolution of captured images was reduced. Detecting the presence of a human operator can be useful, for example, to slow down the robot if the operator enters the workcell and to avoid collisions.



Figure 4. Human detection through exploiting *3D One Sight Model Locator Tool*.

5.2.3. Synchronization between the Two Vision Processes

The two vision processes (2D for human detection, 3D for object recognition), as well as robot motion instructions, must be synchronized with each other. Figure 5 illustrates

a simplified flowchart the collaborative bin-picking program implemented, which was written exploiting the robot's teaching pendant. The main program (purple box in Figure 5) is a multitasking program that calls two subprograms: one that manages the 3D vision process and motion instructions (green box in Figure 5) and the other that manages the 2D vision process (red box in Figure 5).

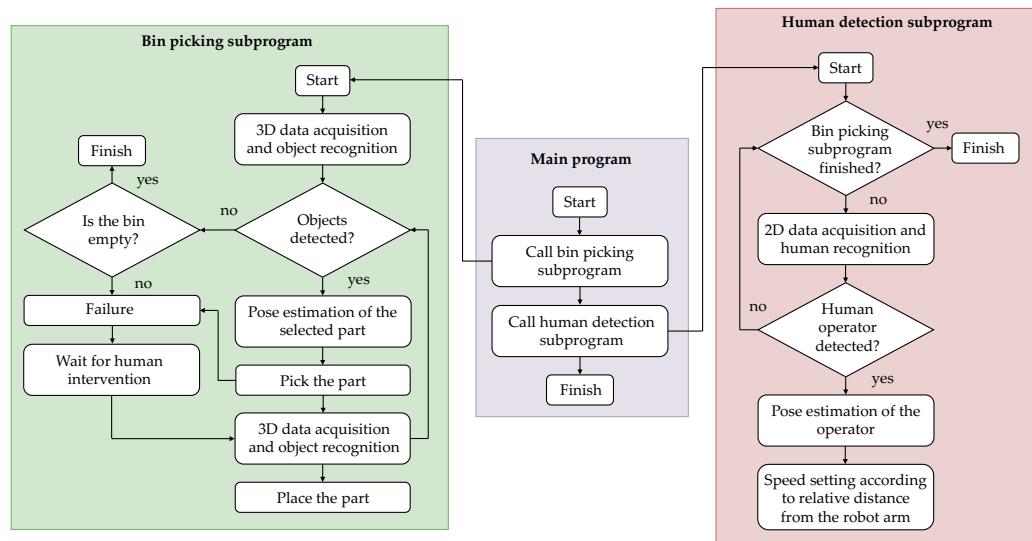


Figure 5. Flowchart of the collaborative bin-picking program implemented.

5.3. Collision Avoidance

The tested bin-picking system performs collision avoidance, both between the robot end effector and between the robot end effector and the human operator.

Collisions between the end effector and the faces of the bin, as well as with the table, are avoided thanks to the pre-defined *Fanuc Interference Avoidance* function. This function checks interference between the robot end effector and fixed objects. It also automatically generates the target position and posture in a specified range if interference occurs in the checked robot position. To use this function, the positions and sizes of objects for which interference must be checked should be set in advance. For the considered bin-picking process, *Fanuc Interference Avoidance* function was set as in Figure 6. As can be seen in Figure 6, the robot is enclosed in bounding volumes as sphere swept lines (SLL) to compute online proximity checks between the robot and the fixed objects, as well as between the robot and its human coworker. The robot tool is also enclosed in an hexahedron-shaped object, whose dimensions are set up by the user. In the considered bin-picking robotic workcell, robot and its end effector must not collide with the table and with the bin: both are represented by a hexahedron-shaped fixed object. In particular, thanks to this integrated function, the robot is able to avoid collisions with the container faces. If the system predicts a collision but the part can still be picked by tilting the end effector, the robot autonomously tilts the end effector when picking that item. If it is not possible to pick the detected part without colliding with the container, the system notifies the user (e.g., by writing a defined value in a register).

