AI-Based Human-Robot Cooperation for Flexible Multi-Variant Manufacturing

Manuel A. Ruiz Garcia, Erwin Rauch and Damiana Salvalai
Industrial Engineering and Automation (IEA)
Free University of Bozen-Bolzano
39100 Bolzano, Italy
ManuelAlejandro.RuizGarcia@unibz.it, erwin.rauch@unibz.it, damiana.salvalai@unibz.it

Dominik T. Matt
Fraunhofer Italia Research
39100 Bolzano, Italy
dominik.matt@fraunhofer.it

Abstract

The flexibilization and intelligent design of multi-variant production systems plays a major role in mass customization. One major difficulty in such manufacturing systems is often the rapid retooling of production lines and workstations in order to respond as flexibly as possible to changing customer requirements and variants. In most cases, this also includes reprogramming of machines and robots, which have to be adapted to new work sequences. However, such a reprogramming overhead can be effectively reduced by endowing the workstation with the ability to autonomously understand the dynamic changes on the assembly sequence. The goal of the present work is to present a concept to automate such a reconfiguring process based on machine learning models and artificial intelligence. The proposed paradigm aims to highlight the role of the human-robot cooperation on the effective achievement of flexible multi-variant manufacturing in shared collaborative workspaces. In particular, the underlying problems of scene understanding, tasks modelling and robot autonomy are deconstructed and discussed. In other words, the autonomous reconfiguration is defined in terms of the process awareness of the collaborative robot through the monitoring, identification and prediction of the environment and operator states in terms of previous, current and predicted assembly steps. Modelling the manufacturing process allows to abstract the spatiotemporal relations between the assembly steps of each variant's sequence, independent of the operator actually executing the task. As a result, a flexible multi-variant assembly is obtained by the dynamic and intelligent adaptation of the collaborative robot integrated in the assembly workstation. The concept is partially implemented through an experimental setup of a collaborative assembly station in a learning factory lab.

Keywords
Industry 4.0; Flexible Manufacturing; Artificial Intelligence, Machine Learning, and Human Robot Collaboration

1. Introduction
The last decades were marked by the further development of modern production systems and the introduction of Industry 4.0 in manufacturing (Rauch et al., 2020). The concepts and technologies of Industry 4.0 are mostly aimed at the networking of production and the efficient design of production systems (Kagermann et al., 2013). Industry 4.0 is also seen as an enabler for the flexibilization of production systems, which allows the production of customized products up to batch size 1 in a highly efficient way and at similar unit costs as in mass production. Thus, Industry 4.0 also represents an important milestone for mass customization (Piller, 2007). The developments of the last years also show a strong progress in sustainability in production (Matt et al., 2020), and the use of artificial intelligence (AI) in manufacturing (Woschank et al., 2013). Finally, according to our hypothesis, AI in combination with existing industry

4.0 technologies such as human-robot collaboration should have an immense potential for the realization of seamless human-robot cooperation for flexible multi-variant manufacturing systems.
Therefore, we would like to present in this paper a concept of human-robot cooperation in assembly to show how machine learning (ML) and AI can revolutionize the flexibility of manufacturing in the future.
The paper is structured as follows: in section 2 we first discuss the theoretical background of multi-variant manufacturing and AI in production. In section 3 we will explain the concept of the AI-based human-robot cooperation. Then in section 4 we will present some preliminary results on the use of deep neural networks for the operator's tracking, as part of the use case in a laboratory study and discuss the potential of our concept in section 5. Finally, the conclusions in section 6 are summarized and an outlook for future research is given.

2. Literature Review

2.1 Multi-Variant Manufacturing and Mass Customization
Increasing product variety can be considered as one possible strategy to enable companies to maintain and increase market shares through satisfying the variety seeking behavior of customers (Goetzfried, 2012). Manufacturing companies are confronted with this and many other factors of a dynamic environment and perceive the corresponding complexity e.g. through the frequent adaption of the manufacturing system in order to meet the current and future market demands. These market demands mainly result from megatrends such as the diversified customer demands, shorter product life cycles, shortage of resources and declined manufacturing depths (Abele and Reinhart, 2011). This is why mass customization has been introduced in manufacturing in order to face these challenges and to adapt manufacturing systems to the increasing product variety. In this sense mass customization is defined to deliver products and services that best meet individual customers' needs with near mass production efficiency (Tseng et al., 1996). A way to achieve this goal is flexibilization of manufacturing systems.
In the last years, a lot of research has been conducted once to better understand the challenges of multi-variant manufacturing. The authors in Johansson et al. (2016) analyzed the challenges for the operator due to increasing complexity in manual assembly of trucks as a consequence of an increase in product variety. Products with high variety make line balancing more challenging in terms of high cycle time variation. A variation in work content makes the production sequence sensitive to production disturbances and makes it harder to increase utilization. Their study has shown the difficultness to design and create robust production systems when manufacturing products with high variety. The authors in Dombrowski et al. (2014) confirm that for multi-variant manufacturing systems it is crucial to manage not only product variety at a product level but also process variety due to differences in assembly sequence, in parts used for the fabrication and assembly as well as the reduction of necessary changes of the manufacturing system. The goal is to achievie such a flexibility to change from one product to the other without the need to stop production for a changeover or other manual adaptations of the manufacturing system.

