Towards Intelligent Pick and Place Assembly of Individualized Products Using Reinforcement Learning

Caterina Neef[1(✉)], Dario Luipers[1], Jan Bollenbacher[2],
Christian Gebel[1], and Anja Richert[1]

[1] Cologne Cobots Lab, TH Köln - University of Applied Sciences,
Betzdorfer Str. 2, 50679 Cologne, Germany
{caterina.neef,dario.luipers,
anja.richert}@th-koeln.de
[2] Institute of Telecommunications Technology, TH Köln - University of Applied
Sciences, Betzdorfer Str. 2, 50679 Cologne, Germany

Abstract. Individualized manufacturing is becoming an important approach to fulfill increasingly diverse consumer expectations. While there are various solutions for the manufacturing process, such as additive manufacturing, the subsequent automated assembly remains a challenging task. As an approach to this problem, we aim to teach a collaborative robot to successfully perform pick and place tasks by implementing reinforcement learning. For the assembly of an individualized product in a constantly changing manufacturing environment, the simulated geometric and dynamic parameters will be varied. Using reinforcement learning algorithms capable of meta-learning, the tasks will first be trained in simulation, and then performed in a real-world environment where new factors are introduced that were not simulated in training to confirm the robustness of the algorithms. A concept comprised of selected machine learning algorithms, hardware components as well as further research questions to realize the outlined production scenario are the results of the presented work.

Keywords: Machine learning · Reinforcement learning · Meta-learning · Individualized manufacturing · Collaborative robotics

1 Introduction

For decades, robots have been used to automate tasks in the industry sector. Conventional industrial robots are taught to perform one task at a time, are competent at executing this single task and can perform thousands of repetitions accurately. While the automation of such tasks has led to an increase in efficiency and a decrease in manufacturing costs for mass production, it is less applicable for the individualized consumer expectations of today's economy. Globalization, digitalization and the resulting growth of markets have led to an increasing number of product variants and shorter product life cycles [1]. Customers now demand highly individualized products that are designed specifically for them. This change is observable in a wide range of industrial fields [2]. An example is the health-care sector, where personalized medicine

is becoming increasingly important, and additive manufacturing is being used for the production of biomaterials, implants and prosthetics [3].

To enable such individualized manufacturing, traditional programming of machines with repetitive tasks is no longer applicable [2]. Robots with inherent collaborative properties, however, are well suited for constantly changing tasks as they are mobile, highly flexible and adaptable, and can quickly be taught new tasks, even by unskilled workers. We therefore propose to use reinforcement learning (RL) algorithms capable of meta-learning (ML) combined with visual and tactile sensor data to enable collaborative robots to accomplish highly individualized pick and place tasks.

2 State of the Art: Machine Learning and Robotics

Recently, RL has achieved great success in a wide range of different tasks and complex games. The implementation of RL and ML seems promising to enable a robot to perform a pick and place task for unknown objectives and destinations. In RL, an agent interacts with the environment and receives its state. Based on this state, the agent takes an action and receives a new state and reward for the chosen action. Each RL algorithm is designed specifically for a certain task in terms of its architecture and training, and the RL agent needs to be trained from scratch for each task.

ML is an approach to overcome this shortcoming by designing an algorithm in such a way that the agent learns how to learn from a broad distribution of similar tasks. Similar to human learning, an ML agent can apply knowledge it has gained from previously solved corresponding tasks to learn a new task with only a small amount of data.

Recent ML algorithms suitable for RL can be divided into two categories depending on their architecture and optimization goals:

1. Model-based ML approaches [4, 5] generalize to a wide range of learning scenarios, seeking to recognize the task identity from a few data samples and adapting to the tasks by adjusting a model's state (e.g. long short-term memory [LSTM] internal states)
2. Model-Agnostic Meta-Learning (MAML) [6] seeks an initialization of model parameters so that a small number of gradient updates will lead to fast learning on a new task, offering flexibility in the choice of models

Another relevant RL approach is introduced by OpenAI, who have trained a robotic hand to solve a Rubik's Cube despite external perturbations [7].