Collision avoidance between the robot arm and the human operator is performed by first locating the operator's wrist and then by calculating its distance from the robot arm. Of course, the bracelet is placed on the operator's wrist, but the potential collisions occur between the operator's fingers and/or arm and the robot end effector. This must be taken into account when implementing collision avoidance algorithms. Moreover, in this calculation, the delay introduced by the time needed to detect the operator must also be taken into account.

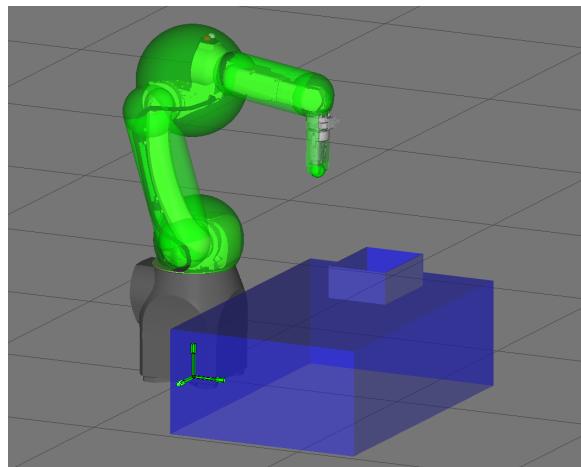


Figure 6. Setup of collision avoidance through *Fanuc Interference Avoidance*.

5.4. End Effector Design

The gripper mounted on the robot flange is a collaborative parallel electric gripper, and the fingers have been specifically designed for this application. In particular, the custom gripper fingers are designed to pick hexagonal nuts from the inside and have been 3D-printed.

5.5. Human–Robot Collaboration in Bin-Picking Failures

Even if the described bin-picking process has a high success rate, bin-picking failures can still occur. When failures occur, human intervention is exploited to clear the fault.

5.5.1. Failures Because of Constraints in Robot Motion

Figure 7 represents a typical case of failure due to constraints in robot motion. The circled nut in Figure 7a is correctly recognized, but because it is close to the edge of the bin, it cannot be picked up, as can be seen in Figure 7b. In fact, picking it would mean colliding with the bin faces, and tilting the robot end effector cannot solve the problem. In this case, after 3D data acquisition and object recognition, the system is able to recognize that the part cannot be picked and the robot does not try to move above the grasping point and pick the part. Instead, it seeks the help of the human operator by turning the gripper LED strip light red. Once the operator notices that the robot is looking for help, he or she manually picks the part from the bin (Figure 7c) and places it where it needs to be placed. When this is done, the operator informs the robot that the fault has been cleared. This is done by simply touching the robot: the robot recognizes the contact and restores bin-picking task execution. Alternatively, the operator can push the button placed on the robot arm, but this typically requires a greater amount of time.

5.5.2. Human–Robot Collaboration during Object Detection Failures

Due to the reflectance properties of hexagon nuts, with changes in lighting conditions, perception failures may also occur. An example of perception failure is depicted in Figure 8a: the circled nut is not recognized, although it must be picked before the green nut underneath (which, instead, has been correctly recognized). In Figure 8a also, a magnification of the acquired 3D point map is reported (3D points are represented in light blue): it can be seen that the map has too few points, and therefore the computer vision algorithms cannot recognize the part. In this situation, the robot seeks help by turning on the red light installed on the robot arm. The human operator then proceeds to manually move the robot above the grasping pose, exploiting the manual guided teaching function. Once this is done, the operator notices the robot that the fault has been cleared. This is also done by simply touching the robot.