2.2 Human-Robot Collaboration as Industry 4.0 Technology to Foster Flexibility in Manufacturing
A collaborative robot is an industrial robot, which is able to interact physically and safe with humans in a shared and collaborative hybrid workspace (Gualtieri et al., 2019, Gualtieri et al. 2020a). Human robot collaboration is one of the nine key technologies of Industry 4.0 (Rüßmann et al., 2015). The use of sensor technology in the robot makes the robot a safe companion and helper for the employee. Where physical barriers used to be necessary to ensure safety in the workplace, cobots can now share a common workspace with the robot, making them flexible assistance systems for flexible and multi-variant production (Mark et al., 2019).
According to Gualtieri et al. (2020b) collaborative robotics in assembly is one of the potential technologies of Industry 4.0 to increase flexibility in modern manufacturing systems. In addition, the possibility to increase the system. However, each robot, also if collaborative robot, needs to be programmed in order to carry out the needed movements and operations in the assembly of products with a high number of variants. Therefore, in research regarding collaborative robotics there is a need for technical solutions to rapidly adapt the assembly workstation (including the robot program and paths) to different product variants.

2.3 Artificial Intelligence and Machine Learning as Solution for Flexible Human-Robot Collaborative Assembly
AI can become in the next years one of the most important drivers for realizing intelligent factories in manufacturing, perhaps more than any other technological innovation in recent years. Very often AI is mentioned together with machine learning (ML) and deep learning (DL). Basically, we can say that these three terms are different concepts on

different levels. In general, deep learning is a subset of machine learning, and machine learning is a subset of artificial intelligence (Garbade, 2018).
AI and ML can be used in human-robot collaboration for example to implement gesture recognition for the identification of the operators movements carrying out assembly tasks. A recent trend of gesture identification is to use learning algorithms, especially for static gesture detection that can be represented in a single frame (Liu and Wang, 2018). Learning algorithms utilise machine learning algorithms to identify gestures from raw sensor data. Learning algorithms such as Support Vector Machine (SVM), Artificial Neural Networks (ANN) and Random Decision Forests (RDF) are e.g. widely applied in gesture recognition systems (Tang et al, 2014; Ronfard et al., 2002; Lee et al., 2003).
Timely context awareness is key to improving operation efficiency and safety in human-robot collaboration (HRC) for intelligent manufacturing. Visual observation of human workers' motion provides information about the specific tasks to be performed (Wang et al., 2018a).

2.4 Research Question
AI can be seen in the collection of current literature, there is currently still a great need to implement the potential of artificial intelligence in production by testing and finally realizing applications. The combination of human-robot collaborative assembly together with computer vision and AI represents a great opportunity to massively increase flexibility in multivariate production and thus enable the cost-effective production of smallest batch sizes and thus individual products.
The research question we therefore ask ourselves in this paper is as follows:

- How flexibility in multi-variant manufacturing can be improved through the exploitation of machine learning and AI in human-robot collaboration?

In order to discuss this research question, we first present the basic concept of our research in the next section, followed by a preliminary implementation and discussion of a laboratory experiment.

3. AI-Based Human-Robot Cooperation in Assembly
The most dangerous risk specific to robots are the unexpected collisions between the robot and the environment (Siciliano and Khatib, 2016). When an unexpected exertion occurs between a collaborative robot and its surrounding environment, impact forces are eased thanks to their lightweight design and compliant mechanisms and control.
However, avoiding unexpected force exertions implies foreseeing dangerous situations, and thus it relies on sensing, situational awareness, planning and decision making capabilities. Therefore, without suitable exteroceptive sensing a collaborative robot cannot be considered as a safe companion in the context of human-robot cooperation. Indeed, to safely interact with a human operator a collaborative robot must predict and prevent any risky circumstances based on its own situational awareness.
To this end it is required to associate to the the human operator a set of meaningful spatio-temporal features that allows –with some degree of accuracy, within a finite time horizon– to model and predict the operator's behaviour. In terms of safety, it is required to sense and predict the operator's motion. In terms of cooperation, is is required to understand and predict the operator's actions and intentions.
We identify three major synergic elements that contribute to this adaptation, as shown in Fig. 1: (i) scene understanding, (ii) tasks modelling and (iii) robot autonomy. Although these general problems can be unreasonable complex, within the context of assembly workstations where different constraints are imposed to the environment and due the cyclic nature of the assembly process, the analysis of each element can be greatly simplified.
In particular, the scene understanding problem can be restricted to the recognition and tracking of assembly objects and to the operator's motion tracking. In this work we will focus only on the latter problem. A throughout treatment of the problem of object recognition in smart manufacturing is found in Riordan et al. (2019). On the other hand, since only the operator and robot actions can cause a process state transition, the tasks modelling problem can be restricted to the recognition and prediction of the actions executed by the operator. Finally, assuming that the robot manipulator is attached to a fixed base, the robot autonomy can be limited to the problems of action planning and safe tasks execution.

