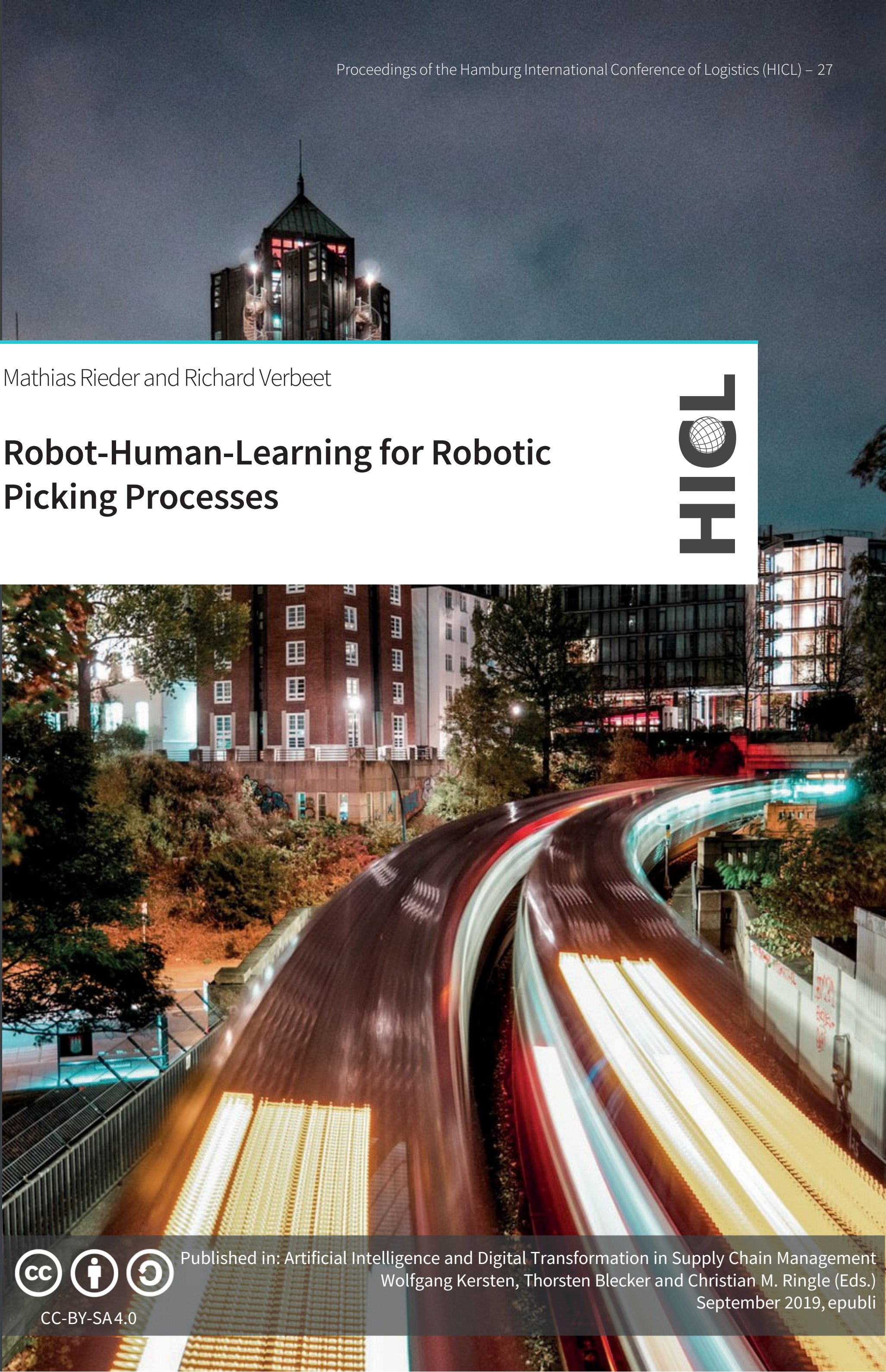
Proceedings of the Hamburg International Conference of Logistics (HICL) – 27



Mathias Rieder and Richard Verbeet

Robot-Human-Learning for Robotic Picking Processes

HICL

Published in: Artificial Intelligence and Digital Transformation in Supply Chain Management
Wolfgang Kersten, Thorsten Blecker and Christian M. Ringle (Eds.)
September 2019, epubli

Robot-Human-Learning for Robotic Picking Processes

Mathias Rieder[1] and Richard Verbeet[1]

1 – Ulm University of Applied Sciences

Purpose: This research paper aims to create an environment which enables robots to learn from humans by algorithms of Computer Vision and Machine Learning for object detection and gripping. The proposed concept transforms manual picking to highly automated picking performed by robots.

Methodology: After defining requirements for a robotic picking system, a process model is proposed. This model defines how to extend traditional manual picking and which human-robot-interfaces are necessary to enable learning from humans to improve the performance of robots' object detection and gripping.

Findings: The proposed concept needs a pool of images to train an initial setup of a convolutional neural network by the YOLO-Algorithm. Therefore, a station with two cameras and a flexible positioning system for image creation is presented by which the necessary number of images can be generated with little effort.

Originality: A digital representation of an object is created based on the generated images of this station. The original idea is a feedback loop including human workers after a not successful object detection or gripping which enables robots in service to extend their ability to recognize and pick objects.