The main points of this approach are:

- An actor-critic consisting of an artificial neural network (ANN) equipped with LSTM cells to install internal memory
- Automatic domain randomization (ADR) to generate diverse environments with randomized physics and dynamics (e.g. weight and size of the manipulated object)

This results in a system with high robustness and high success rates in the transfer from simulation to testing in the real-world environment. Due to the combination of internal memory and ADR this approach also shows signs of emerging ML.

In RL, it is necessary to provide the learning agent with an extrinsic reward signal. This enables the agent to determine if the actions applied to the environment have a positive effect in the long run. Extrinsic reward signals are called sparse if the reward for a certain action is temporally disentangled from the reward, e.g. only a positive reward is given after every successful task. To tackle this problem of sparse extrinsic rewards, we can divide the approaches in literature into two classes. First, by changing the reward function, e.g. using curiosity-driven exploration [8], which introduces an intrinsic reward function. Second, hierarchical RL methods which try to divide the main task into a sequence of sub-goals can be used. While the main goal is to successfully perform the task, the agent first learns to find a policy for the sub-goals. One popular candidate for this are FeUdal Networks [9] in which the agent is split into two parts. The manager learns to formulate goals and the worker is intrinsically rewarded to follow the goal. A similar approach is Hierarchical Actor-Critic [10] in which the agent learns to set sub-goals to reach the main goal.

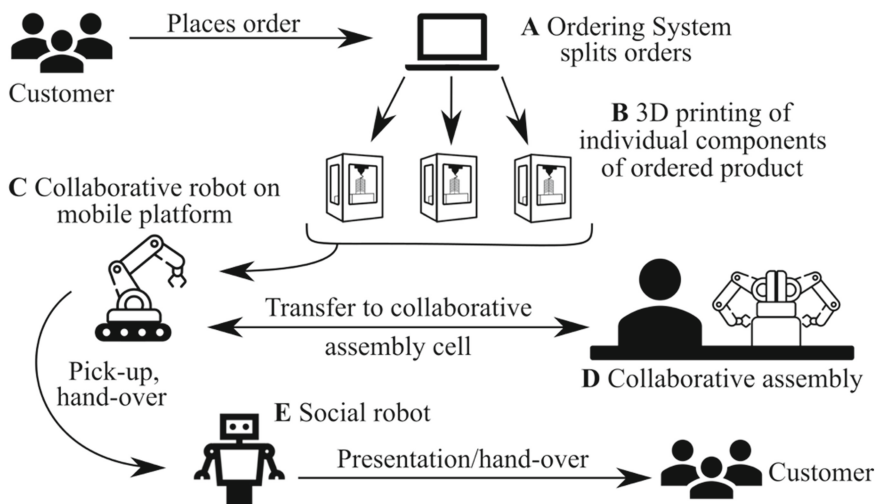## 3 Concept for Manufacturing Scenario

Fig. 1. A customer places an order for an individualized product using the ordering system (A). This system splits the order into individual components, which are manufactured using additive manufacturing (B). Once the manufacturing process is completed, the collaborative robotic arm Panda (Franka Emika, Germany), mounted on an automated guided vehicle (C), removes the components from the 3D printers and transports them to the collaborative assembly cell (D). This cell consists of the collaborative robot YuMi (ABB, Switzerland) and a human worker collaboratively assembling the individual components into the final product. Once the assembly process is completed, the mobile robotic arm transports the finished product and hands it over to a social robot (E), which presents and hands over the product to the customer.

The development of intelligent pick and place tasks for the assembly of products using RL is an integral part of the manufacturing scenario we are setting up in the Cologne Cobots Lab, shown in Fig. 1. It combines individualized production using additive manufacturing, autonomous mobile systems that transport components, and collaborative and social robotics.
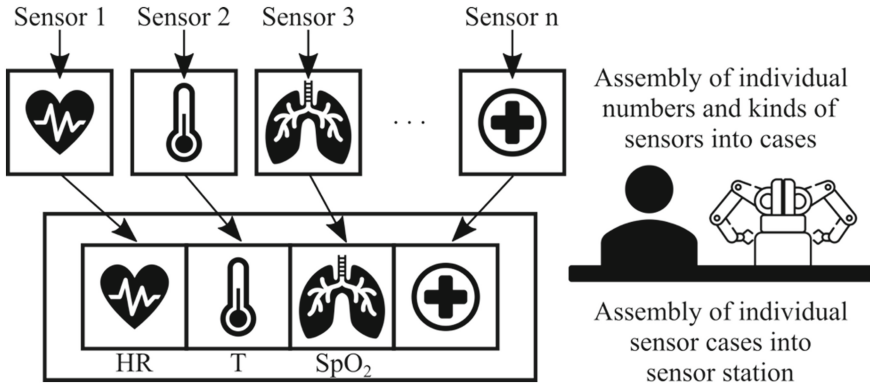

