Progress in Autonomous Picking as Demonstrated by the Amazon Robotic Challenge

Dominik Colling  
Karlsruhe Institute of Technology

Jonathan Dziedzitz  
Karlsruhe Institute of Technology

Kai Furmans  
Karlsruhe Institute of Technology, kai.furmans@kit.edu

Patric Hopfgarten  
Karlsruhe Institute of Technology

Kai Markert  
Karlsruhe Institute of Technology

Follow this and additional works at: https://digitalcommons.georgiasouthern.edu/pmhr_2018

Part of the Industrial Engineering Commons, Operational Research Commons, and the Operations and Supply Chain Management Commons

Recommended Citation
https://digitalcommons.georgiasouthern.edu/pmhr_2018/29

This research paper is brought to you for free and open access by the Progress in Material Handling Research at Digital Commons@Georgia Southern.  
It has been accepted for inclusion in 15th IMHRC Proceedings (Savannah, Georgia. USA – 2018) by an authorized administrator of Digital Commons@Georgia Southern. For more information, please contact digitalcommons@georgiasouthern.edu.
Progress in Autonomous Picking
as demonstrated by the Amazon Robotics Challenge

Dominik Colling, Jonathan Dziedzitz, Kai Furmans, Patric Hopfgarten, Kai Markert
Karlsruhe Institute of Technology (KIT)
Institute for Material Handling and Logistics (IFL)
Karlsruhe, Germany
kai.furmans@kit.edu

Abstract—The automation of the picking process is a long-held dream in intra-logistics, while the expert judgment in recent years was, that there is no immediate applicability to be expected for applications in areas, where the variety of products is high. However, the recent advances in computer science, mechanical engineering and material handling systems design have moved the field substantially forward. In this article we want to show, how the Amazon Robotic Challenge has driven this development process and how capable the technology is. The Amazon robotic Challenge (ARC) and its increasingly difficult challenges are described first. Then the applied technology in soft- and hardware and its performance increase are described. The performance of important participants in the past is estimated and finally the approach of the team IFLpiro is presented. Finally, the impact of the results on the future implementation of material handling systems is discussed.

Index Terms—Amazon Robotics Challenge, automated order picking, autonomous picking

I. INTRODUCTION

In 2016, members of our institute started to study the Amazon Robotics Challenge by visiting the competition in Leipzig / Germany. In 2017 a small team participated and learned a lot. Even without a strong background in robotics, the team was able to compete with more experienced teams in Nagoya, which should be a message to all academics and practitioners in the field of order picking. This article is the English translation of a contribution to the "Materialflusskongress 2018" in Munich and the related publication in the proceedings of this conference [1]. In order to learn from the the most successful team read their publications about the grasping system and the image segmentation in [2] and [3].

II. MOTIVATION

Flexibility in terms of quantity and range of articles are two important challenges in the automation of picking processes. For homogeneous goods (e.g. pharmaceuticals in homogeneous packaging), automated order picking is already state of the art. With heterogeneous articles and a changing range of articles, however, there is still no robust and economical solution on the market. However, there would be a great need here in high-wage countries. Ever shorter product life cycles in connection with smaller quantities and an increasing variety of variants require a growing flexibility of the companies. At the same time, competition from low-wage countries is growing, while in Germany personnel costs are high and the shortage of skilled workers and demographic change are increasingly creating a bottleneck. However, there has been some recent progress in the development of automated yet flexible intralogistics systems, e.g. for internal material handling. AGVs can now be quickly put into operation and expanded. They are also no longer bound to any track guidance and can navigate freely in space. Standardized load carriers such as containers, boxes and pallets are already available on the market today so that they can be automatically removed and stored [4]. The next step in production and logistics is the handling of a wide variety of products in any shape and with any material properties. In the future, for example, even flexible parts that are difficult to recognize and grip due to their different contours will have to be handled automatically. If this is successful, commissioning can be fully automated. Significant progress has been made in this area in recent years. This publication exemplifies this progress with the Amazon Robotics Challenge in the years 2015 to 2017. To this end, the competition itself and its development will be presented. In the following, the technical progress in the areas of hardware and software is examined in more detail and its effects on the challenge are described. This is followed by a discussion of what these innovations mean for the order picking of the future, what consequences they will have for the logistics of the future in general and how they will affect society and the labour market. Finally, an outlook is given on the new challenges that can be derived from this.