Keywords: Picking Robots, Machine Learning, Object Detection, Computer Vision, Human-Robot-Collaboration

First received: 19.May.2019 **Revised:** 23.June.2019 **Accepted:** 26.June.2019

1 Introduction

Finding staff for carrying out logistic tasks is getting harder and harder for companies as a survey of Kohl and Pfretzschner (2017) showed. Combined with developments in engineering and Artificial Intelligence there is a trend to integrate machines into the execution of logistic tasks, either to support workers or to automate them completely (Schneider, et al. 2018). Different to tasks for transport or manufacturing standardization, it is more challenging in picking tasks because a high amount of flexibility is needed to complete these tasks. This is the main reason why there is a low level of automatization in picking processes of just 5% in warehouses, 15% are mechanized and 80% are still run manually (Bonkenburg, 2016). Fully automated picking processes, besides fully automated storage, offer several advantages: saving of space and labor cost, availability of personnel instead of robots, savings on operational cost as heating or lighting (de Koster, 2018) and facing lack of personnel in logistics.

For here discussed robots in logistics there is a suitable definition of Bonkenburg (2016) in contrast to all other robotic solutions like robotic vacuum cleaner or nursing robot: "A Robot with one or more grippers to pick up and move items within a logistics operation such as a warehouse, sorting center or last-mile”.

Picking of known objects in dynamic environments by robots is a major task as shape and position of an object may change since the last visit of a robot at the object's storage location. If the position of an object is constant, e.g. for welding robots in automotive production systems, robots complete their jobs very well. Therefore, no understanding of their surroundings is necessary. But if the robot must work in cooperation with humans there are changes to the environment as, to the objects, the shelf, and the position of the objects within the shelf or their orientation. Furthermore, even the object itself can be different since the last handling process due to changing object design caused by changed package sizes or modernization of styles as it is common business. In retail there is also a constantly changing product range by introducing respectively discontinuing promotional or seasonal goods. So, a robot must adapt to this situation by object detection.

The cooperation of robots and humans is necessary because the number of objects robots can pick is very small (Schwäke, et al., 2017). A promising approach is to assign those picking orders to robots they can recognize and grip while humans pick the leftovers (Wahl, 2016). These both sections could be separated in different areas, but this would cause two major disadvantages: humans are not able to pick objects from the robots' working area, e.g. in case of a capacity bottleneck during seasonal peaks, and robots can't enlarge the number of pickable objects by working with and learning from human colleagues.

In addition to this, cooperation between robots and humans may be the answer if partial automation is desired or even required because of lack of personnel. To enable such a picking setup a process model is proposed which allows cooperation between humans and robots to guarantee robust processes and learning robots. The first step in this model is to generate the necessary data for robots' object detections. But especially for jobs in logistic environments there is a lack of data sets for training object detection systems which are essential for robot picking (Thiel, Hinckeldeyn and Kreutzfeldt, 2018). Out of these data sets the object detection system pulls "knowledge" about the objects. If data quality is low the resulting model will also be inadequate and furthermore, if there are objects not being part of the input data, they cannot be identified by the model. This data set must initially be created which means a lot of work, for comparison COCO-dataset contains 330,000 images for differentiating 80 object categories which took about 70,000 worker hours (Lin, et al., 2015). Therefore, an adaptive system is necessary whose data can represent the latest status to successfully work on picking orders.

Besides the question on how to get data for training there is a very central point mentioned by Hui (2018): "The most important question is not which detector is the best. It may not possible to answer. The real question is which detector and what configurations give us the best balance of speed and accuracy that your application needed." The two central aspects of characterizing an object detection algorithm is accuracy and speed (Hui, 2018).

Considering the heterogeneous landscape of objects combined with the variety of cameras, algorithms, impacts from surroundings like lighting and robotic and computing hardware the comparison of existing solutions is a very challenging task. This results in a need for experiments and testing. There must be a specific set of training data for each solution approach which represent the target area of the algorithm. The best way for gathering such data representing operational processes is using these processes themselves. To support the task of gathering information, external data input is needed from humans to tell the system about changes on objects as a computer system cannot reliably recognize the consequences of chances to objects for object detection and picking. There is also a need for an efficient way to collect images to train an initial object detection model which must work successful on training data before it can be used in picking customer orders in operational processes.

This leads to the question how to transform manual picking processes into highly automated ones in an efficient way ensuring operational order fulfillment. To answer this question, it is necessary to set up an object detection system as basis for robot picking evaluating which object detection algorithm is suited best for a specific logistic environment. Comparing possible algorithms, a specific training and testing data set is necessary which is not existing yet. During building up this data set, there will be further questions to be answered belonging the data set itself, e. g. how many images of each object must be recorded, how to face changes to objects and their appearance or which angles and rotations of object images are more useful to training.

Summarizing the central questions in short:

1. How to transform manually operated picking processes into highly automated ones?
2. Which algorithm(s) support object detection in logistics best?
3. What must a data set for object detection in logistics look like?
4. How can changes on the objects and within object range be handled?

Deriving from these questions the goal of the current research work is to work out a process environment that makes adaption to changes possible using human-robot-cooperation. Besides there must be an answer on how to collect images of objects on an efficient way to make a comparison of object detection algorithms possible to choose the best suitable one.

This paper is structured as follows. Section 2 reviews related work in areas of picking robots and object detection. In section 3 the chosen object detection algorithm is introduced shortly. Section 4 proposes a process model that handles the need for image data to enable object detection models including the introduction of a picture recording machine. In section 5 first results are presented and shortly discussed. Section 6 presents the conclusion. Section 7 shows further questions for research work on picking robots and object detection.