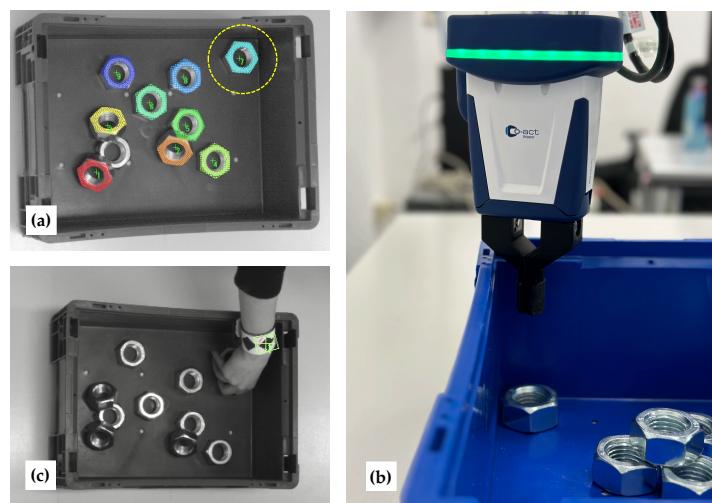


Figure 7. Example of a failure because of constraints in robot motion: (a) results of object recognition (the highlighted nut is recognized but cannot be picked), (b) constraints in robot motion (bin faces), and (c) resolution of the fault by manual pick of the part.

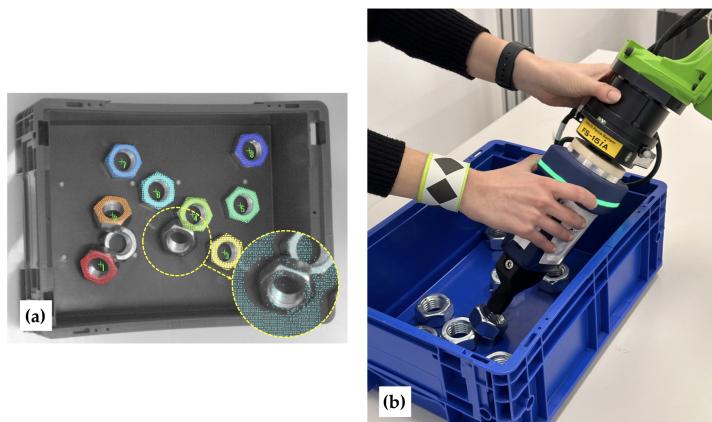


Figure 8. Example of a perception failure: (a) results of object recognition and magnification of the 3D point map acquired (it has too few points due to reflectance properties of the surface). (b) Fault resolution by manual hand guidance to the pick pose.

6. Results and Discussion

The experimental test revealed that robot collaborative functions can be effectively exploited when performing bin-picking tasks, and they allow a wide variety of possible applications. For example, manual hand guidance of the robot end effector can be exploited when clearing object-recognition faults, and the ability of a cobot to detect contacts can be used to inform the robot that the fault has been cleared by simply touching it. In addition, the experimental tests proved that it is, in fact, possible to use the same vision system for object detection and human recognition in collaborative human–robot bin-picking tasks. This is done by implementing a multitasking program that manages the two vision processes: the one for object recognition is a 3D vision process, while the one for human recognition is a 2D vision process. Three-dimensional map acquisition takes a significant amount of time, around 2 s, while object recognition takes 650 ± 50 ms. Two-dimensional image acquisition and human detection takes 914 ± 6 ms. It may seem a high value compared to the time required for 3D object recognition, but this depends on the searching area: reduced to the bin in 3D object recognition, extended to the whole field of view of the camera in human recognition. However, considering that the maximum TCP speed when

working in collaborative mode is 250 mm/s, the time required to detect the operator was short enough to ensure safety during the tests, as no collisions occurred.

The results obtained suggest that human–robot collaboration when performing bin-picking tasks increases flexibility and improves fault tolerance by combining human perception skills and robots' ability to perform repetitive tasks and lift heavy objects. In [6–8], human perception skills were already exploited to solve object recognition failures in bin-picking. In this case, the robot initiates a call to a remotely located human operator to ask for help in resolving perception system failures during bin-picking operations. This requires a user interface that allows information exchanges between the robot and the human operator that detects the objects that need to be picked from the bin. In the proposed layout, the human operator is already working alongside the robot, and he or she can more quickly resolve the fault by manually moving the robot end effector above the grasping position without the need for an interface. Manual hand guidance of the robot to the pick position can be particularly useful when resolving perception failures of heavy objects that are difficult to pick manually. Moreover, in this work, human perception skills are exploited not only for fault resolution, but also for fault prediction. In fact, the operator can predict when a potential failure may occur and autonomously disentangle entangled objects or manually pick a part that cannot be picked due to constraints in robot motion.