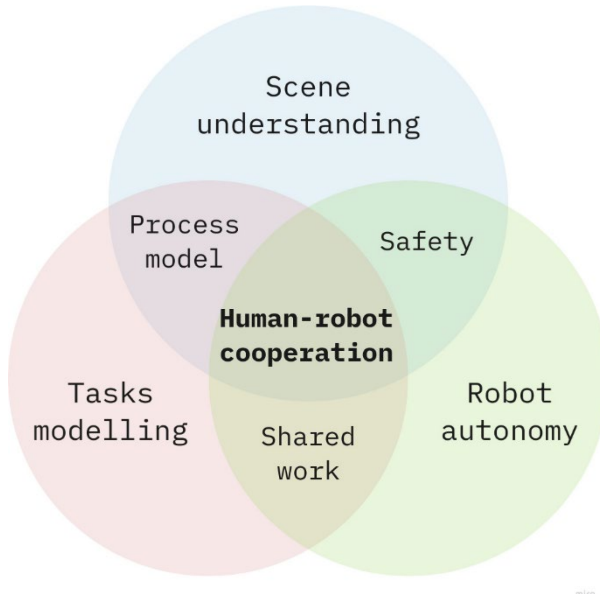

Figure 1. AI-based human-robot cooperation

3.1 Operator's Tracking

Due the stochastic nature of the operator's body, head, arms, etc., movements while executing an assembly task, the problem of monitoring and predicting the operator's motion can be considered as a particular instance of a filtering problem (Tan and Arai, 2011). That is, based on a set of past possibly noisy observations of the operator's pose, determine the best estimate of the current operator's motion.

Today on research and industry exist a wide range of different sensing technologies that can be used to measure the operator's pose. The current technological trend points toward multiple networking range sensing devices or computer vision based systems providing analogous measurements (Ferrari et al., 2018; Gkournelos et al., 2018; Agethen et al., 2016). The use of multiple sensors not only ensures a better accuracy of the estimation but also accounts for the decreasing point-density at far distances of a single sensors. Moreover, different view points are required for a reliable identification of features or markers.

However, state-of-the-art deep learning models for pose estimation in RGB images (Cao et al., 2017) and 2D lidar data (Weinrich et al., 2014) allows to reach high-levels of accuracy. Indeed, the larger field of view of lidar sensors allows to track the operator beyond the field of view of the RGB-D sensor.

3.2 Operator's Intentions Prediction

In industrial manufacturing scenarios, the problem of task prediction is greatly simplified by the cyclic nature of the operator's work. Any manufacturing cycle is indeed defined by a finite set of atomic tasks. However, the order in which such atomic tasks are performed by the operator to conclude the cycle, in general, is not uniquely defined. Therefore, for a machine to be aware of the current state of an assembly cycle, it is required to recognize any atomic task executed by the operator and to model the transitions between them (Alati et al., 2019a).

Identifying a task implies understanding the actions being performed by the operator, while understanding the transitions between tasks implies predicting the operator's intentions. Based on this idea, the prediction of intentions problem can be analysed in terms of two distinct processes: (i) action recognition, and (ii) action prediction.

Action recognition refers to the prompt identification of the current task executed by the operator. The aim of this process is to continuously monitoring the operator's actions on real-time. The action identification can be driven by different cues, like gestures (Carrasco and Clady, 2010), scene objects being manipulated (Koppula and Saxena, 2015) or environmental information (Casalino et al., 2018). Action recognition has been also extensively studied in terms of whole body motion tracking and segmentation (Natola et al., 2015; Tome et al., 2017).

Action prediction refers to the total or partial reconstruction of the possible sequence of actions that the operator would execute just after concluding the current task. Consequently, this process implies the generation and constant refinement of an action transition model (Zanchettin and Rocco, 2017; Zanchettin et al., 2018). In general it can be also assumed that all operator's states and actions are fully observable and that only one action a time can be executed. Although there are no specific manufacturing datasets for the evaluation of action recognition models. In recent years different deep network architectures had been demonstrated high levels of accuracy on totally unrelated but similar manipulation tasks, like the one proposed by the Epic Kitchens challenge (Wang et al., 2018b; Damen et al., 2018).