2 Related Work

For the topic of this paper there are two sections of great interest: picking robots which will do the physical job by gripping objects and object detection algorithms to determine where the robots must grip the targeted objects.

2.1 Picking Robots

For several years many research efforts have been done on flexible picking robots, e. g. for harvesting vegetables and fruits like oranges (Muscato, et al., 2005), cucumber (Van Henten, et al., 2005) or strawberry (Hayashi, et al., 2010). Current robotic applications are driven by four technology trends that enable and enhance the applicable solutions of robots in logistics. These are feet (mobility), hands (collaboration and manipulation), eyes (perception and sensors) and brains (computing power and cloud) where each trend has shown many improvements in recent years (Bonkenburg, 2016).

Nowadays there are several companies offering mobile picking robots as IAM Robotics (2019), Fetch Robotics (Fetch Robotics, 5Inc., 2019) or Magazino (Magazino GmbH, 2019) and many more supporting logistics processes by partwise process automation (Britt, 2018). But there will be more solutions supported by developments within the technology trends mentioned before, which, for example, support robots for more and more flexible gripping by tactile sensors which make robotic grippers more adaptable to their use case (Costanzo, et al., 2019). Besides better sensors, grippers are getting increasingly adaptable as the presentation of a gripper construction kit for robots by Dick, Ulrich and Bruns (2018) shows. The central component, the brain of the robot, is also being refined. Besides the constant improvement on brain's processing and architecture there is a continuous work on algorithms to extract information from sensor input to detect object positions and the object's gripping point faster and with higher accuracy (Tai, et al., 2016).

Motion as the job of moving around is solved adequate for autonomous guided vehicles since many years and works on first picking robots successfully as Magazino's Toru shows (Wahl, 2016).

To sum up: existing systems partially face the problem of picking automation and may deliver viable solutions in stable environments but there is a lack of flexibility adapting to chances in the environment.

2.2 Object Detection

For picking processes it is essential to know where the target object is located. For this job object detection algorithms determine the position of the target from sensor data - usually images from cameras - within an image that contains multiple objects (Agarwal, 2018). Semantic Segmentation, classification, localization and instance segmentation are other jobs working on images besides object detection. These tasks of Computer Vision are shown in Figure 1 outlining the differences of these.

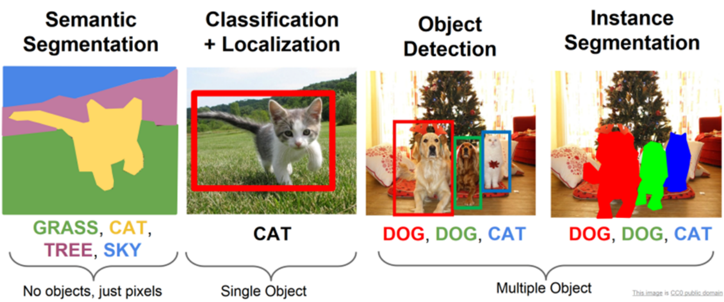


Figure 1: Comparison of semantic segmentation, classification and localization, object detection and instance segmentation (Li, Johnson and Yeung, 2017)

To improve the efficiency of object detection, Machine Learning can be used to extract information from an existing set of data to predict on unknown data (Witten, Frank and Hall, 2011). This can be applied to different domains where a great amount of data exists - the more data the better, which is the case for object detection (Domingos, 2012).

Research of recent years within the field of object detection has developed approaches based on Deep Learning, a special kind of Machine Learning.

They offer the advantage of finding features automatically, for example a neural network is taught using training data for object detection (Ouaknine, 2018a). Several algorithms for object detection were developed using Deep Learning, a comparison of these algorithms for different applications is presented by Zhao, et al. (2019). Deep Learning is used to train a neural network which later, after the training is finished, can be used for object detection tasks. But such neural networks have problems at object detection with object not being part of the training data as these tend to be identified as objects containing in the data set (Colling, et al., 2017). Machine Learning can also be used for gripping point detecting which outperforms hand-set configurations (Lenz, Lee and Saxena, 2015).

Object detection requires input data distinguished in 2D-images or 3D-information which depends on what information is available, which kind of objects must be distinguished or what accuracy is needed. If the objects look different but have identical geometric shape a combination of images and distance information may be needed, as RGB-D data, to define gripping points (Lenz, Lee and Saxena, 2015). Another option is the computation of 3D-information from 2D-images (Jabalameli, Ettehadi, and Behal, 2018).

Like other object detection algorithms YOLO is trained by images. To gather images there are different approaches like turning an object in front of a camera to take images. This idea is used for different purposes like 3D-Scanning as basis for 3D-Printing (Rother, 2017), 360-degree images for web shops (Waser, 2014) or master data and image capturing (Kraus, 2018). A similar setup to the proposed process model in this work was created by Hans and Paulus (2008). Their focus of research is on color within the images (Hans, Knopp, Paulus, 2009). Another specific setup is designed to record 360-degree-images of motor vehicles (Ruppert, 2006). A similar approach is moving the camera around an object and take images from different positions which is proposed by Zhang et al. (2016).

