Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 106 (2022) 102–107

www.elsevier.com/locate/procedia

9th CIRP Conference on Assembly Technology and Systems

Operator training framework for hybrid environments: An Augmented Reality module using machine learning object recognition

George Apostolopoulos[a], Dionisis Andronas[a], Nikos Fourtakas[a], Sotiris Makris[a,*]

[a]*Laboratory for Manufacturing Systems and Automation, Department of Mechanical Engineering and Aeronautics, University of Patras, Patras 26504, Greece* *
*Corresponding author. Tel.: +30-261-091-0160; fax: +30-261-099-7314. E-mail address: makris@lms.mech.upatras.gr

Abstract

As market demands are characterized by more customized products with shorter lifecycles, it is obligatory for modern operators to manage recurrent product or manufacturing system changes. In contrary to previous years, adaptation to such changes prerequires memorization of more information, and familiarization with more complex systems and resources in a shorter period of time. This manuscript presents a novel operator training framework based on Augmented Reality (AR) technology. More specifically, intuitive instructions enhanced with machine learning-based physical object detection are used for making steeper learning curves and providing hands-on experience to operators. The implemented application also supports a walkthrough mode where users can get familiarized with Information and Communication Technologies (ICT) data streams besides fenceless Human-Robot coexistence in collaborative schemes. An automotive case study is used for evaluating the performance of the training framework through a Human-Robot Collaboration (HRC) assembly scenario.

Keywords: Augmented Reality, Machine learning, Human Centered Design, Training, Assembly, Human Robot Collaboration

1. Introduction

In recent years, a balance between productivity, cost, quality and most importantly flexibility is attained by industry in order to satisfy modern market requirements [1]. Competitiveness and smaller product lifecycles are stressing manufacturing systems that are shifting from mass production towards mass customization. Those demands are covered by Industry 4.0 and Internet of Things (IoT) services where smart systems are able to adapt to dynamic production changes through knowledge-based schemes [2].

In this context, significant changes are observed between present operators' workspaces and the traditional ones. Currently, workers are expected to deal with assembly or manufacturing operations of multiple products during their shift. Moreover, the frequency that they are introduced to new assembly routines is increased due to shorter product lifecycles. However, environment's complexity proves to be the most integral aspect as operators need to use or work close to sophisticated tools, machinery, cooperative robots, mobile manipulators, Human Machine Interfaces (HMIs), etc. [3]. In such systems, it is obligatory that operators follow process planning, execution and monitoring procedures since their improvisation skills could lead to potential system stop, loss of data, quality and Key Performance Indicators (KPIs) fluctuations or even safety issues [4].

Despite the complexity of the new industrial paradigm as well as the necessity of quick operator adaptation, training procedures are mostly limited to traditional means such as operation manuals, seminars, workshops, etc. Workers have to adapt to production changes and information streams relying on counterintuitive training material besides their institute knowledge. In [5], it is indicated that swift and immense growth of the Industry 4.0 does not permit businesses to establish training schemes in advance. Moreover, one of the main difficulties when companies implement training activities is the

10.1016/j.procir.2022.02.162

limited or even the lack of operators' interest on participating in such activities. Subsequently, advances in technology must be used for delivering pioneer training methods that ensure high operator interest and well-being together with steeper learning curves for shorter training periods. In parallel, those methods need to familiarize operators with huge information streams, sensors and resources of modern Cyber-Physical Systems (CPS) of IoT era.

This manuscript presents a novel operator training framework based on Augmented Reality (AR) technology. An application that supports multiple modals for prompt interaction and control is implemented aiming to familiarize operators with hybrid systems, IoT technologies and new assembly operations. The overall structure of this study takes the form of six sections, including this introductory section. Section two lays down the challenges, outcomes, limitations, and conclusions of previous studies. Section three analyses the architecture and methodology of the proposed training tools. Section four continues with the implementation of software modules and features of the multi-modal training framework, whereas section five presents a case study that is used for evaluating the framework's performance. Finally, sections six and seven discuss about the results as well as conclusions and future steps of this work.

2. Relevant work

Operators in manufacturing are mostly involved in manual assembly operations due to their improvisation, dexterity and flexibility. With Industry 4.0 paradigm and the establishment of hybrid-assembly workstations, the position of operators has to be improved since collaborative systems' successful implementation is mostly depended on operator's safety and acceptance [6]. Acceptance is compromised by heavy information streams and customization of products leading to serious physical and cognitive workload [7].