3.3 Action Planning in Collaborative Workstations

The assembly process modelling extends the action transitions modelling by superimposing the robot states on top of it, allowing the real-time analysis and generation of the robot plan. In other words, the robot's collaborative behaviour is achieved by dynamically allocating its tasks, in terms of the predicted operator's actions and relative transitions (Alati et al., 2019b).

As a result, the objective of the action planning is to reach a designated assembly process goal state. This implies that any goal state includes both the operator's and robot's states. Therefore, it is expected that one particular goal can be reached from a finite set of initial candidate states, each one depending on the particular sequence of actions performed by the operator.

Consequently, in human-robot collaborative environments, the action planning process deals with the robot behaviour adaptation (Mitsunaga et al., 2008) to the time-varying set of constraints imposed the operators actions. In turns, imposed by the customer requirements, diversity of the available manufacturing variants and operator's task execution preferences (Munzer et al., 2017). Therefore, the robot adaptation should provide a proactive (anticipatory) collaborative behaviour driven by the different forms of human-robot interaction associated to each target goal (Mason and Lopes, 2011). In other words, robots working alongside humans should model how to anticipate a belief about possible future human actions (Hsieh et al., 2015).

In complete analogy to the operator's case, the cyclic nature of the assembly process implies that there's a predefined number of goals that the robot can reach, a finite set of deterministic actions that it can perform and a finite set of states that it can have. Moreover, it can be also assumed that the robot can only execute one action a time. However, in general, the execution time of any planned cannot be defined in advance since it also depends on the current state of the assembly process. Specially in the cases when the action execution requires explicit synchronization with the operator.

4. Experimental Setup

In this section we describe our testing and still under development experimental setup for human-robot cooperation in flexible manufacturing. In particular, we describe our implementation for addressing the operator's motion tracking problem. This work is built on top of a collaborative assembly workstation (see Fig. 2 and 3) developed at the Smart Mini Factory Laboratory (SMF) of the Free University of Bozen-Bolzano. The assembly tasks consist of the assembly of different variants of pneumatic cylinders. The workstation is equipped with a mobile workbench, a block-and-tackle for lightweight applications, an integrated Kanban rack, a working procedures panel, a double lighting system, an industrial screwer and a knee lever press. Further the operator is supported by a Universal Robot UR3 cobot. The collaborative robot takes over non-value adding tasks – from a lean management stand point (Matt, 2009; Rauch et al., 2017) – like pick-and-give tasks to eliminate handling time of the operator.

The sensing system is composed by a ZED-mini stereo camera and a PSENscan 2D lidar scanner with an opening angle of 275 degrees and a measurement range of up to 5.5 meters. The laser scan is aligned with the ground plane and at a fixed height of 45 cm above the ground.

We model the operator motion with a constant acceleration Kalman filter for each kinematical joint of the operator body, including legs, arms and head. Only the upper part of the operator's body can be observed by the stereo camera and only the legs are visible by the lidar scan. We rely on the intrinsic data fusion capabilities of the filter to exploit such complementary sources of information, allowing an accurate estimate of the pose of the operator's body. The upper body pose is obtained on the left RGB image of the stereo camera based on the OpenPose deep pose regression network (Agethen et al., 2016) and then projected into the sensor's point cloud, as can be observed in Fig. 2. The leg detector uses an AdaBoost classifiers ensemble of binary decision trees. The operator's legs are detected based on the GANDALF leg detector (Cao et al., 2017). In terms of safety, tracking the operator's motion allows to understand where critical body parts (see Fig. 3) are located in real-time. Therefore, it is possible to promptly modulate the relative velocity between the operator and the robot or to stop the robot motion so to prevent any risky circumstance.