To compare the outcome of different algorithms meta-data is added to the data sets containing information about the set. Different data sets are used for different learning jobs like images, text or speech (Stanford and Iriondo, 2018). For an image data set this information defines the objects in the images and where within the picture the objects can be found. This enables a comparison of the output of object detection algorithms and what they should discover within the images. Latest research in object detection focuses on COCO-dataset (Common Objects in Context) which was presented by Lin, et al in 2015 including metrics measuring the performance of object detection algorithms on test images. Redmon et al. (2016) characterized the performance of their YOLO-algorithm (You Only Look Once) using different datasets (ImageNet 2012, VOC 2007, VOC 2012, Picasso, People-Art). For version 3 of YOLO there is given a comparison on COCO-dataset only (Redmon and Farhadi, 2018). YOLO is mentioned here as it is the "fastest general-purpose object detector in the literature" (Redmon, et al., 2016).

3 YOLO-Algorithm

Processing 45 images per second YOLO-algorithm can be part of a real-time object detection system (Redmon, et al., 2016). YOLO trains a convolutional neural network (CNN) with training data on a loss function (Redmon, et al., 2016). The functional principle of YOLO is shown in Figure 2 originating with the publication of Redmon et al. (2016). Within the CNN images are split into grids cells where each cell is analyzed for possible objects, marking them with a bounding box and equipping each with according confidence. Afterwards the bounding boxes of each grid cell are combined with a class probability. Each grid cell can only contain one object so the bounding box with the highest confidence is chosen. Neighboring cells respectively bounding boxes containing the same object are summarized by non-maximal suppression. In difference to other algorithms YOLO works on the whole image what makes it being so fast. But there are limitations which are detecting small objects appearing in groups or objects in new respectively unusual environments (Redmon, et al., 2016).

4 Proposal for Adaptive Generation of Learning Data

As mentioned above classifying models must deal with a changing range of products especially in businesses where product design is changed for promotional purpose. This is a major problem for a robotic picking system. It needs reliable input from the object detection system because failure in picking can be expensive causing delayed order completion or even non-fulfillment of customer orders.

This requirement leads to the need for picking processes being adaptable to changes in the environment.

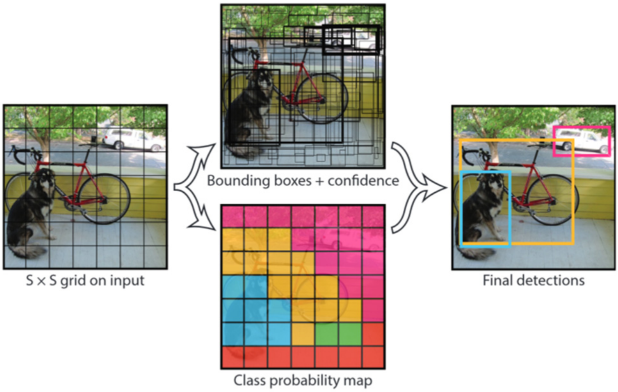


Figure 2: Working principle of YOLO-algorithm (Redmon, et al. 2016)

4.1 Process Modell

The process model consists of the two parts Learning and Operation as shown in Figure 3. Within the first part no robotic equipment is needed as it aims to build a detection model for objects which can be calculated on external computing resources. Therefore, images of the objects to differentiate must be generated. A lot of pictures are needed to calculate such a model so that images of an object from different perspectives and angles of rotation must be created. The different rotation angles and perspectives are needed as the object can appear in every orientation in a warehouse (compare Zhang, et al., 2016). For subsequent calculation of the detection model from the images they are stored in a database.

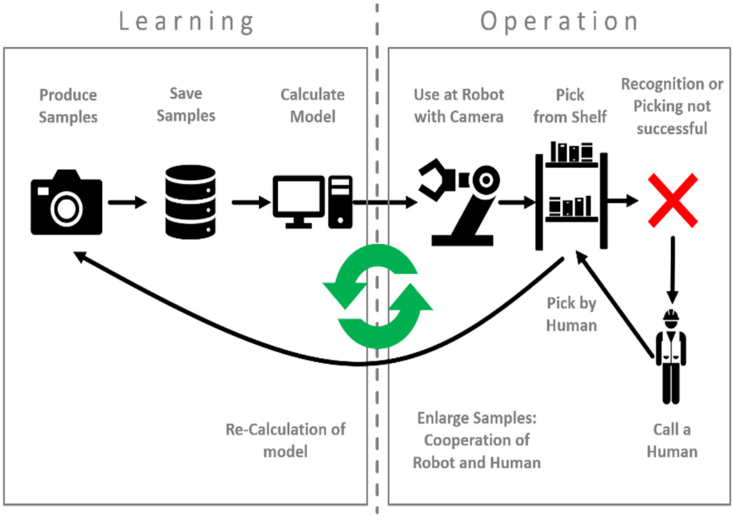


Figure 3: Process Model

If the object detection model exceeds the defined performance indicators it is used in a real picking environment which is describe by Operation in the process model. These performance indicators must be defined evaluated during testing. There the object detection model is applied to the robot control to find objects the robot must pick. An image of the target shelf is recorded by a camera mounted at the mobile robot. The model locates the target object within the images and defines grasping points from the orientation of the object and possible grasping points from the database where master data is saved. If the robot succeeds everything is fine.

If a problem occurs, e.g. the target object isn't detected in the shelf because of changes of its design or it is obscured by another object, the robots calls for a human picker. The human fulfills two important tasks.

If the object is in the shelf, he completes the order by picking the object. Furthermore, he must give feedback describing why detection was not possible according to the system's error message and, if the object is in the image, where it is located. The system uses this information to improve the detection model for the next try by including the additional images recorded in cooperation with the human picker at the shelf for model calculation. But as the calculation of such a model on a standard computer lasts several days re-calculation cannot be done in real-time on the robot as could be observed during testing.