For this reason, previous studies have focused on technologies for increasing attractiveness and performance of training, namely via: a) simulation [8], b) Virtual Reality (VR) [9], and c) Augmented Reality (AR) [10]. AR is a technology that digital items are overlayed on the physical world enriching the information that users can get inside their point of view [3]. In this regard, training methods using AR can have a serious head start as operators gain experience directly in a physical (or their working) environment. Moreover, when smart glasses that support voice and gesture-based control are used, hands-free working is feasible even during production [11]. This is ideal for Training on the Job (ToJ). ToJ means that employees are gaining competence during their shift because information is provided right at the workplace and there is no longer necessity of searching training documents and material away from the production line [12]. Findings indicate that AR tools can surpass training modules based on conventional screens despite ergonomic handicaps [13].

In this regard, a considerable amount of literature has focused on the development of AR applications for operator training and support. Overlayed textual annotations, frames and pointing arrows are used for increasing worker's understanding of the manufacturing process [14]. In cases of human robot co-existence, additional features such as robot trajectories, voice commands, notifications and gestures can also be used for greater worker well-being and adaptability in fenceless environments [15]. Despite benefits, wide industrial implementation will be achieved after headset manufactures address limitations in the field of view, long-term wearability and interaction methods [16].

Previous works also focus on providing overlaid instructions on tracked physical objects. This presents challenges on its own due to computational performance of wearable devices, visual sensor quality, dynamic scenes, etc. This can be overpassed by static monitoring devices and edge image processing [17]. Focusing on image processing, the identification of objects in regions of frames has driven academia and industry in the development of numerous methods that can be classified into "traditional object recognition methods" and "deep learning-based detection methods" [18]. Former methods are favored for object detection towards operator support, with a lot of applications being based on SSD [19], YOLO [20], R-CNN [21].

3. Approach

The proposed methodology aims on providing an intuitive training framework that facilitates familiarization of both amateur and experienced workers with a) Industry 4.0 workstations and more specifically hybrid workstations and b) frequent product and assembly procedure updates. Moreover, it is integral that implemented framework stimulates worker's interest through proper communication of information and easy controls. Last but not least, training should be feasible in the physical workstation, even during production itself, in order to establish ToJ and steeper learning curves.

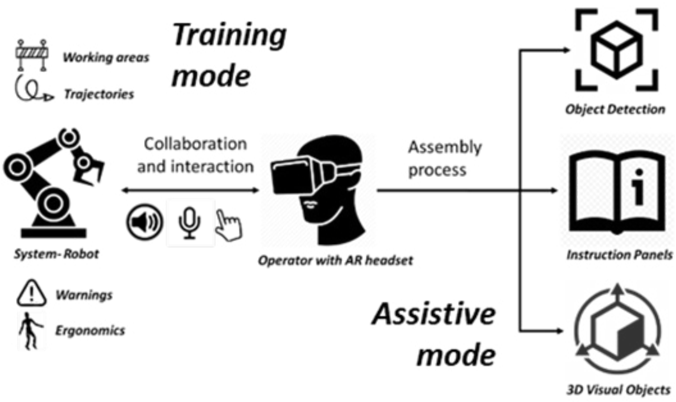

Fig. 1. Training framework's modes and main features

The information that an operator has to compile for keeping up with modern industrial paradigm is huge and stresses adaptation and memorization procedures. In this regard, the proposed training framework distinguishes training sessions into two modes (Fig. 1). The first mode is called "training mode" and focuses on introducing the operator in hybrid workstations whereas the "assistive mode" elaborates on assembly instructions for new products. The methodologies of each training mode as well as important design parameters are analyzed on following subsections.

3.1. Environment design

In order to make training more effective, a "gamification" of the procedure is attained. Training steps and visuals are designed for giving a playful perspective. Previous studies indicate that this is a technique that motivates people on reaching predefined goals and subsequently knowledge and skill transmission is greater [22]. Moreover, the level of detail that operations or systems are described should balance the skills of the user in combination with the difficulty that each training step has. There is a fine line or "flow channel" where this balance is maintained, and any deviation from it can lead to anxiety or boredom [23].