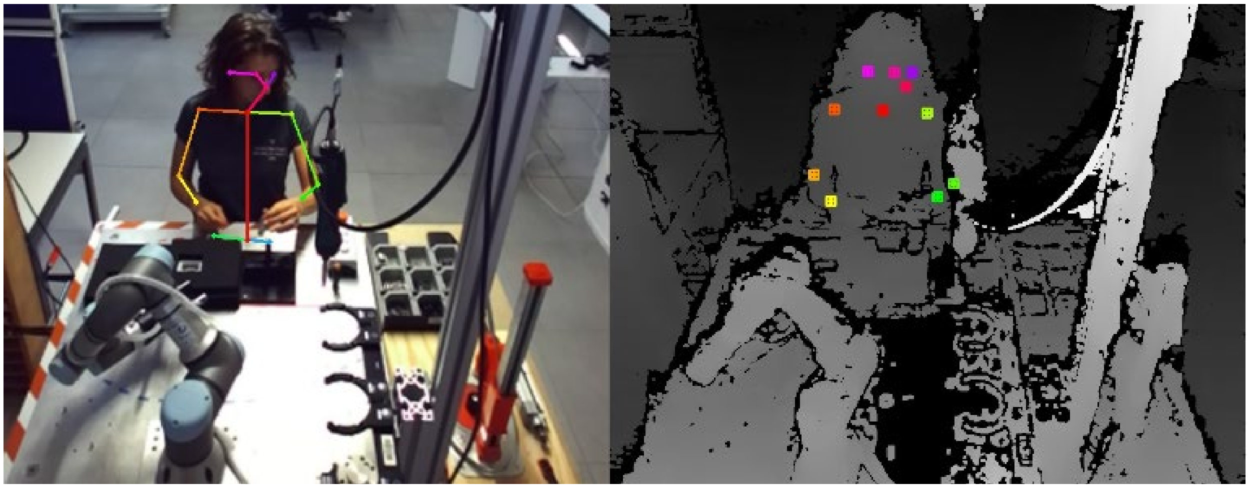

Figure 2. Operator's tracking. The upper body pose is obtained directly from the RGB stream based using the OpenPose framework (left). Such 2D pose is then back-projected into the point cloud to obtain the 3D pose (right).

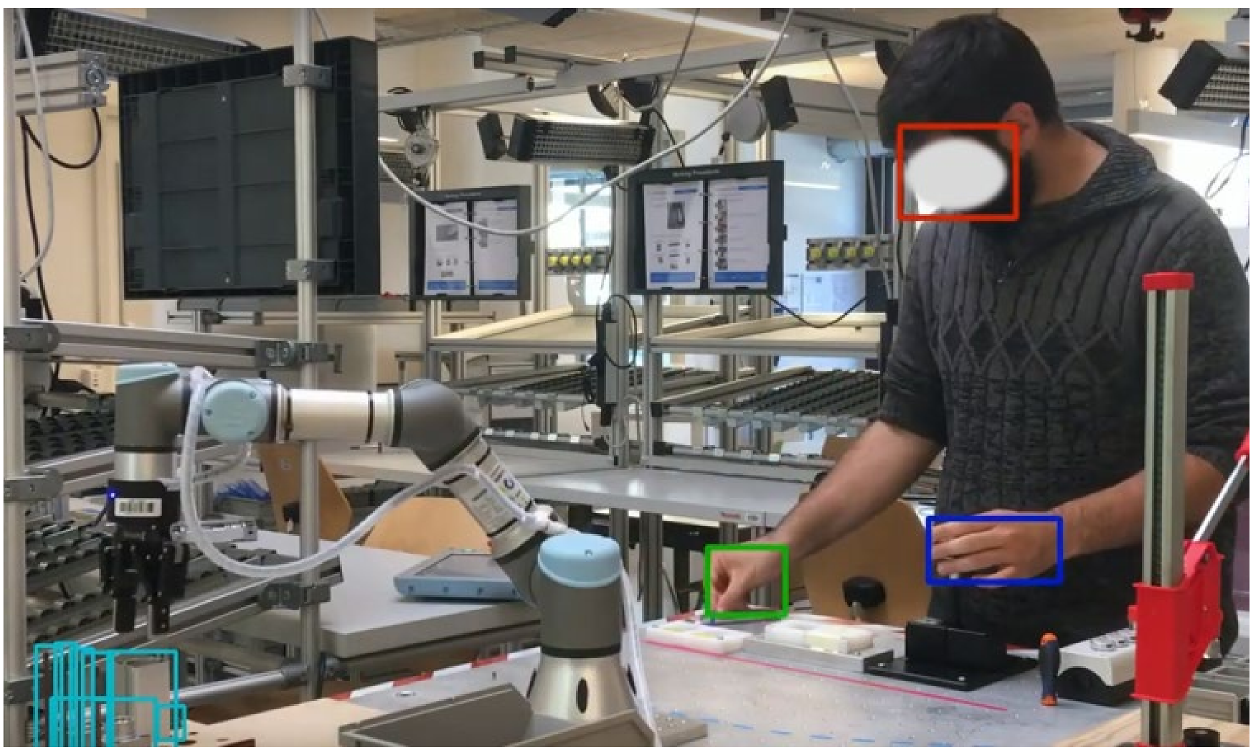

Figure 3. Operator's tracking. Critical body parts are monitored in real-time to ensure a safe human-robot cooperation during the assembly task

5. Discussion

5.1 Potentials for the Future Application of AI-based Human-Robot Collaboration in Industrial Assembly

The potential for applying the concept of AI-based human-robot collaboration described in Section 3 and the preliminary tests with the experimental setup shown in Section 4 is promising and manifold. In summary, the following potentials for future application in industrial assembly can be derived:

- **Automatic adaptation of the production system in real time** (in the specific case of the collaborative robot), by means of image processing and AI intelligence detecting from the operator's work sequence when a product or variant change occurs.
- **Product flexible and multi-variant production** and thus the production of different products or variants without the necessity of batch size formation.
- **Minimizing idle times** of the robot system and thus increasing productivity by enabling the robot to predict work sequences before they start and thus to start tasks for handling components early.
- **Reduction of the cognitive workload of the operator**, who does not have to think about changing the production system and thus a full concentration on the value-adding activities that require the indispensable skills of the human.
- **Monitoring of process quality** by monitoring critical processes to see if the worker is following or deviating from the optimal work sequence