If an object detection model performs very poor the object will be sent back to Learning: more images must be recorded with the Picture Recording Machine being introduced in the following chapter and the detection model must be trained once again.

4.2 Picturing Recording Machine

As first step to implement a setup of the introduced process model it is necessary to gather images of the different objects. As doing this manually is a time-consuming job this task is partial automated by a Picture Recording Machine which is shown in Figure 4.

Besides this the machines enables collecting images from precise orientations in a repeatable way. For gathering images on different locations, it is mobile and enables imaging of objects from different perspectives and angles of rotation. To get different angles of rotation between 0 to 360 degrees there is a turning table in the center of the machine which is driven by a stepper motor (42SHD0216-20B).

Figure 4: Picturing Recording Machine

Furthermore, the camera system mounted on the rocker is also moved by a separate stepper motor (Nema 23, 60STH86-3008B). By this, images between 0 and 90 degrees can be taken. Each motor is controlled by its own motor driver (PoStep60) triggered by a microcontroller (Raspberry Pi 3 Model B v1.2). The electronic setup is shown in Figure 5.

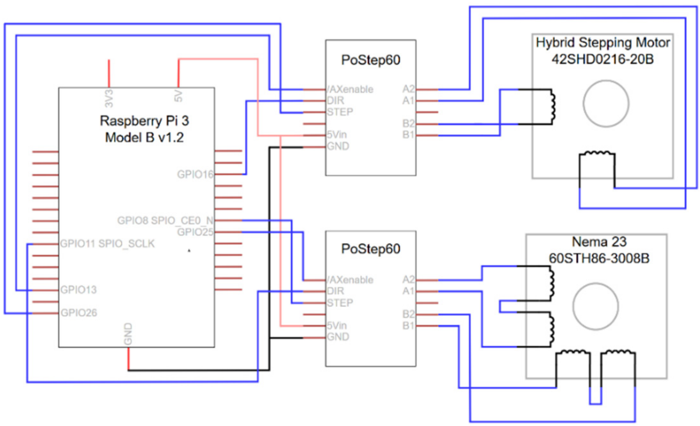


Figure 5: Wiring Diagram for the Picturing Recording Machine

This setup allows the collection of images in a semi-automatic way as there is only the change of the objects and the input of the number of images and angles to be done manually. The system gets its input by the GUI that is shown in Figure 6. This GUI is connected to the database where images and master data are stored. It is possible to use the Picture Recording Machine also for collection of master data. By clicking on the barcode-icon the object's barcode is scanned by the user and the system checks if there is existing master data and images according to this object-ID. If there is existing data, it is loaded from database if not, a new object-ID is generated. Loading data from the database also includes loading a representative view of the object in the GUI what makes comparison of existing data and the present object possible for the user of the machine.

There are options for recording images of the object as the number of images taken in one rotation and the number of perspectives. For perspectives there is the option to decide whether to include 0- or 90-degree's view or not.

The mounting on the rocker is designed to simultaneously support two different camera systems for imaging. Besides a comparison of camera-hardware it enables also comparison of different algorithms as there may be one recording 2D-data and the other one generating 3D-information.

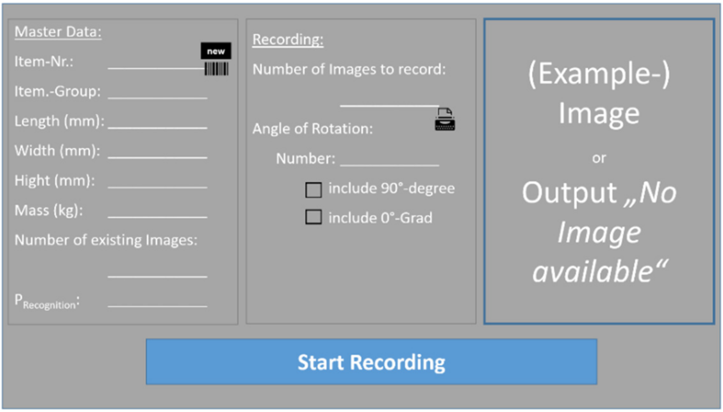


Figure 6: GUI for the Picturing Recording Machine

These images are taken by two different cameras: Microsoft Kinect One and Photoneo PhoXi 3D-Scanner M giving the option to compare training and testing of object detection algorithms on different equipment and testing the algorithm working independent on different hardware (Photoneo s. r. o., 2018; Microsoft Corporation, 2018). These ones are chosen as there will be a comparison of the industrial camera solution which comes with better features and a higher price (Photoneo) as the cheap home solution (Kinect). Both are each chosen in their category after a market research under the condition of each camera system providing color images (e.g. RGB) and depth-information.

Currently there is a Microsoft Kinect One and a Photoneo PhoXi 3D-Scanner M mounted on the rocker (Photoneo s. r. o., 2018; Microsoft Corporation, 2018).

As creation of own data sets for training is suggested by Thiel, Hinckeldeyn and Kreutzfeldt (2018) the Picture Recording Machine is an essential part of this research work and will support the follow-up steps to create an adaptive and learning environment for robots in logistics.