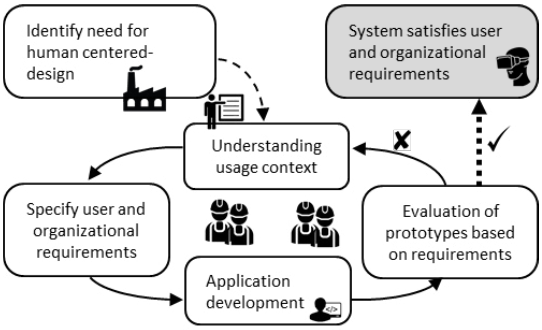

Fig. 2. User-centric designing approach

The whole design process (Fig. 2) was based on the widely used "User-Centered Design" (UCD) method [24]. The objective was the implementation of a training module that is highly usable by users. In this regard, potential users were involved in the design process aiming to: a) identify important features that should be included in the application and b) provide feedback about the accomplishments of design goals. The users originate from two "samples" and provided important feedback after the release of each prototype version. The first group of users are aware about the assembly process of the product, but they are unexperienced on collaborative systems, whereas the second group of users is exactly the opposite. During each "design loop" both teams specified gaps, contributed to solutions and performed evaluation from their perspective until all "what an unexperienced user would need" and "what as unexperienced user I need" constraints are fulfilled. Feedback was provided either through internal questionnaires or though sessions with developers based on point of view recordings.

3.2. Training mode for Industry 4.0 hybrid workstations

Fenceless coexistence with robots emerges a series of challenges related to safety and synchronization that need to be addressed. Effective and safe collaboration prerequires that the intentions and actions of each cooperative resource (i.e. robot, human) are known from the other one(s). Subsequently, advanced planning systems, Human Machine Interfaces (HMIs) and monitoring modules are deployed. The performance of those systems relies on their correct usage from the operator, proper following of monitoring procedures and most importantly rapid compile of CPS information streams by the worker.

A state-of-the art hybrid workstation consists of HMIs that propagate to the operators: a) tasks, b) instructions and c) robot status information. Some pioneer implementations also involve visualization of robot trajectories, safety zones and performance metrics. On the other hand, operators must report their actions through HMIs, physical buttons, or specialized tool operation. The exchange of information in those systems can take place through different modals (e.g. visuals, holograms, audio, gestures, voice, touch, etc.) via different devices (e.g. screens, keyboards, buttons, AR glasses, smartwatches, projectors, tablets, etc.) [25]. The volume of data and the number of devices or tools that worker has to interact with could confuse operators and reduce hybrid system acceptance in total.

The "training mode" of the application involves a "walk-through" session where all functionalities of the HMIs and resources are analyzed sequentially. The operator can remotely or physically navigate in the workstation and receive augmented instructions about available resources and controls. Moreover, the system can overlay digital items about potential events in the production (e.g. safety or security incidents, customized product assemblies, etc.) and analyze the information that workers would receive. The types of HMIs, resources as well the design of the workstations themselves is out of the scope of this manuscript.

3.3. Assistive mode for assembly operation training

As discussed, operators are challenged to memorize assembly operations of multiple products frequently. In contrast with traditional training material, the proposed application aims on providing an intuitive environment that users can get familiarized with joining operations between various components. The overall objective is to provide tools where trained persons can interact with the components' virtual copies and unveil important parameters for assembly before gaining "hands-on" experience. In difference with operation manuals, the holograms can effortless be manipulated with gestures and operators can see in first-sight details from CAD models. Besides components, the holograms need also to contain relative workpiece movements and assistive notifications.

An aspect that most training applications or materials lack is interaction with physical components. It is possible that operators are confusing assembly related components and loose time until they identify proper ones, and they are certain of their selection. In this regard, the "assistive mode" must support physical object recognition and provide augmented identification and instructions based on the physical component's positioning. For ensuring setup easiness of the application, it is attained that object detection uses AR headset's vision sensor. Given operator's head movements, random positioning of components and dynamic environment conditions, a machine learning object recognition algorithm should be used.

4. Implementation

Aiming greater industrial implementation, one of the most important parameters is the selection of the host device. The attained controlling and communication modals, hardware specifications, autonomy, ergonomics and field of view contributed in the selection of Microsoft HoloLens 2 as the most appropriate device. This headset supports the functionalities that were defined in the early design process for both training modes. For facilitating the reader, there is one sub-section per mode where the deployment of related features is analyzed.