III. AMAZON ROBOTICS CHALLENGE

Amazon and its subsidiary Amazon Robotics recognized the opportunities and possibilities of automated order picking early on and therefore organized the Amazon Picking Challenge (APC) for the first time in 2015 to accelerate development in automated order picking. The first competition took place during the ICRA (International Conference on Robotics and Automation) in Seattle, USA. The task was to pick 12 out of 24 articles from a shelf into a crate within 20 minutes with a self-sufficient robot (Pick Task). Points were awarded for every correctly completed task. If wrong articles were grabbed or articles were damaged, there was a point deduction. All articles were available to the teams in
advance as a training set for several weeks, including 2D and 3D article data. An analysis of this competition can be found in [5], the winning team described their approach in [6]. The team, which came in second published information about their system in [7] In 2016 the competition took place in Leipzig as part of the RoboCup. The number of articles on the shelf has been increased to 50, more than doubling. Each compartment contained one to ten articles. Within 15 minutes, 12 defined articles had to be put into a crate. On another day, 12 items had to be placed on a shelf filled with 40 items (Stow Task). The Nimbro team, which came in on second and third place, described their system in [8].

In 2015 and 2016, the rack with the articles was a standard rack, which Amazon uses for its manual picking, for instance movable racks which are brought by KIVA to the picking station. As a specific difficulty, in these racks items might be located behind each other. In 2017, the teams were allowed to design their own racks, with some limits on the size and the technology used. The length of the storage space was limited to 125 cm, the totals floor space to 0.5 square meters and the volume to 0.095 cubic-meters.

<table>
<thead>
<tr>
<th>TABLE I</th>
<th>OVERVIEW OF THE MAJOR FEATURES OF THE AMAZON ROBOTICS CHALLENGES BETWEEN 2016 AND 2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>task</td>
<td>no. locations</td>
</tr>
<tr>
<td>2015</td>
<td></td>
</tr>
<tr>
<td>Pick</td>
<td>12</td>
</tr>
<tr>
<td>Pick</td>
<td>12</td>
</tr>
<tr>
<td>2016</td>
<td></td>
</tr>
<tr>
<td>Pick</td>
<td>12</td>
</tr>
<tr>
<td>Stow</td>
<td>12</td>
</tr>
<tr>
<td>2017</td>
<td></td>
</tr>
<tr>
<td>Pick</td>
<td>2 - 10</td>
</tr>
<tr>
<td>stow</td>
<td>2 - 10</td>
</tr>
<tr>
<td>fin. stow</td>
<td>2 - 10</td>
</tr>
<tr>
<td>fin. pick</td>
<td>2 - 10</td>
</tr>
</tbody>
</table>

Under the name Amazon Robotics Challenge (ARC), the 2017 extended competition took place in Nagoya, Japan. The design and construction of the shelving system was now up to the participants and thus part of the task. The size, area and volume of the shelving system have been limited. It was allowed to contain sensory (with a value up to 50$) and mechanical components, but no actuators. The shelving system had to be transportable by FTS - you can see the reference to KIVA here. In contrast to previous years, the robots now also had to store and retrieve articles in the various competitive tasks with which they could not practice in advance. Only half of the articles were known, the other half of the articles were made available to the teams for 30 minutes shortly before the respective task. This made the use of neural networks for article identification considerably more difficult. During the two competition days with one pick and one stow task each, 8 of 16 teams qualified for the final round task. The robot first had to fill up the half-filled shelf further in order to then pick 10 articles into three different shipping cartons. The rules for the individual years can be found in [9], [10] and [11]. An overview of selected general conditions of all competitions is shown in I.