5 Results and Discussion

Evaluating YOLO-algorithm images of all-day office objects are taken which could be picked in a retail commissioning (coffee package, stapler, different bottles, different beverage packs, cookie pack). Figure 7 shows a first test on a manually taken image which shows the result localizing three objects within the image, giving a confidence score for each object: ice_tea (58%), water bottle (91%) and coffee (92%). The predic-tion needs 5.63 seconds which is about the time quoted by Redmon et al. and shows only predictions with confidence higher than 20% (Redmon, 2016). The complete results of the first try are shown in Table 1.

The Convolutional Neural Network for this purpose is trained with 329 training images containing one or two of the objects (254 images with one object and 75 with two). Preparation and recording these images take approximately 10 men-days. Training images are taken with Photoneo PhoXi

3D-Scanner M, testing images for evaluating object detection are shot with MS Kinect.

Even though "ice_tea" exists in 154 images (50 of them with another object) it had a bad score comparing "water bottle" and "coffee", except "stapler" which seems having not enough images or otherwise being only single in training images.

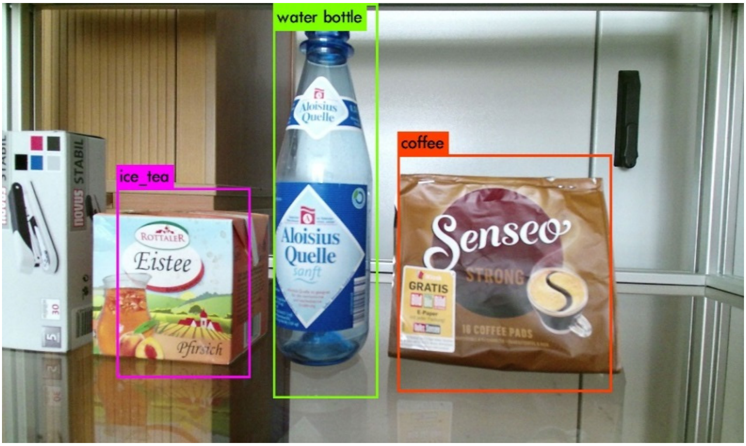


Figure 7: Testing YOLO on manually taken images

A very positive result of these test is the section of "false positives". A false positive result is an object the object detection algorithm tells it would be another object. As errors in picking are expensive this score must be low. There must be distinguished between known (part of training data) and unknown objects appearing within an image. Removing unknown objects from "false positives" gives the actual number of errors. The scores of 0% to 6% in this category resulting from only a few training images indicates YOLO to be a very promising approach.

The other way round "false negatives" tell about objects being in an image but not being detected. This is not a problem for the proposed process model as the human fallback level handles this type of error. For false negatives' score there is the chance of getting better by human feedback during operation.

Table 1: Results of testing YOLOv3 in first images

	ice_tea	**water bottle**	**coffee**	**stapler**
training images	154	100	100	50
single object	104	50	50	50
with one other object	50	50	50	0
test images	50	50	50	50
positive detections	31 (62%)	32 (64%)	46 (92%)	1 (2%)
false negatives	19 (38%)	18 (36%)	4 (8%)	49 (98%)
false positives	0 (0%)	1 (2%)	13 (26%)	1 (2%)
known objects	0 (0%)	1 (2%)	3 (6%)	1 (2%)
unknown objects	0 (0%)	0 (0%)	10 (20%)	0 (0%)
mean confidence for positive detections	69,1%	72,8%	82,7%	63,0%

To sum up the results YOLO-algorithm seems to be very promising for our purpose as there were only few images for training and testing leading to quite good results. So, there will be further research about this approach.

Supporting training for object detection the Picture Recording Machine will help by contributing efficient picture recording. First test had shown that the machine is able to record about 20 images a minute under high repeating accuracy and by systematically saving the images with according recording information (ID, angle, rotation, camera type). The number on images per minute depends on how many pictures are taken during one rotation and the number of the angles of recording. The bigger the number of images the shorter the distance to move between two recordings.

6 Conclusion

The proposed model will help to transform manual picking processes to a robotic picking system for a wide range of objects, whereby the problem of missing worker or legal restriction (working on Sundays) can be solved. For making picking in warehouses by robots possible continuous update of object data is necessary. Cooperation with humans is essential for reliable working on picking orders. In this work an approach for cooperation between robots and humans for picking processes is presented to guarantee stable output of commissioning and to support the robotic system by updating and extending their object data to increase the rate of objects being picked by robots.

The presented process model provides the following advantages which ensures robust processes as well as a learning environment for gripping robots:

1. Generation of missing image data respectively the needed number of images
2. Continuous updating of image data

3. Decoupling of generation of images, model testing and robotic installation
4. Reliable picking processes by human fallback level

Making robots learn to grip new objects with the presented process model may work slowly for a big number of objects but it guarantees stable output of the commissioning.

A part of the process model is the generation of image data of related objects which is a common problem for today's application working on object detection. Generating many images of many objects is possible by the Picture Recording Machine what tackles the problem of missing data sets. In combination with feedback from human-robot collaboration the basic data can be enriched by images from the process which makes object detection training results more stable.

Having many images also enables a comparison of different object detection algorithms which again enables to choose the best one for the specific object detection task.

7 Future Research

After gathering images of different objects there will be a closer look on which parameters affecting the output of different object detection algorithms. Therefore, tests with the recorded data set on known algorithms will be done. Another question to answer is which input has which impact on the output of an algorithm: number of images, how to handle similar looking objects, which degrees of rotation and which camera angles are more helpful than others supporting object detection. A further question to answer is if 0- and 90-degrees' view is how helpful for object detection models.