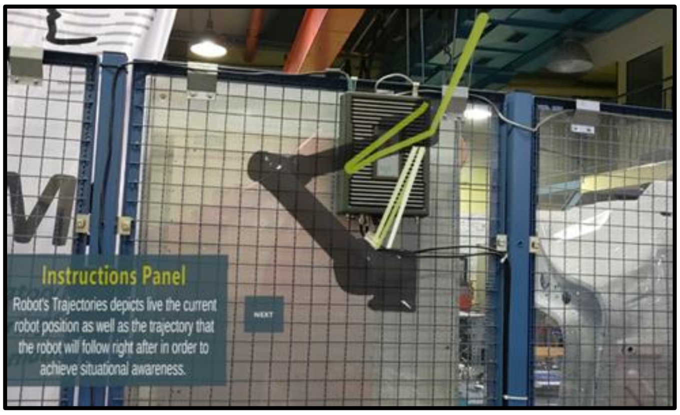

Fig. 3. Training mode introducing robot resources and visualization of trajectories

4.1. Training mode for Industry 4.0 hybrid workstations

This mode recreates a walkthrough experience where the operator is informed about specific aspects of the collaborative workstation on each step. As soon as the operator is confident that provided insights are comprehended, a "next" button can be used for loading the next instruction scene. The training session (refer to Fig. 3 and Fig. 4) elaborates on a) the resources that the workers are going to interact with during their shift, and b) the controls and the provided information of state-of-the art multi-modal interfaces [15]. Extensive description of those interfaces is out of the scope of this manuscript, however they included features that can be summarized as:

- operator registration and login
- customization of interface environment
- robot trajectories and status
- real-time ergonomic assessment results
- task information and assembly instructions
- safety zone visualization
- physical object recognition
- cyber-security incident notifications
- audio alerts and warning notifications
- physical, touch or augmented reality buttons
- voice commands and gesture control

Fig. 4. Training mode introducing a) ergonomic analysis results and b) assembly instruction panels

4.2. Assistive mode for assembly operation training

The second mode is oriented on product assembly training. Like the first mode, the worker is informed about the product assembly through a step-by-step walkthrough experience. Having as basis the related bill of processes, the operator is instructed about the assembly operation sequence and related workpieces (Fig. 5). On each training step, information is provided in the form of augmented: a) textual instructions, b) 3D workpiece visualization, c) arrows and signs, d) workpiece relative movements, e) physical object identification.

Fig. 5. Assistive mode representing assembly process through augmented components

For both modes, Unity Game engine was used for application development. Mixed Reality Toolkit supported the implementation of most augmented items and controls. Physical object detection was performed using OpenCV (Open-Source Computer Vision Library). It was selected as a widespread open-source software library where multiple machine learning applications have been released. The detection method itself is based on Paul Viola et. al [26] machine learning approach for visual object detection. This approach is capable of processing images extremely rapidly and achieve high detection rates.

The performance of this method relies on the quality of the used dataset and the quality of classifiers training. A variant of AdaBoost is used to select a small set of features and train the classifier. In order to achieve robust results of cascade training, thousands of images are captured per object. Objects with complex geometries or multiple poses of equilibrium on a plane require significantly more images. Given that the detection sensor is integrated in the AR headset, the dataset should include multiple figures of objects from different angles, point of views or distances as well as lighting conditions. Moreover, for detecting an object wherever it is, the method must be repeated for different backgrounds. The next step involves manual sorting of negative and positive figures, with the former ones requiring further processing. More specifically, for each positive figure, a square must be placed around the object of interest. Finally, cascade training involves the comparison of positive and negative images (Fig. 6).

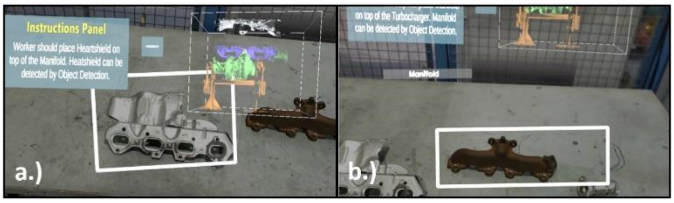

Fig. 6. Assistive mode physical object tracking during training

5. Case study and experimental setup

The training application was evaluated on a use case originating from automotive industry. A hybrid manufacturing cell deals with the assembly of a powertrain component and more specifically a turbocharger. Tasks are dispatched between operators and robots that are orchestrated in a fenceless environment (refer to [27]). Despite improvements in quality, cycle time and ergonomics, operators need to adapt to a collaborative environment where robots manipulate tools and work on the same workpiece. On the other hand, the assembly operation itself is challenging due to the complex geometries of related components. In this context, a new worker would struggle to get familiarized with a) the assembly operation and b) multi-modal interface controls and information streams that aim increased awareness and safety. Subsequently, this cell is a fertile ground for evaluating the performance of the training framework's two modes.