Fig. 2. The goal of the assembly task is to pick individual health care sensors, (e.g. sensors by Maxim Integrated (USA): electrocardiogram (EKG, MAX30003), body temperature (T, MAX30205), blood oxygen saturation (SpO$_2$), MAX30102), place them into their cases, then place the sensor cases in a health care sensor station for the monitoring of health conditions. As the sensor combination changes with each user, reinforcement learning will be used to successfully train an agent to perform each individual assembly task.

As described in Sect. 2, we aim to perform object manipulation tasks using RL in a real-world scenario, in which a specific and useful product is manufactured and assembled. Our goal is to assemble individualized sensor cases for different health care sensors (e.g. to measure body temperature, heart rate or blood oxygen saturation), as shown in Fig. 2. These cases will be used in our research concerned with social robots conducting health assessments [11]. The assembly of individualized products is desirable, as different users with different health conditions require different kinds of information. The long-term goal regarding these sensors is to create individualized wearable devices with different kinds and numbers of health care sensors. These parameters offer a promising approach to a hybrid job shop scheduling or action planning system in which human and robot actions are combined in an optimized way.

4 Approach: Hardware and Machine Learning

In our manufacturing scenario, we will be using various hardware and software/ machine learning components, described in the following.

The individual sensor cases will be manufactured using fused deposition modeling (FDM)/fused filament fabrication (FFM). For the assembly, we will use the YuMi collaborative robot (ABB, Switzerland), which has two arms with 7 degrees of freedom (DoF) each, a payload of 0.5 kg and a pose repeatability of 0.02 mm. To receive additional feedback during object manipulation tasks, its gripper will be outfitted with tactile sensors. This will improve the efficiency and robustness of the grasping task. The sensor provides feedback regarding the grip quality confidence and enables slip detection of the grasped object, which can then be counteracted by improving the applied force of the gripper on the object. For the object detection, we will use a 3D scanner to create heightmaps of the objects. Several area scan cameras will be implemented for vision-based information from various angles and to determine the orientation of the objects.

In order to successfully teach the robot to perform the pick and place task, the trained machine learning algorithm needs to recognize which produced element belongs to the corresponding case. To accomplish this, the tools presented in Sect. 2 will be implemented and combined. By implementing ADR [7], the agent will be trained to realize a robust system with a high success rate in the transfer from simulation to the real world. Additionally, due to the implementation of ML and solutions for sparse reward, the learning time of the agent will be decreased. The main consideration for this approach is:

- How can we apply (a combination of) machine learning algorithms to generalize pick and place assembly tasks (i.e. various weights, sizes, geometries, quantities) for individualized products?

The Meta-Learners by Finn, Duan, and Mishra [4–6, 12] have been successfully taught to perform simple RL tasks, such as locomotion tasks. In Meta-World [13], these algorithms have been tested for a variety of advanced robotic tasks. However, the results for the presented problems have a maximum rate of successfully solved tasks of 50%. This does not seem high enough to confidently and robustly perform the presented pick and place operations. We therefore aim to reach a higher success rate with the following concept:

1. Generate and define the described collaborative tasks in the simulated environment using Mujoco (Roboti LLC, USA)
2. Find a suitable configuration and combination of ML algorithms which will solve the pick and place task
3. Implement and fine tune the trained algorithm in the real-world production scenario, applying the concept of transfer-learning

5   Conclusion and Outlook

In this paper, we propose an approach to successfully perform intelligent pick and place tasks for the assembly of individualized products using RL algorithms capable of ML.