The level of difficulty and the scope of the tasks were increased with each challenge. Nevertheless, the tasks were solved better and by more teams. At the first APC 2015, the winning team RBO from Berlin relocated 10 articles, but with only three correctly grabbed articles a team still came in third place. Only half of the 26 teams managed to grab a correct article at all. In 2016 and 2017, the teams moved closer together with their performance. While in 2015 only two teams scored more than 50% of the winning team’s points, in Pick Task 2016 there were already four teams and in 2017 six teams. The situation is similar for the teams at the lower end. In 2015, for example, 14 out of 26 teams achieved zero points, in 2016 and 2017 there were only four and three out of 16 teams without points in the pick task ([5]).

IV. TECHNICAL ADVANCES OF THE COMPONENTS

A. Hardware

An important component for successful participation in the ARC is the imaging sensorics. Without a camera the grip of an article is almost impossible. Since a free position on the shelf must be determined for the Stow task, 3D sensor technology was generally used, which determines a depth image of the surroundings in addition to the RGB image. The release of the X-Box Kinect in 2010, originally conceived as a 3D sensor for gesture control of an X-Box console, brought about a change in the field of 3D sensor technology, but was quickly discovered by universities and research institutes and used for a variety of applications. The decisive factor here was the low price compared to existing 3D sensors. Since then, many companies have developed their own 3D sensors, e.g. Intel. Intel has been releasing a new version of its Intel RealSense 3D camera annually since 2015 [12]. With the current D400 [13] series, Intel combines two processes, infrared camera and stereo camera to produce more stable depth images (see figure 1).

Fig. 1. Intel RealSense Depth Camera D415 and D435 [13]

The data collected must then be processed. Image processing requires a lot of computing power for this, since the majority of the teams at least partially relied on artificial neural networks (ANN). Image processing with the help of ANN generally uses a large number of convolutions and thus a very large number of floating point operations in the calculation. Ordinary processors can no longer solve this task in real time. These calculations are now taken over by graphics cards. With
the release of Nvidia’s Pascal architecture in 2016 (see 2, the available performance increased significantly. For example, the teams using graphics cards at ARC 2017 were able to learn new articles in a short time by training neural networks and using them for image recognition. GPU servers with more than four graphics cards were used for this purpose. With the announcement of the next generation of Nvidia, called Volta (see Figure 2) for 2018, and announcements by AMD at the Financial Analyst Day 2017 on their 7nm format architecture navigation system, another rapid development is expected that will bring more performance in smaller volumes. Nvidia offers an example of integrated graphics cards in a small space with its embedded platform Jetson [14]. It is easy to integrate into a variety of applications.

![Fig. 2. Number of Single precision floating General Matrix Multiply (SGEMM) per Watt of different architectures of Nvidia [15]](image)

The robotic arms used in the competitions did not change much over the years. Some teams, like the MIT team, even used the same robots across multiple challenges. In 2017, the KUKA LBR iiwa [16] was used for the first time by Team Panasonic. This is a sensitive 7-axis robot with integrated impedance control. This opened up new fields of application. The main problem, especially for research institutes, is the high purchasing costs of a robot arm. Franka Emika GmbH [17] with its robot arm Panda could help in the future. Like the KUKA LBR iiwa, the Panda is a sensitive 7-axis robot at a low cost and has been in production since October 2017. For this they were awarded the German Future Prize 2017 [18].

### B. Software Architecture

In addition to the technical progress of the hardware, the open source concept with open source code and simple extensibility is responsible for the rapid development in robotics. Since 1999 the Open Source Computer Vision Library (short: OpenCV) is developed. The basic idea of the open source library is the simple handling of complex evaluation and filter algorithms in relation to machine vision. Meanwhile the library contains over 2500 optimized algorithms, which can be used via different interfaces in C++, Python, Java or Matlab under Linux, Windows and MAC OS [19]. The integration of NVIDIA CUDA [15] into OpenCV in 2010 created new possibilities for real-time calculation of image recognition algorithms.

While in 2010 mainly conventional image recognition methods such as Support Vector Machines (SVMs) in combination with various classifiers such as histogram-of-oriented Gradients (HOG) or Scale-Invariant-Feature-Transform (SIFT) features were used, these methods have now been replaced by Deep Learning methods for object recognition and classification. For the simple development of neural networks, different frameworks have been created, such as Theano [20], Torch [22] or TensorFlow [23].

