Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 54-1 (2021) 540–545

Indoor positioning, artificial intelligence and digital twins for enhanced robotics safety

Tomi Pitkäaho, Tero Kaarlela, Sakari Pieskä, Sami Sarlin

Centria University of Applied Sciences, Vierimaantie 7, 84100 Ylivieska, Finland, (e-mail: tomi.pitkaaho@ centria.fi).

Abstract: Flexible robotics safety solutions allowing the implementation of fenceless robot cells are becoming a reality nowadays. Safety approved sensors such as light curtains, safety scanners, and safety cameras have been deployed already successfully in various industrial robotic solutions. Still, as these safety systems are installed in fixed locations, monitoring predefined regions, the systems can be rigid and inflexible. This paper introduces a novel hybrid safety solution. The solution comprises safety-approved sensors, additional sensors, and artificial intelligence analysis. The system increases flexibility, especially in cases where collaborating humans and robots need monitoring in larger areas. Typically, in such environments, work objects are large and heavy, introducing additional challenges. In addition, the proposed system includes a digital twin implementation that allows a connection between the real and virtual worlds. Already virtual models and robot simulation have been used for designing safe robot applications. However, the efficient use of digital twins in safety planning and safety monitoring is still uncommon.

Keywords: Robotics technology, Perception and sensing, Positioning systems, Machine learning, Intelligent interfaces

1. INTRODUCTION

The utilization of robotics in industrial environments is highly regulated by standards such as ISO 10218 [ISO (2011)] and legislation. The human safety within robotic applications is ensured by following the ISO 12100 standard [ISO (2010)], and the actual safety-approved sensors that can be used to implement such systems, need to fulfill the ISO 13849 [ISO (2016a, 2012)] or IEC 61508 [IEC (2010a,b)] standards. These strict and somewhat conservative regulations might prevent the wider utilization of robots. As a consequence, traditionally, industrial robots are isolated behind fences to prevent human injuries.

A true human-robot collaboration (HRC) enables a combination of human skills and robot qualities. When implemented appropriately, it is expected to allow dynamic, more efficient performance leading to increased productivity than either one alone could achieve [see for example Michalos et al. (2015); Robla-Gómez et al. (2017)]. However gaining the true HRC is not easy to achieve, at least partially due to conventional safety requirements, that restrict dynamic and efficient collaboration between the robots and people. Fortunately, the new safety requirements have been taken into account in the updated ISO 10218 standards. In addition, ISO/TS 15066:2016 [ISO (2016b)] is solely aimed for collaborative robotics (cobots). Cobots with limited payloads have been implemented successfully in many manufacturing tasks such as pick-and-place operations, packing and palletizing, machine tending, welding, screw driving, and quality inspection.

In this specification, the robot safety modes are defined as follows: hand guiding, speed and separation monitoring (SSM), safety-rated monitored stop, and power and force limiting. Nevertheless, safety assurance can be considered as the most critical requirement for cobots as stated by Bi et al. (2020).

In this paper, an additional safety system is introduced. These components and the system are not safety-approved. Instead, these components can be used together with safety-approved sensors to increase safety and monitoring capabilities. The system consists of a Bluetooth Low Energy (BLE) indoor positioning system (IPS), digital twin (DT), and a combination of a 360-degree camera and machine learning (ML) for detecting people. Additionally, the implementation of a DT can be used for safety training, risk analysis, and accident analysis. The overall system is shown in Fig. 1.

2. RELATED RESEARCH

In literature, safety aspects have been reported widely. A review by Halme et al. (2018) is a comprehensive review of approaches for implementation of safe HRC focusing on vision-based safety systems. Only a limited number of them have been used successfully in HRC. A survey for methods for safe human-robot interaction (HRI) was conducted by Lasota et al. (2017). Bi et al. (2020) performed a survey of cobots and especially safety-related to them. As another example of HRC being an activate research area, a software tool for designing safe HRC was implemented by Saenz et al. (2020).