5.2 Implications for Academia and Practitioners

The research presented in this paper on AI-based human-robot cooperation is far from being complete and exhaustive. For academia, this means, that there are still many points that need to be investigated and further developed in the future. However, the presented approach shows how the use of AI and ML in manufacturing can have a great impact on the flexibility of future manufacturing systems.
From a company's view, the presented concept of having a workstation with an operator, a robot and an intelligent system for the observation and analysis of human and robot movements offers an ideal possibility how future workstations can be applied in flexible factories for multi-variant manufacturing and mass customization. By means of dynamic (real-time) recognition of the parts currently being processed from the analysis of operators activities, the production system (in this case the robot) can be adapted dynamically to a product change in quasi real-time. A future industrial application of such systems therefore presents an important step towards the vision for flexible automation in manufacturing and assembly of products with lot size one or smaller lot sizes.

6. Conclusion and Outlook

This paper introduces the concept of AI-based human-robot cooperation in manual or hybrid assembly. First of all it is explained how computer vision can be used for scene understanding and to track objects and the operator. Next, it shows how task modelling can be supported by the recognition and AI-based prediction of the actions executed by the operator. All this facilitates also the implementation of human-robot cooperation and an autonomous behavior and reaction of the robot system in case of a dynamic change of products or variants during assembly. The concept is supported by a preliminary experimental setup in a laboratory environment.
Finally, the future potentials of this technology in the field of industrial assembly are derived and summarized. The research in this area is still in an early stage, so this contribution aims to motivate other researchers to do further research and practitioners to collaborate with research institutions for conducting tests on practical applications in real case studies.

Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 734713.

References

Abele, E. and Reinhart, G., *Zukunft der Produktion*, Hanser, 2011.

Agethen, P., Otto, M., Mengel, S. and Rukzio, E., Using marker-less motion capture systems for walk path analysis in paced assembly flow lines. *Procedia CIRP*, vol. 54, pp. 152-157, 2016.

Alati, E., Mauro, L., Ntouskos, V. and Pirri, F., Help by Predicting What to Do. *Proceedings of 2019 IEEE International Conference on Image Processing* (ICIP), pp. 1930-1934, 2019.

Alati, E., Mauro, L., Ntouskos, V. and Pirri, F., Anticipating Next Goal for Robot Plan Prediction. In *Proceedings of SAI Intelligent Systems Conference*, pp. 792-809, Springer, 2019.

Cao, Z., Simon, T., Wei, S. E. and Sheikh, Y., Realtime multi-person 2d pose estimation using part affinity fields. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7291-7299, 2017.

Carrasco, M. and Clady, X., Prediction of user's grasping intentions based on eye-hand coordination. *Proceedings of 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 4631-4637, 2010.

Casalino, A., Messeri, C., Pozzi, M., Zanchettin, A. M., Rocco, P. and Prattichizzo, D., Operator awareness in human–robot collaboration through wearable vibrotactile feedback. *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 4289-4296, 2018.

Damen, D., Doughty, H., Maria Farinella, G., Fidler, S., Furnari, A., Kazakos, E., Moltisante, D., Munro, J., Perrett, T., Price, W. and Wray, M., Scaling egocentric vision: the Epic-Kitchens dataset. In *Proceedings of the European Conference on Computer Vision* (ECCV), pp. 720-736, 2018.

Dombrowski, U., Krenkel, P., and Ebentreich, D., Adaptability within a Multi-variant Serial Production. *Procedia CIRP*, vol. 17, pp. 124-129, 2014.

Ferrari, E., Gamberi, M., Pilati, F. and Regattieri, A., Motion Analysis System for the digitalization and assessment of manual manufacturing and assembly processes. *IFAC-PapersOnLine*, vol. 51, no. 11, pp. 411-416, 2018.

Garbade, D. M. J., Clearing the Confusion: AI vs Machine Learning vs Deep Learning Differences. *Towards Data Science*, September, 14, 2018.

Gkournelos, C., Karagiannis, P., Kousi, N., Michalos, G., Koukas, S. and Makris, S., Application of wearable devices for supporting operators in human-robot cooperative assembly tasks. *Procedia CIRP*, vol. 76, pp. 177-182, 2018.

Goetzfried, M., Complexity Index: Assessing the impact of external and internal complexity on production plant performance, 2012.

Gualtieri, L., Rauch, E., Vidoni, R., and Matt, D. T., An evaluation methodology for the conversion of manual assembly systems into human-robot collaborative workcells. *Procedia Manufacturing*, vol. 38, pp. 358-366, 2019.