The ImageNet Large Scale Visual Recognition Challenge (ILSVRC) provides a database of more than 14 million images divided into over 17,000 categories. Every year, researchers from all over the world can measure themselves against this with their image recognition algorithms. Figure 3 clearly shows that conventional detection algorithms have been completely overtaken by AI-algorithms since 2012. In 2015, for the first time, the human detection rate was surpassed by that of a computer program and in 2016 humans were clearly beaten. Similar to the comparison of conventional image recognition methods, in which image characteristics conceived and developed by human beings are extracted, the structure, i.e. the number of neurons and the number and type of layers, have so far been determined by human beings in the case of the neuronal networks.

![Fig. 3. Rating of the ILSVRC of the last six years [24]](image)

With the publication of AutoML by Google, the structure of the KNN can be created automatically and even adapted to certain boundary conditions, such as calculation time criteria. With this method, the automatically designed NASNet was able to beat all man-made networks in the ImageNet classification in 2017 (see Figure 4).

If one considers the progress of hardware in connection with many open source and powerful libraries, it can be seen that a uniform interface between hardware and software of robot systems is missing as the middle piece. With this basic idea in mind, research work such as the Stanford Artificial Intelligence Robot (STAIR) and Willow Garage’s first robotics project laid the foundation for an entire robotics framework. Already in 2009 the first version of the Robot Operating System (ROS) was released under the name ROS 0.4 mango tango. Until today the framework has been extended by a multitude of packages, message types and algorithms and allows the fast integration of complex software on different hardware [26]. ROS 2.0 is currently being developed to provide real-time
Fig. 4. Comparison of different neural networks in ImageNet classification [25]

capability for industrial applications and a revised message concept with a focus on IOT applications.

V. 4TH TECHNICAL PROGRESS OF THE APC/ARC FROM 2015 - 2017

The technical advances in hardware and software also made themselves felt in the competitions between 2015 and 2017. In order to master the tasks of the competitions, two main challenges have to be solved:

- Recognition and identification of articles.
- Handling of the articles.

The greatest progress was achieved in the recognition speed of the articles. This is mainly responsible for the picking time per article. The time per pick for the best teams in 2015 was still over 77 seconds and could be reduced to less than 29 seconds in 2017 (see Figure 5). Although cameras have also become more powerful in recent years, the greatest benefits in object recognition result from the advances in ANN and the increase in performance of the graphics cards on which these networks compute.

Fig. 5. Average pick time of the two best teams from 2015 to 2017

Regarding the handling of the articles there were some remarkable developments with the grippers. As with other components, additive manufacturing processes (3D printing) were increasingly used here. The teams mainly used suction pads, as the majority of the articles could be vacuumed due to their nature. Equipped with an additional finger or alternatively usable 2-finger grippers, the manageable article aspect could be increased. No significant further developments were observed in the robots. At ARC 2017, the shelving system itself had to be designed and built for the first time. Most teams opted for a system of boxes instead of standing, conventional shelving systems. This enabled gripping from above instead of from the side, which made gripping much easier.

VI. THE APPROACH OF TEAM IFL PiRo

At the Institute of Conveying Technology and Logistics Systems (IFL) of the Karlsruhe Institute of Technology (KIT), an autonomous order picking robot was already under development in one project. It was therefore a good idea for the institute to take part in the Amazon Robotics Challenge in order to gain experience and speed up the development process. At the end of 2016, the IFL PiRo team applied to participate, was selected and was allowed to compete with 15 other international teams in July 2017 in Nagoya, Japan.