10.1016/j.ifacol.2021.08.062

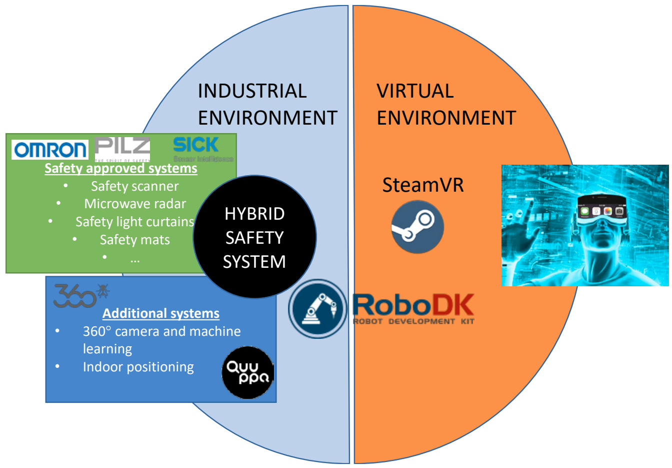

Fig. 1. Overall system. Industrial environment contains safety approved and additional system that together are used for increased safety and monitoring. Indoor positioning system (Quuppa) provides position of people and mobile robots to RoboDK that acts as a link between real and virtual worlds. SteamVR is used for displaying the real world in VR.

If the relative position of humans and robots need to be monitored and measured for a safe operation, vision-based methods are a natural choice. Perhaps, one of the most common computer vision-based safety systems is SafetyEYE (Pilz GmbH & CO. KG, Ostfildern, Germany) that can be installed on a robot cell where the system can monitor multiple regions simultaneously. A human entering an unsafe region inside the cell leads to a predefined action of the robot; the movement speed can be reduced or the robot's motion can be completely halted. Another vision safety system based on a combination of lidar and RGB-camera was recently reported by Rashid et al. (2020).

Besides vision-based methods, other approaches such as microwave radars (e.g. safeRS from SICK AG, Waldkirch, Germany) and more conventional safety solutions like light curtains and laser scanners are commercially available as off-the-shelf solutions.

To allow more flexible HRC, a novel safety monitoring system was introduced by Pieskä et al. (2020). This system allows efficient and safe human-robot cooperation and is especially relevant for industrial robotic applications where large regions need to be monitored. Another paper by Magrini et al. (2020) introduced a layered control architecture with additional sensors.

As one novelty of the proposed approach reported in this paper, the system provides feedback from the real environment to a digital model of the environment. This additional component (DT) can be used for risk-free training of new employees and visitors as reported by Kaarlela et al. (2020). Demonstrations of emergency stop function, safety area violation, and procedures on emergency scenarios such as co-worker injury can be safely carried out. DTs can be utilized to increase the level of safety awareness of workers and this in turn can prevent accidents.

The next three sections contain a short introduction to the key technologies of the proposed approach.

2.1 Indoor positioning

Global Positioning System (GPS), Global Navigation Satellite System (GLONASS), and Galileo are widely acknowledged positioning systems. However, these systems work poorly inside built environments. For this reason, different indoor positioning technologies were invented. Such systems are based on radio waves, acoustic signals, light, or other forms of electromagnetic radiation [Mainetti et al. (2014)].

The Bluetooth 5.1 standard (BLE) includes direction finding feature that has been utilized successfully for centimeter accuracy indoor positioning system [Martin (2019)]. The positioning is based on the detection of the signal's angle of arrival. This technology can be utilized to improve the safety of robotized environments. With an accurate IPS, both mobile robots or people can be tracked in real-time.

2.2 Machine learning

ML is part of artificial intelligence techniques that somehow mimic human problem-solving. One part of ML is deep learning [see for example LeCun et al. (2015)] that has been used successfully in a vast number of different applications. Probably, most popular applications can be found in applications where images are used in one way or another. In this domain, convolutional neural networks (CNNs) have been used successfully for image classification [Krizhevsky et al. (2017)], object detection and localization [Girshick et al. (2014); Redmon et al. (2016)], semantic segmentation [Long et al. (2015)], and image captioning [Chen et al. (2017)] among other applications.

2.3 Digital twin

DTs introduced by Grieves (2014) have the potential for avoiding the laborious phase of artificially creating real-time sensor information from robots, machines, and humans into the virtual world. DTs have a real-time sensor bridge between the real and virtual worlds. The DT sensor information bridges physical qualities such as pose, location, and speed of machines and humans from the production floor to the virtual world[Malik and Bilberg (2018); Kousi et al. (2019)]. When combined with virtual reality (VR), DTs can elevate the level of VR to a more realistic level. Safety planning, risk analysis, and safety training conducted in a DT are very realistic as movements and functionalities of machines, robots, and humans in the training environment correspond to the real world. Because DTs include real-time sensor data gathering, there is no need for the artificial creation of movement and functionality of machines or humans.

Furthermore, it is possible to store sensor information gathered from machines and human tracking devices into a database. This stored information can be later, if needed, utilized for replaying scenarios such as accidents or otherwise dangerous situations. The stored information has a true potential to be a powerful tool for accident analyses and for creating training scenarios of emergencies.