A further step towards more efficient learning could be the automated generation of coordinates of an area where the object within an image is located. Now this is done manually and very time consuming so there will be attempts for a higher degree of automation.

A very important part of the process model is the human-robot interface where information is generated that supports the learning process. There must be research on how this interface must be designed that human pickers will accept the co-working with their robotic colleagues. Besides the information the humans can give as feedback to the system must be specified that the learning system can understand what problems made detecting the object not possible.

The automation of image gathering could also enable meta-learning comparing different object detection algorithms by providing each with image data automatically and evaluating their results.

Finally, the proposed concept must be tested in a realistic field test ideally in a real commissioning of an industrial partner which has to be found.

Acknowledgements

This research work is done within Post Graduate School "Cognitive Computing in Socio-Technical Systems" of Ulm University of Applied Sciences and Ulm University. This work is part of the ZAFH Intralogistik, funded by the

European Regional Development Fund and the Ministry of Science, Research and the Arts of Baden-Württemberg, Germany (F.No. 32-7545.24-17/3/1).

Financial Disclosure

The Post Graduate School is funded by the Ministry for Science, Research and Arts of the State of Baden-Württemberg, Germany.

References

Agarwal, R., 2018. Object Detection: An End to End Theoretical Perspective - A detailed look at the most influential papers in Object Detection [online] Available at <https://towardsdatascience.com/object-detection-using-deep-learning-approaches-an-end-to-end-theoretical-perspective-4ca27eee8a9a> [Accesses 13 May 2019]

Bonkenburg, T., 2016. Robotics in Logistics - A DPDHL perspective on implications and use cases for the logistics industry, Bonn: Deutsche Post DHL Group

Britt, P., 2018. Whitepaper: 10 Robots That Can Speed Up Your Supply Chain, Framingham: Robotics Business Review

Colling, D., Hopfgarten, P., Markert, K., Neubehler, K., Eberle, F., Gilles, M., Jung, M., Kocabas, A., Firmans, K, 2017. PiRo - Ein autonomes Kommissioniersystem für inhomogene, chaotische Lager, Logistics Journal. Proceedings, DOI: 10.2195/lj_Proc_colling_de_201710_01

Costanzo, M., De Maria, G., Natale, C., Pirozzi, S., 2019. Design and Calibration of a Force/Tactile Sensor for Dexterous Manipulation, Sensors, 19(4), 966

de Koster, R. B. M., 2018. Automated and Robotic Warehouses: Developments and Research Opportunities, Logistics and Transport No 2(38)/2018, pp.33-40, DOI: 10.26411/83-1734-2015-2-38-4-18

Dick, I., Ulrich, S., Bruns, R., 2018. Autonomes Greifen mit individuell zusammengestellten Greifern des Greifer-Baukastens, Logistics Journal. Proceedings, ISSN: 2192-9084

Domingos, P., 2012. A Few Useful Things to Know about Machine Learning, Communications of the ACM, 55, 10, pp.78-87, DOI: 10.1145/2347736.2347755

Fetch Robotics, Inc., 2019. Autonomous Mobile Robots That Improve Productivity, [online] Available at <https://fetchrobotics.com/> [Accessed 12 May 2019]

Hans, W. Paulus, D., 2008. Automatisierte Objektaufnahme für Bilddatenbanken. In: Helling, Stephan; Brauers, Johannes; Hill, Bernhard; Aach, Til: 14. Workshop Farbbildverarbeitung. RWTH Aachen: Shaker. S. 143-151.

Hans, W., Knopp, B., Paulus, D., 2009. Farbmetrische Objekterkennung. In: 15. Workshop Farbbildverarbeitung. Berlin: GfAI. S. 43-51.

Hayashi, S., Shigematsu, K., Yamamoto. S., Kobayashi, K., Kohno, Y., Kamata, J., Kurita, M., 2010. Evaluation of a strawberry-harvesting robot in a field test, Biosystems Engineering, Volume 105, Issue 2, 2010, Pages 160-171, ISSN: 1537-5110,

Hui, J., 2018. Object detection: speed and accuracy comparison (Faster R-CNN, R-FCN, SSD, FPN, RetinaNet and YOLOv3), [online] Available at < https://medium.com/@jonathan_hui/object-detection-speed-and-accuracy-comparison-faster-r-cnn-r-fcn-ssd-and-yolo-5425656ae359> [Accesses 13 May 2019]

IAM Robotics, 2019. Making flexible automation a reality, [online] Available at <https://www.iamrobotics.com/> [Accessed 12 May 2019]

Jabalameli, A., Ettehadi, N., Behal, A., 2018. Edge-Based Recognition of Novel Objects forRobotic Grasping, arXiv:1802.08753v1 [cs.RO]

Kohl, A.-K., Pfretzschner, F., 2018. Logistimonitor 2018 Der Wirtschaftszweig in Zahlen - Ergebnisse einer Expertenbefragung von Statista und der Bundesvereinigung Logistik (BVL) e.V.. Bremen

Kraus, W., 2018. Digitale Prozesse im Großhandelsunternehmen: Logistik 4.0 - Roboter im Warenlager, Stuttgart, 09 October 2018. Stuttgart: Fraunhofer IPA

Lenz, I., Lee, H., Saxena, A., 2015. Deep Learning for Detecting Robotic Grasps, The International of Robotics Research, 34(4-5), pp.705-724

Li, F.-F., Johnson, J., Yeung, S., 2017. Detection and Segmentation, CS231n: Convolutional Neural Networks for Visual Recognition,Stanford University, [online] Available at < http://cs231n.stanford.edu/slides/2017/cs231n_2017_lecture11.pdf> [Accessed 10 June 2019]