Fig. 7. Subjects training on a) specialized environment and b) on production line-relevant environment

When the application is started, the user selects between the two described modes. Focusing on the "training mode" a series of scenes were implemented for describing how the user can customize the HMIs, how to use them, what resources are available and what kind of information is communicated during production. In the same basis, panels and augmented items are used for analyzing assembly procedure and parameters within the "assistive mode". The main challenge of the second mode is the deployment of physical object recognition and augmented notifications on tracked objects. The related components required several thousand of photos for creating reliable datasets due to the complex geometries and reflecting surfaces.

6. Results

A series of tests were conducted with subjects originating from two different "experience groups" equally split (refer to Section 3.1). The objective of those tests was to estimate overall training framework's performance, expose handicaps and identify features that next training application prototype should include for greater training intuitiveness. All persons participated in training sessions by using both training modes and they provided evaluation feedback. Questionnaires, based on System Usability Scale (SUS) [28], were shared to a sample of eight users. Having a total score of 82.19%, the proposed training framework proves to be an effective mean of training for experienced or non-experienced users.

Table 1. SUS Analysis Results

User	1	2	3	4	5	6	7	8
Score	85	85	70	82.5	85	90	67.5	92.5
Total mean score		82.1						

In a more qualitative way, the users expressed that both training modes were very intuitive and easy to use. The enhancement of traditional training material into augmented items, that operators can interact with, helped users to get familiarized with collaborative systems and most importantly with CPS information streams. More specifically, 62% of individuals uttered that getting used to a hybrid system during actual production process would be very stressing for them. The capability of training in a physical environment with augmented reality items was one of the biggest advantages of the framework as operators didn't experienced any cognitive workload. This is because AR instead of Virtual reality provides digital information only when it really maters increasing concertation during training. Moreover, the interaction with physical workpieces and resources during training (refer to Fig. 7) was helpful especially with augmented notifications on physical objects of interest. Focusing on learning-based object recognition, the algorithm demonstrated rapid tracking of parts, however accuracy was achieved after increasing the size of used datasets for training. When accuracy levels were low false positives and tracking of false or inexistant objects could confuse users.

7. Conclusions and future steps

The proposed training framework proved that Augmented Reality is a technology that can enhance training practices. Either for familiarization with Industry 4.0 systems or new assembly operations, augmented reality can lead to steeper learning curves and faster operator adaptation. Testing proved that is due to the intuitive representation of training information as well as intuitive controls that the proposed application offers.

Augmentation of physical world instead of navigation into a virtual environment led to limited cognitive and physical overload and gave users the confidence that they comprehended all training aspects. Training in the robotic cell itself or a relevant environment as well as interaction with real assembly parts facilitated end-to-end understanding of the assembly operation. Augmented a) textual information, b) CADs with joining relative movements, c) notifications on real parts in combination with hands-free controls made the application appealing and users could reuse it if any training procedure is required. The main handicap of the application is the preparation time that is required in case a new product needs to be introduced in the training session. Augmented visuals or object detection requires significant preparation time from developers. In this regard, future work would involve connectivity of the application with CAD/ CAM systems for reducing releasing periods between application versions. Focusing on the human-centric architecture of the system, "gamification" could be increased by integrating user profiles and an experience metric system based on points.

Such training frameworks would be valuable for businesses as operators can be adapted rapidly to production system changes and job openness can be increased. Industry could benefit as training periods are reduced and workers can keep up with production and customization demands.

8. Acknowledgments

This work has been partially funded by the EC research project "SHAREWORK-Safe and effective HumAn-Robot coopEration toWards a better cOmpetiveness on cuRrent automation lacK manufacturing processes" (Grant Agreement: 820807).