The DT also provides a method for safe risk-analysis [Bellalouna (2019)] and safety planning based on real-time data. Engineers planning safety measures can step into the DT and conduct risk-analysis without interfering with the actual production. The DT is a stress and risk-free environment for safety planning, inspection, and risk analysis.

VR has been implemented in safety training already by Sim et al. (2019). It has also been proven as a more efficient training method compared to traditional training methods such as lecturing. Training taking place in VR can be seen as an intuitive, immersive, and interactive way of learning [Le et al. (2015)].

Unfortunately, building blocks for virtual experience are often passive three-dimensional models of entities from a production environment. These three-dimensional models contain physical data but lack functional and movement data of their real-world duplicates. The addition of this missing data to virtual entities requires the artificial creation of movement and functionality for machines, robots, and humans involved. This phase in the creation of virtual experience is usually time-consuming and also requires deep knowledge of game programming. Game programming skills are mandatory as functionality in the virtual environment is created by utilizing game engine such as Unity (2020). Each training scenario can be seen as an individual programming project. Also because the virtual environment is based on imported three-dimensional model files, updating layout or product data usually leads to complete re-creation of training scenarios [Bellalouna (2019)].

3. APPROACH

3.1 Indoor positioning

The used commercially available IPS, based on the BLE technology, was Quuppa Intelligent Locating System™(Quuppa Oy, Espoo, Finland). The positioning system consisted of five locators evenly spread in the laboratory where each locator was calibrated using three distinct points. After the calibration, tracking of spatial (x, y, z) positions was performed in real-time.

3.2 Machine learning

With 360°cameras, a single camera shot contains information of horizontal and vertical directions. A single image is a projection of a scene on a unit sphere. When viewed as a planar image, the image contains relatively strong geometric distortions mainly on both poles. This is not favorable for CNNs, that usually learn rectangular or square convolutional filters during the training process. This challenge poses a difficulty for the transfer learning and already trained networks cannot be used as-is for such data. To overcome this challenge, we used the method developed by Eder et al. (2020) where tangent images are extracted from equirectangular images. This procedure is shown in Fig. 2.

We used 360fly (360fly 4K, 360Fly Inc., Canonsburg, USA) that was installed on the ceiling and a desk (Fig. 3). These two different installations were used to simulate two different monitoring possibilities, monitoring a robot cell (the ceiling installation), and monitoring the immediate surroundings of a robot (the on-desk installation). We captured 360°videos at 2880 × 2880 pixel resolution. Of these videos, spherical image frames were extracted forming a database of 715 tangent images. Each image was 512 × 512 pixels. Images were manually annotated using a browser-based labeling tool Labelbox (Labelbox Inc., San Francisco, USA). The labeling was performed using rectangular bounding boxes. The data was split into training and testing data with 90/10% ratio, and the total number of annotated objects was 4 400.

YOLOv4 by Bochkovskiy et al. (2020) implemented in PyTorch deep learning framework [Paszke et al. (2017)] was used for the training and inference. YOLOv4 is an improvement over YOLOv3 that was originally introduced by Redmon et al. (2016). The training was performed using the Adam optimizer [Kingma and Ba (2014)] over 500 epochs. Data was augmented using mosaic images from four images, random horizontal and vertical flips, and random HSV gain.

3.3 Digital twin

The DT presented here was compiled using three different components:

(1) RoboDK robot simulation software [Mihai (2015)],
(2) Quuppa IPS, and
(3) SteamVR VR platform [Hewlett-Packard (2015)].

The RoboDK robot simulation software was acting as a bridge between sensor information from robots, machines, and humans to the virtual world. The Quuppa IPS provided the location information of humans to RoboDK, and the SteamVR VR platform acted as a bridge of user sensor information to the virtual twin. As the DT allows the bi-directional data stream, the implemented DT can be used for safety monitoring. Physical positions of people are tracked using the indoor positioning system and physical robot dimensions are obtained directly from robots. These two sources of information are combined in the DT and digital signals are used to control robot speed according the spatial distances of people to pre-defined danger zones.

The first step during implementing the DT presented in this paper was the laser scanning of the robotics laboratory at Centria university (Fig. 4). Laser scanning was conducted using the Leica 3D-laser scanner (ScanStation 2, Leica Geosystems AG, St Gallen, Switzerland). The resulting point cloud was transformed into a three-dimensional solid model in order to import the model in RoboDK simulation software. At this point, the DT contained all passive fixed elements such as walls, floor, robot fixtures, and tables. The robot models matching the real world were added in RoboDK from the library and the model of human employees was created utilizing RoboDK tools.