Lin, T.-Y., Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., Perona, P., Ramanan, D., Zitnick, C. L., Dollár, P., 2015. Microsoft COCO: Common Objects in Context, Computer Vision and Pattern Recognition, arXiv:1405.0312v3 [cs.CV]

Magazino GmbH, 2019. Intelligente Robotik und Lagerlogistik, [online] Available at <https://www.magazino.eu/> [Accessed 12 May 2019]

Microsoft Corporation, 2018. Kinect für Windows. [online] Available at: <https://developer.microsoft.com/de-de/windows/kinect> [Accessed 11 De-cember 2018]

Muscato, G., Prestifilippo, M., Abbate, N., Rizzuto, I., 2005. A prototype of an orange picking robot: past history, the new robot and experimental results, In: Industrial Robot: An International Journal, Vol. 32 Issue: 2, pp.128-138, DOI: 10.1108/01439910510582255

Ouaknine, A., 2018. Review of Deep Learning Algorithms for Image Semantic Segmentation, [online] Available at < https://medium.com/@arthur_ouaknine/review-of-deep-learning-algorithms-for-image-semantic-segmentation-509a600f7b57> [Accessed 1 May 2019]

Photoneo s. r. o., 2018. PhoXi® 3D-Scanner M. [online] Available at: <https://www.phtoneo.com/prduct-detail/phoxi-3d-scanner-m/?lang=de> [Accessed 11 November 2018]

Redmon, J., 2016. YOLO: Real Time Object Detection, [online] <https://github.com/pjreddie/darknet/wiki/YOLO:-Real-Time-Object-Detection> [Accessed 07 June 2019]

Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016. You Only Look Once: Unified, Real-Time Object Detection. Computing Research Repository (CoRR), arXiv:1506.02640 [cs.CV]

Redmon, J., Farhadi, A., 2018. YOLOv3: An Incremental Improvement. arXiv:1506.02640 [cs.CV]

Rother, H., 2017. 3D-Drucken...und dann? - Weiterbearbeitung, Verbindung & Verede-lung von 3D-Druck-Teilen, München: Carl Hanser Verlag

Ruppert, W., Patentanwälte Freischem, 2006. Verfahren zur Aufnahme digitaler Abbildungen. Köln. EP 1 958 148 B1 (active).

Schneider, J., Gruchmann, T., Brauckmann, A., Hanke T., 2018. Arbeitswelten der Logistik im Wandel: Automatisierungstechnik und Ergonomieunterstützung für eine innovative Arbeitsplatzgestaltung in der Intralogistik. In: B. Hermeier, T. Heupel, S. Fichtner-Rosada, S., ed. 2018. Arbeitswelten der Zukunft. Wiesbaden: Springer Fachmedien GmbH. pp.51-66

Schwäke, K., Dick, I., Bruns, R., Ulrich, S., 2017. Entwicklung eines flexiblen, vollautomatischen Kommissionierroboters, Logistics Journal. Proceedings, ISSN: 2192-9084

Stanford, S., Iriondo, R., 2018. The Best Public Datasets for Machine Learning and Data Science - Free Open Datasets for Machine Learning & Data Science, [online] Available at < https://medium.com/towards-artificial-intelligence/the-50-best-public-datasets-for-machine-learning-d80e9f030279> [Accessed 13 May 2019]

Tai, K., El-Sayed, A.-R., Shahriari, M., Biglarbegian, M., Mahmud, S., 2016. State of the Art Robotic Grippers and Applications, Robotics, 5, 11; DOI: 10.3390/robotics5020011

Thiel, M., Hinckeldeyn, J., Kreutzfeldt, J., 2018. Deep-Learning-Verfahren zur 3D-Objekterkennung in der Logistik. In: Wissenschaftliche Gesellschaft für Technische Logistik e. V., 14. Fachkolloquium der WGTL. Wien, Austria, 26-27 September 2018, Rostock-Warnemünde: Logistics Journal

Van Henten, E.J., Van't Slot, D.A., Hol, C.W.J., Van Willigenburg, L.G., 2009. Optimal manipulator design for a cucumber harvesting robot, Computers and Electronics in Agriculture, Volume 65, Issue 2, 2009, Pages 247-257, ISSN 0168-1699, DOI: 10.1016/j.compag.2008.11.004.

Wahl, F., 2016. Pick-by-Robot: Kommissionierroboter für die Logistik 4.0. Future Manufacturing. 2016/5. Frankfurt am Main: VDMA Verlag. pp.16-17

Waser, G., 2014. 360-Grad-Bilder: eine Attraktion für Webshops. Marketing & Kommunikation, 01/2014

Witten, I.H., Frank, E., Hall, M.A., 2011. Data mining - Practical machine learning tools and techniques, 3 ed., Amsterdam: Morgan Kaufmann Publishers Inc.

Zhao, Z.-Q., Zheng, P., Xu, S.-t., Wu, X., 2019. Object Detection with Deep Learning: A Review, arXiv:1807.05511v2 [cs.CV]

Zhang, H., Long, P., Zhou, D., Qian, Z., Wang, Z., Wan, W., Manocha, D., Park, C., Hu, T., Cao, C., Chen, Y., Chow, M., Pan, J., 2016. DoraPicker: An Autonomous Picking System for General Objects, arXiv:1603.06317v1 [cs.RO]