Gualtieri, L., Palomba, I., Wehrle, E. J. and Vidoni, R., The Opportunities and Challenges of SME Manufacturing Automation: Safety and Ergonomics in Human–Robot Collaboration. In: Matt D., Modrák V., Zsifkovits H. (eds) *Industry 4.0 for SMEs.* Palgrave Macmillan, pp. 105-144, 2020.

Gualtieri, L., Palomba, I., Merati, F. A., Rauch, E., and Vidoni, R., Design of Human-Centered Collaborative Assembly Workstations for the Improvement of Operators' Physical Ergonomics and Production Efficiency: A Case Study. *Sustainability*, vol. 12, no. 9, 3606, 2020.

Hsieh, M. A., Khatib, O. and Kumar, V., *Experimental Robotics: The 14th International Symposium on Experimental Robotics*, vol. 109, Springer, 2015.

Johansson, P. E., Mattsson, S., Moestam, L., and Fast-Berglund, Å., Multi-variant truck production-product variety and its impact on production quality in manual assembly. *Procedia CIRP*, vol. 54, pp. 245-250, 2016.

Kagermann, H., Helbig, J., Hellinger, A., & Wahlster, W., *Recommendations for implementing the strategic initiative INDUSTRIE 4.0: Securing the future of German manufacturing industry*; final report of the Industrie 4.0 Working Group. Forschungsunion, 2013.

Koppula, H. S. and Saxena, A., Anticipating human activities using object affordances for reactive robotic response. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 1, pp. 14-29, 2015.

Lee, S. J., Ouyang, C. S., and Du, S. H., A neuro-fuzzy approach for segmentation of human objects in image sequences. *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 33, no. 3, pp. 420-437, 2003.

Liu, H. and Wang, L., Gesture recognition for human-robot collaboration: A review. *International Journal of Industrial Ergonomics*, vol. 68, pp. 355-367, 2018.

Mark, B. G., Hofmayer, S., Rauch, E., and Matt, D. T., Inclusion of Workers with Disabilities in Production 4.0: Legal Foundations in Europe and Potentials Through Worker Assistance Systems, *Sustainability*, vol. 11, no. 21, 5978, 2019.

Mason, M. and Lopes, M., Robot self-initiative and personalization by learning through repeated interactions. In *Proceedings of 6th ACM/IEEE International Conference on Human-Robot Interaction* (HRI), pp. 433-440, 2011.

Matt, D. T., Design of lean manufacturing support systems in make-to-order production. *Key Engineering Materials*, vol. 410, pp. 151-158, 2009.

Matt, D. T., Orzes, G., Rauch, E., & Dallasega, P., Urban production–A socially sustainable factory concept to overcome shortcomings of qualified workers in smart SMEs. *Computers & Industrial Engineering*, vol. 139, 105384, 2020.

Mitsunaga, N., Smith, C., Kanda, T., Ishiguro, H. and Hagita, N., Adapting robot behavior for human--robot interaction. *IEEE Transactions on Robotics*, vol. 24, no. 4, pp. 911-916, 2008.

Munzer, T., Toussaint, M. and Lopes, M., Preference learning on the execution of collaborative human-robot tasks. In *Proceedings of IEEE International Conference on Robotics and Automation* (ICRA), pp. 879-885, 2017.

Natola, F., Ntouskos, V., Sanzari, M. and Pirri, F., Bayesian non-parametric inference for manifold based MoCap representation. In *Proceedings of the IEEE International Conference on Computer Vision*, pp. 4606-4614, 2015.

Piller, F. T., Observations on the present and future of mass customization. *International Journal of Flexible Manufacturing Systems*, vol. 19, no. 4, pp. 630-636, 2007.

Rauch, E., Dallasega, P., & Matt, D. T., Critical factors for introducing lean product development to small and medium sized enterprises in Italy. *Procedia CIRP*, vol. 60, pp. 362-367, 2017.

Rauch, E., Unterhofer, M., Rojas, R. A., Gualtieri, L., Woschank, M. and Matt, D. T., A Maturity Level-Based Assessment Tool to Enhance the Implementation of Industry 4.0 in Small and Medium-Sized Enterprises. *Sustainability*, vol. 12, no. 9, pp. 3559, 2020.

Riordan A, Toal D, Newe T and Dooly G. Object recognition within smart manufacturing. *Procedia Manufacturing* vol. 38, pp. 408-414, 2019.

Ronfard, R., Schmid, C. and Triggs, B., Learning to parse pictures of people. In *European Conference on Computer Vision*, pp. 700-714, Springer, 2002.

Rüßmann, M., Lorenz, M., Gerbert, P., Waldner, M., Justus, J., Engel, P., and Harnisch, M., Industry 4.0: The future of productivity and growth in manufacturing industries. *Boston Consulting Group*, vol. 9, no. 1, pp. 54-89, 2015.