Tests have shown that it is useful, as suggested in [9], that the robot's speed is adjusted according to the relative distance between the robot end effector and the operator. However, in [9], a laser sensor was used, while in this work the bracelet allows a more accurate and almost costless location of the operator.

The experimental tests carried out showed that it is possible to effectively use the same robot-integrated vision sensor for both object and human recognition, proving the inexpensiveness of the proposed strategy. This has the advantage of minimizing the hardware required, as well as the connections to the robot controller. It is worth noting that the 3D scan usually takes a couple of seconds and during this time the human presence cannot be detected. This is clearly a limitation, but when a structured light vision sensor acquires a 3D map it is quite visible, so the operator can easily recognize the event and avoid intervening. Additionally, the use of a robot-integrated vision sensor eliminates the need to implement complex computer vision algorithms. Indeed, the use of robot-integrated computer vision algorithms poses some limitations: for example, since it is not possible to implement complex customized computer vision algorithms, we were forced to use a bracelet with geometric shapes to detect and locate human operators rather than using a hand/arm recognition algorithm. This is obviously a limitation, but it does not introduce risks. The operator may forget to wear the bracelet, but safety is still ensured by using a collaborative robot. Clearly, the presence of the operator cannot be detected, and its/her position cannot be tracked. Extensions of the proposed work may consider the use of an external 2D or 3D camera for human recognition. This solution, though more expensive, allows the implementation of a wide range of more complex computer vision algorithms that can detect the hand and possibly the whole body without the need for any additional accessories or markers on the operator. Using an external PC for image processing may also speed up recognition time, increasing safety. On the other hand, using an external camera for human detection requires the setup of a communication between the camera and the robot controller, for example, through TCP/IP or EtherCAT connection.

7. Conclusions

Bin-picking is a complex task and is very difficult to fully automate. The success rate of state-of-the-art bin-picking solutions is still not high enough, and it is not very effective to implement them in industrial applications. Starting from a general-purpose industrial bin-picking device composed of a 3D-structured light vision system and a collaborative robot, we showed how human–robot collaboration can improve success rates, increase system flexibility and productivity, as well as reduce downtimes. The hardware and software requirements necessary to implement a collaborative bin-picking cell were also defined.

The proposed strategy was then tested, showing that using the same vision sensor for object recognition and human recognition can be cheap and effective, and that collaborative functions can be successfully exploited to solve typical bin-picking failures.

Future works will include a comprehensive analysis of collaborative human–robot bin-picking performances compared to traditional or robotic bin-picking performances in terms of failures, success rate, downtimes, and productivity using some representative assembly tasks.

Author Contributions: Conceptualization, G.B., T.S. and A.T.; methodology, G.B., T.S. and A.T.; software, T.S.; writing—original draft preparation, T.S.; writing—review and editing, G.B., T.S. and A.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research received funding through the program “SISTEMA” (MUR-Dipartimenti di Eccellenza), and is coherent with the objectives of the Innovation Ecosystem “iNEST” funded by NextGenerationEU (ECS_00000043-SP.5).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