References

[1] G. Chryssolouris, Manufacturing Systems: Theory and Practice, 2006.
[2] D. Mourtzis, S. Fotia, N. Boli, P. Pittaro, Product-service system (PSS) complexity metrics within mass customization and Industry 4.0 environment, International Journal of Advanced Manufacturing Technology. (2018).
[3] S. Makris, Cooperating Robots for Flexible Manufacturing, Springer International Publishing, Cham, 2021. https://doi.org/10.1007/978-3-030-51591-1.
[4] S. Monsell, Task switching, Trends in Cognitive Sciences. 7 (2003) 134–140.
[5] T.C. Piñol, S.A. Porta, M.C.R. Arévalo, J. Minguella-Canela, Study of the training needs of industrial companies in the Barcelona Area and proposal of Training Courses and Methodologies to enhance further competitiveness., Procedia Manufacturing. 13 (2017) 1426–1431.
[6] G. Michalos, S. Makris, P. Tsarouchi, T. Guasch, D. Kontovrakis, G. Chryssolouris, Design considerations for safe human-robot collaborative workplaces, in: Procedia CIRP, Elsevier B.V., 2015: pp. 248–253.
[7] D. Romero, P. Bernus, O. Noran, J. Stahre, Å. Fast-Berglund, The Operator 4.0: Human Cyber-Physical Systems & Adaptive Automation Towards Human-Automation Symbiosis Work Systems, IFIP Advances in Information and Communication Technology. 488 (2016) 677–686.
[8] D.M. Oliveira, S.C. Cao, X. Fernández Hermida, F. Martín Rodríguez, Virtual Reality System for Industrial Training, (n.d.).
[9] I. Galvan-Bobadilla, A. Ayala-Garcia, E. Rodriguez-Gallegos, G. Arroyo-Figueroa, Virtual reality training system for the maintenance of underground lines in power distribution system, 2013 3rd International Conference on Innovative Computing Technology, INTECH 2013. (2013) 199–204.
[10] D. Mourtzis, V. Siatras, J. Angelopoulos, N. Panopoulos, An Augmented Reality Collaborative Product Design Cloud-Based Platform in the Context of Learning Factory, Procedia Manufacturing. 45 (2020) 546–551.
[11] J. Peddie, Augmented Reality, (2017).
[12] S.R. Sorko, M. Brunnhofer, Potentials of Augmented Reality in Training, Procedia Manufacturing. 31 (2019) 85–90.
[13] T. Sutar, S. Pawar, Smart Glasses: Digital Assistance in Industry, Lecture Notes in Electrical Engineering. 703 (2021) 169–182.
[14] S. Webel, U. Bockholt, T. Engelke, M. Peveri, M. Olbrich, C. Preusche, Augmented Reality Training for Assembly and Maintenance Skills, BIO Web of Conferences. 1 (2011) 00097.
[15] D. Andronas, G. Apostolopoulos, N. Fourtakas, S. Makris, Multi-modal interfaces for natural Human-Robot Interaction, Procedia Manufacturing. 54 (2021) 197–202.
[16] A. Syberfeldt, O. Danielsson, P. Gustavsson, Augmented Reality Smart Glasses in the Smart Factory: Product Evaluation Guidelines and Review of Available Products, IEEE Access. 5 (2017) 9118–9130.
[17] Z.H. Lai, W. Tao, M.C. Leu, Z. Yin, Smart augmented reality instructional system for mechanical assembly towards worker-centered intelligent manufacturing, Journal of Manufacturing Systems. 55 (2020) 69–81.
[18] X. Zou, A Review of object detection techniques, Proceedings - 2019 International Conference on Smart Grid and Electrical Automation, ICSGEA 2019. (2019) 251–254.
[19] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, A.C. Berg, SSD: Single Shot MultiBox Detector, Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 9905 LNCS (2016) 21–37.
[20] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, You Only Look Once: Unified, Real-Time Object Detection, (2016) 779–788. http://pjreddie.com/yolo/ (accessed September 24, 2021).
[21] S. Ren, K. He, R. Girshick, J. Sun, Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, (n.d.). https://github.com/ (accessed September 24, 2021).
[22] S. Gilotta, S. Spada, L. Ghibaudo, M. Isoardi, A Technology Corner for Operator Training in Manufacturing Tasks, Advances in Intelligent Systems and Computing. 824 (2018) 935–943.
[23] Mihaly. Csikszentmihalyi, Beyond boredom and anxiety, (1975) 231.
[24] L. Gamberini, Human-computer interaction: fondamenti teorici e metodologici per lo studio dell'interazione tra persone e tecnologie, (2012).
[25] P. Kumari, L. Mathew, P. Syal, Increasing trend of wearables and multimodal interface for human activity monitoring: A review, Biosensors and Bioelectronics. 90 (2017) 298–307.
[26] P. Viola, M. Jones, Rapid object detection using a boosted cascade of simple features, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001.
[27] D. Andronas, A. Argyrou, K. Fourtakas, P. Paraskevopoulos, S. Makris, Design of Human Robot Collaboration workstations - Two automotive case studies, in: Procedia Manufacturing, Elsevier B.V., 2020: pp. 283–288.
[28] J. Brooke, SUS-A quick and dirty usability scale, n.d.