The developed demonstrator will then be used to answer the following research questions, which are both of technical and socio-technical nature:

- Which algorithms have which impact on the robustness of the system? How can we assure that the robustness reached in simulation can be transferred to the real-world environment, e.g. using ADR?
- How can we implement a dynamic work space for the robot when working collaboratively with a human? How can a human be integrated into the collaborative assembly process in a way that is both sensible and effective?

In the future, we plan on fully implementing the developed assembly process into our manufacturing scenario described in Sect. 3. The manufacturing scenario includes the transportation of individual parts using automated guided vehicles and presenting the final product to the customer. A further goal is to study the collaborative assembly process between humans and robots. This is the focus of another research project in our lab, which aims to achieve adaptive human-robot collaboration through the implementation of sensors to detect the user's status (e.g. focus, stress). The combined results of these projects will contribute to an optimal collaborative working process.

References

1. Brettel, M., Friederichsen, N., Keller, M., Rosenberg, M.: How virtualization, decentralization and network building change the manufacturing landscape: an Industry 4.0 perspective. Int. J. Mech. Ind. Sci. Eng. **8**, 37–44 (2014)
2. Outón, J.L., Villaverde, I., Herrero, H., Esnaola, U., Sierra, B.: Innovative mobile manipulator solution for modern flexible manufacturing processes. Sensors **19**, 5414 (2019). https://doi.org/10.3390/s19245414
3. Zadpoor, A.A., Malda, J.: Additive manufacturing of biomaterials, tissues, and organs. Ann. Biomed. Eng. **45**, 1–11 (2017). https://doi.org/10.1007/s10439-016-1719-y
4. Duan, Y., Schulman, J., Chen, X., Bartlett, P.L., Sutskever, I., Abbeel, P.: RL^2: fast reinforcement learning via slow reinforcement learning. arXiv:1611.02779 [cs, stat] (2016)
5. Mishra, N., Rohaninejad, M., Chen, X., Abbeel, P.: A simple neural attentive meta-learner. arXiv:1707.03141 [cs, stat] (2018)
6. Finn, C., Abbeel, P., Levine, S.: Model-agnostic meta-learning for fast adaptation of deep networks. In: Proceedings of the 34th International Conference on Machine Learning-Volume 70, pp. 1126–1135 (2017). http://www.JMLR.org
7. OpenAI, Akkaya, I., Andrychowicz, M., Chociej, M., Litwin, M., McGrew, B., Petron, A., Paino, A., Plappert, M., Powell, G., Ribas, R., Schneider, J., Tezak, N., Tworek, J., Welinder, P., Weng, L., Yuan, Q., Zaremba, W., Zhang, L.: Solving rubik's cube with a robot hand. arXiv:1910.07113 [cs, stat] (2019)
8. Pathak, D., Agrawal, P., Efros, A.A., Darrell, T.: Curiosity-driven exploration by self-supervised prediction. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 16–17 (2017)
9. Vezhnevets, A.S., Osindero, S., Schaul, T., Heess, N., Jaderberg, M., Silver, D., Kavukcuoglu, K.: Feudal networks for hierarchical reinforcement learning. In: Proceedings of the 34th International Conference on Machine Learning-Volume 70, pp. 3540–3549 (2017). http://www.JMLR.org
10. Levy, A., Konidaris, G., Platt, R., Saenko, K.: Learning multi-level hierarchies with hindsight. arXiv:1712.00948 [cs] (2019)

11. Richert, A., Schiffmann, M., Yuan, C.: A nursing robot for social interactions and health assessment. In: International Conference on Applied Human Factors and Ergonomics, pp. 83–91. Springer (2019)
12. Finn, C., Xu, K., Levine, S.: Probabilistic model-agnostic meta-learning. In: Bengio, S., Wallach, H., Larochelle, H., Grauman, K., Cesa-Bianchi, N., Garnett, R. (eds.) Advances in Neural Information Processing Systems 31, pp. 9516–9527. Curran Associates, Inc. (2018)
13. Yu, T., Quillen, D., He, Z., Julian, R., Hausman, K., Finn, C., Levine, S.: Meta-world: a benchmark and evaluation for multi-task and meta reinforcement learning. arXiv:1910.10897 [cs, stat] (2019)