- Martinez, C.; Chen, H.; Boca, R. Automated 3D vision guided bin picking process for randomly located industrial parts. In Proceedings of the IEEE International Conference on Industrial Technology, Seville, Spain, 17–19 March 2015; pp. 3172–3177.
- Rahardja, K.; Kosaka, A. Vision-based bin-picking: Recognition and localization of multiple complex objects using simple visual cues. In Proceedings of the IEEE International Conference on Intelligent Robots and Systems, Osaka, Japan, 8 November 1996; Volume 3, pp. 1448–1457.
- Liu, M.; Tuzel, O.; Veeraraghavan, A.; Taguchi, Y.; Marks, T.K.; Chellappa, R. Fast object localization and pose estimation in heavy clutter for robotic bin picking. *Int. J. Robot. Res.* **2012**, *31*, 951–973. [[CrossRef](#)]
- Schraft, R.D.; Ledermann, T. Intelligent picking of chaotically stored objects. *Assem. Autom.* **2003**, *23*, 38–42. [[CrossRef](#)]
- Bencak, P.; Hercog, D.; Lerher, T. Evaluating robot bin-picking performance based on Box and Blocks Test. *IFAC-PapersOnLine* **2022**, *55*, 502–507. [[CrossRef](#)]
- Kaipa, K.N.; Thevendria-Karthic, S.S.; Shriyam, S.; Kabir, A.M.; Langsfeld, J.D.; Gupta, S.K. Resolving automated perception system failures in bin-picking tasks using assistance from remote human operators. In Proceedings of the IEEE International Conference on Automation Science and Engineering, Gothenburg, Sweden, 24–28 August 2015; pp. 1453–1458.
- Kaipa, K.N.; Kankanhalli-Nagendra, A.S.; Kumbla, N.B.; Shriyam, S.; Thevendria-Karthic, S.S.; Marvel, J.A.; Gupta, S.K. Enhancing robotic unstructured bin-picking performance by enabling remote human interventions in challenging perception scenarios. In Proceedings of the IEEE International Conference on Automation Science and Engineering, Fort Worth, TX, USA, 21–25 August 2016; pp. 639–645.
- Kaipa, K.N.; Kankanhalli-Nagendra, A.S.; Kumbla, N.B.; Shriyam, S.; Thevendria-Karthic, S.S.; Marvel, J.A.; Gupta, S.K. Addressing perception uncertainty induced failure modes in robotic bin-picking. *Robot. Comput.-Integr. Manuf.* **2016**, *42*, 17–38. [[CrossRef](#)]
- Olesen, A.S.; Gergaly, B.B.; Ryberg, E.A.; Thomsen, M.R.; Chrysostomou, D. A collaborative robot cell for random bin-picking based on deep learning policies and a multi-gripper switching strategy. *Procedia Manuf.* **2020**, *51*, 3–10. [[CrossRef](#)]
- Fuchs, S.; Haddadin, S.; Keller, M.; Parusel, S.; Kolb, A.; Suppa, M. Cooperative bin-picking with time-of-flight camera and impedance controlled DLR lightweight robot III. In Proceedings of the IEEE/RSJ 2010 International Conference on Intelligent Robots and Systems, Taipei, Taiwan, 18–22 October 2010; pp. 4862–4867.
- Roy, M.; Boby, R.A.; Chaudhary, S.; Chaudhury, S.; Roy, S.D.; Saha, S.K. Pose estimation of texture-less cylindrical objects in bin picking using sensor fusion. In Proceedings of the IEEE International Conference on Intelligent Robots and Systems, Daejeon, Republic of Korea, 9–14 October 2016; pp. 2279–2284.
- Oh, J.; Baek, K.; Kim, D.; Lee, S. Development of structured light based bin picking system using primitive models. In Proceedings of the 2009 IEEE International Symposium on Assembly and Manufacturing, ISAM 2009, Seoul, Republic of Korea, 17–20 November 2009; pp. 46–52.
- Radhakrishnamurthy, H.C.; Murugesapandian, P.; Ramachandran, N.; Yaacob, S. Stereo vision system for a bin picking adept robot. *Malays. J. Comput. Sci.* **2007**, *20*, 91–98. [[CrossRef](#)]
- Dolezel, P.; Pidanic, J.; Zalabsky, T.; Dvorak, M. Bin picking success rate depending on sensor sensitivity. In Proceedings of the 2019 20th International Carpathian Control Conference, ICCC 2019, Krakow-Wieliczka, Poland, 26–29 May 2019.