Fig. 6. Team IFL PiRo at the ARC 2017 in Nagoya, Japan

The first fundamental decisions were on the storage system and the arrangement of the components. IFL PiRo has opted for a storage system consisting of crates open from above, as this allows identical access to the storage system and shipping cartons. It is also easier to store items rather than having to put them on a traditional shelf. The boxes and cartons were arranged around the robot arm. This made it possible to monitor the level above the crates using a laser scanner. For the gripper, IFL PiRo decided to combine a 2-finger parallel gripper with a suction cup. If necessary, the suction cup was gripped by the finger gripper. This made it possible to select the appropriate handling method for each article. A commercially available vacuum cleaner was used to generate the suction flow. The camera used was a DepthSense DS 325/525, which has a 2D RGB sensor and a 3D module based on the time-of-flight
principle (ToF). Several methods were used simultaneously for object recognition. Two convolutional neuronal networks (CNN) were used for the recognition of articles known from the training set. For the articles, which were only known shortly before the task, two feature-based algorithms were used (SIFT: Scale-Invariant Feature Transformation, or ASIFT: Affine SIFT). For the calculations an NVIDIA GeForce 1070 GTX (graphics card with Pascal architecture) was used. Team IFL PiRo reached the final round at the ARC 2017 with its robot and storage system and finished 7th overall. For more details, see [27].

![IFL PiRo at the Institute of Materials Handling and Logistics Systems (IFL)](image)

**VII. IMPORTANCE OF AUTOMATED ORDER PICKING IN THE FUTURE**

It can be assumed that the trend towards globalization will continue, even if it is hindered in the short term by political considerations. Globalization, which leads to globally distributed production, obviously results in an increasing number of distribution centers, which establish the logistical link between the manufacturers, who may be operating on another continent, and the end customers. At the same time, various studies show that the share of e-commerce in retail sales will continue to increase. For Germany, growth rates of e-commerce sales are estimated to decrease from 13.1% (2017) to 4.5% (2022). The penetration rate increases from 60.6% to 65.6% (all data [28]). Both phenomena together, the increasing share of trade due to globalization and the growing share of e-commerce are making efficient order picking an increasingly important task in the national economies (for the USA see [29]). The study "A Future that Works: Automation, Employment, and Productivity" [30] clearly shows that the automation of simple, physically repetitive tasks will increase and is also necessary to compensate for the demographically induced reduction of the working share of the total population. Picking of articles is certainly one of the simpler and, if necessary, well structured tasks in logistics. By concentrating this task in distribution centers, which are handled either by large trading platforms or by logistics service providers offering their services to various companies, an increasingly attractive basis for automation is developing:

- The total volume of work to be done in one place is increasing.
- The technology for automated order picking of heterogeneous item-specific items has improved significantly in recent years and is more cost-effective.
- The results of the last challenge show that automated order picking is feasible under suitable conditions. Due to the specifics of the Amazon approach, the competitive conditions for order picking are more difficult than they would be for a warehouse that is geared towards automation.

The winning teams’ pick time of 28.7s/pick is slow, especially since only one article was handled at a time. However, numerous measures are conceivable to make better use of the existing technology. The following boundary conditions could be used to further increase performance:

- Previous knowledge exists as to which article is in which subject.
- When assigning storage locations, clear identifiability could be taken into account.
- The learning of new articles by the neural networks can typically take place in advance, e.g. ideally when adding new articles to the assortment.
- Good accessibility of the articles for the picking process, if necessary by removing a container with articles to be picked from the picking station.
- Parallelization of processes: recognition, localization, gripping planning and own gripping. During the conveying process from the storage area to the picking station, image recognition and planning of the gripping strategy could already take place, so that the robot station would only be occupied for the time of the actual gripping process.

That is why we at IFL concluded that the fundamental questions of the state of the art are sufficiently clarified, so that industrialization can concentrate on solving the questions of robustness and profitability. This would make it possible to implement bearings that can be operated 24 hours a day, with remote support if necessary, and even under more unpleasant environmental conditions (temperature, lighting). We actually wanted to end this presentation with a preview of the next Amazon-Robotic Challenge, since we would also have liked to have participated in it and wanted to do better with the combined abilities of KIT. Now Amazon has changed the funding format and will no longer host another challenge. For us, this is a clear indication that Amazon also considers the maturity level of automated order picking technology to be advanced. The challenges now lie in making the process safe, robust and economical.

**REFERENCES**