Siciliano, B., and Khatib, O., Springer handbook of robotics. Springer, 2016.

Tan, J. T. C. and Arai, T., Triple stereo vision system for safety monitoring of human-robot collaboration in cellular manufacturing. In 2011 *IEEE International Symposium on Assembly and Manufacturing* (ISAM), pp. 1-6, 2011.

Tang, D., Jin Chang, H., Tejani, A. and Kim, T. K., Latent regression forest: Structured estimation of 3d articulated hand posture. In *Proceedings of the IEEE conference on computer vision and pattern recognition,* pp. 3786-3793, 2014.

Tome, D., Russell, C., and Agapito, L., Lifting from the deep: Convolutional 3d pose estimation from a single image. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2500-2509, 2017.

Tseng, M. M., Jiao, J., and Merchant, M. E., Design for mass customization. *CIRP Annals*, vol. 45, no. 1, pp. 153-156, 1996.

Wang, P., Liu, H., Wang, L. and Gao, R. X., Deep learning-based human motion recognition for predictive context-aware human-robot collaboration. *CIRP Annals*, vol. 67, no. 1, pp. 17-20, 2018.

Wang, L., Xiong, Y., Wang, Z., Qiao, Y., Lin, D., Tang, X. and Van Gool, L., Temporal segment networks for action recognition in videos. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 11, pp. 2740-2755, 2018.

Weinrich, C., Wengefeld, T., Schroeter, C. and Gross, H., People detection and distinction of their walking aids in 2D laser range data based on generic distance-invariant features. P*roceedings of the 23rd IEEE International Symposium on Robot and Human Interactive Communication*, pp. 767-773, 2014.

Woschank, M., Rauch, E., & Zsifkovits, H., A Review of Further Directions for Artificial Intelligence, Machine Learning, and Deep Learning in Smart Logistics. *Sustainability*, vol. 12, no. 9, pp. 3760, 2020.

Zanchettin, A. M., and Rocco, P., Probabilistic inference of human arm reaching target for effective human-robot collaboration. In *IEEE/RSJ International Conference on Intelligent Robots and Systems* (IROS), pp. 6595-6600, 2017.

Zanchettin, A. M., Casalino, A., Piroddi, L. and Rocco, P. Prediction of human activity patterns for human–robot collaborative assembly tasks. *IEEE Transactions on Industrial Informatics*, vol. 15, no. 7, pp. 3934-3942, 2018.

Biographies

Manuel A. Ruiz Garcia is postdoctoral research fellow at the Smart Mini Factory lab for Industry 4.0 at the research group Industrial Engineering and Automation (IEA) of the Free University of Bolzano (Italy). His research interests are in computer vision, robotic control, machine learning and natural language processing.

Erwin Rauch is an Assistant Professor in Manufacturing Technology and Systems at Free University of Bolzano, Italy. He received his BSc in Logistics and Production Engineering from the Free University in Bolzano (Italy). He also holds an MSc in Mechanical Engineering from the Technical University Munich (TUM) and a Master in Business Administration from the TUM Business School. Since 2007, he worked in a Management Consultancy as a Consultant for Lean management. Later he obtained his PhD in Mechanical Engineering at the University of Stuttgart in Germany. Since 2014, he is Assistant Professor at the Free University of Bolzano and in 2017 Visiting Researcher at Worcester Polytechnic Institute (WPI) in Worcester, MA (USA). Further, he is Head of the Smart Mini Factory laboratory for Industry 4.0 application at the Free University of Bolzano. His research interests include industry 4.0, axiomatic design, agile and reconfigurable manufacturing systems as well as optimization in SMEs.

Damiana Salvalai is research fellow at the Smart Mini Factory lab for Industry 4.0 at the research group Industrial Engineering and Automation (IEA) of the Free University of Bolzano (Italy). Her research interests are in artificial intelligence, deep learning and their application in manufacturing.

Dominik T. Matt is a Full Professor for Manufacturing Technology and Systems at the Free University of Bolzano and Head of the Industrial Engineering and Automation (IEA) research group. He is also Head of Fraunhofer Italia IEC (Innovation Engineering Center) in Bolzano/Italy. He studied Mechanical Engineering at the Technical University of Munich (TUM) and achieved a Ph.D. in Industrial Engineering at the Karlsruhe Institute of Technology (KIT). From 1998 he worked as research project manager for a US company and for the BMW Group in Munich. In 2004, he was appointed to the post of a Professor for Manufacturing Systems and Technology at the Polytechnic University of Turin, Italy. In 2008, he accepted a call of the Free University of Bolzano. Since 2010, Professor Matt holds the Chair of Manufacturing Engineering at the Free University of Bozen-Bolzano and is the Head of Fraunhofer Italia IEC. He is member in national and international scientific organizations and published more than 100 papers and book contributions.