15. Alonso, M.; Izaguirre, A.; Graña, M. Current Research Trends in Robot Grasping and Bin Picking. *Adv. Intell. Syst. Comput.* **2019**, *771*, 367–376.
16. D’Avella, S.; Tripicchio, P.; Avizzano, C.A. A study on picking objects in cluttered environments: Exploiting depth features for a custom low-cost universal jamming gripper. *Robot. Comput.-Integr. Manuf.* **2020**, *63*, 101888. [[CrossRef](#)]
17. Roy, D. Development of novel magnetic grippers for use in unstructured robotic workspace. *Robot. Comput.-Integr. Manuf.* **2015**, *35*, 16–41. [[CrossRef](#)]
18. Shintake, J.; Cacucciolo, V.; Floreano, D.; Shea, H. Soft Robotic Grippers. *Adv. Mater.* **2018**, *30*, 1707035. [[CrossRef](#)] [[PubMed](#)]
19. Mahler, J.; Matl, M.; Liu, X.; Li, A.; Gealy, D.; Goldberg, K. Dex-Net 3.0: Computing Robust Vacuum Suction Grasp Targets in Point Clouds Using a New Analytic Model and Deep Learning. In Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA), Brisbane, QLD, Australia, 21–25 May 2018; pp. 5620–5627.
20. Sanders, K.; Danielczuk, M.; Mahler, J.; Tanwani, A.; Goldberg, K. Non-Markov Policies to Reduce Sequential Failures in Robot Bin Picking. In Proceedings of the IEEE International Conference on Automation Science and Engineering, Hong Kong, China, 20–21 August 2020; pp. 1141–1148.
21. Fujita, M.; Domae, Y.; Kawanishi, R.; Ricardez, G.A.G.; Kato, K.; Shiratsuchi, K.; Haraguchi, R.; Araki, R.; Fujiyoshi, H.; Akizuki, S.; et al. Bin-picking robot using a multi-gripper switching strategy based on object sparseness. In Proceedings of the IEEE International Conference on Automation Science and Engineering, Vancouver, BC, Canada, 22–26 August 2019; pp. 1540–1547.
22. Pochylý, A.; Kubela, T.; Singule, V.; Cihák, P. 3D vision systems for industrial bin-picking applications. In Proceedings of the Proceedings of 15th International Conference MECHATRONIKA, Prague, Czech Republic, 5–7 December 2012.
23. Albu-Schäffer, A.; Ott, C.; Hirzinger, G. A unified passivity-based control framework for position, torque and impedance control of flexible joint robots. *Int. J. Robot. Res.* **2007**, *26*, 23–39. [[CrossRef](#)]
24. Matheson, E.; Minto, R.; Zampieri, E.G.G.; Faccio, M.; Rosati, G. Human-robot collaboration in manufacturing applications: A review. *Robotics* **2019**, *8*, 100. [[CrossRef](#)]
25. Ikeura, R.; Inooka, H. Variable impedance control of a robot for cooperation with a human. In Proceedings of the IEEE International Conference on Robotics and Automation, Nagoya, Japan, 21–27 May 1995; Volume 3, pp. 3097–3102.
26. Safeea, M.; Bearee, R.; Neto, P. End-Effector Precise Hand-Guiding for Collaborative Robots. In ROBOT 2017: Third Iberian Robotics Conference, Proceedings of the ROBOT 2017, Sevilla, Spain, 22–24 November 2017; Advances in Intelligent Systems and Computing; Springer: Cham, Switzerland, 2018; Volume 694, pp. 595–605.
27. Mezquita, Y.; González-Briones, A.; Wolf, P.; Prieto, J. Computer Vision: A Review on 3D Object Recognition. In *Distributed Computing and Artificial Intelligence, Special Sessions, 19th International Conference, DCAI 2022, Proceedings of the International Symposium on Distributed Computing and Artificial Intelligence, L’Aquila, Italy, 13–15 July 2022; Lecture Notes in Networks and Systems*; Springer: Cham, Switzerland, 2023; Volume 585, pp. 117–125.
28. Wohlhart, P.; Lepetit, V. Learning descriptors for object recognition and 3D pose estimation. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3109–3118.
29. Khatib, O. Real-time obstacle avoidance for manipulators and mobile robots. *Int. J. Robot. Res.* **1986**, *5*, 90–98. [[CrossRef](#)]
30. Mohammed, A.; Schmidt, B.; Wang, L. Active collision avoidance for human–robot collaboration driven by vision sensors. *Int. J. Comput. Integr. Manuf.* **2017**, *30*, 970–980. [[CrossRef](#)]

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