Real-time sensor information from machines and robots was bridged to the virtual twin running on RoboDK. In practice, physical connections between the PC running RoboDK and various robot controllers can be either through serial or ethernet ports. In this case, the ethernet connection was utilized in order to bridge the pose and velocity information from robot controllers to the DT.

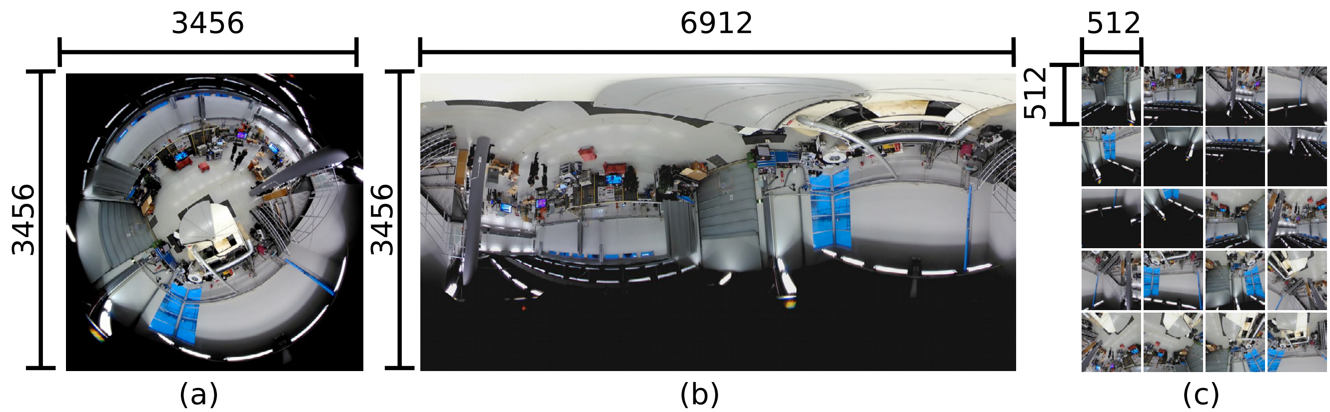

Fig. 2. From spherical image to tangent images. (a) native format of Fly360 camera, (b) spherical image converted to equirectangular format, (c) extracted tangent images. Image sizes are shown besides images.

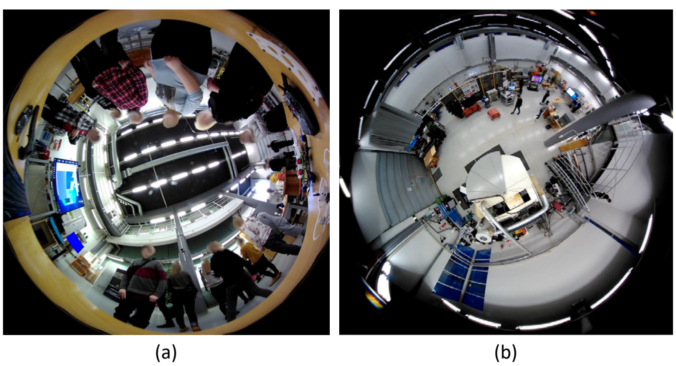

Fig. 3. Spherical images obtained from two different installation locations. (a) on-desk, (b) on ceiling.

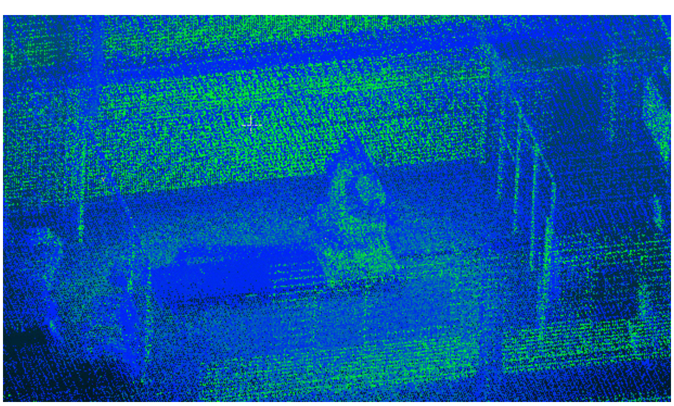

Fig. 4. The scanned point cloud of the robotics laboratory.

In addition, the location information of humans inside the robotics laboratory were bridged to the DT. The Quuppa positioning system provides location information of tags in plain (x, y, z) coordinates through the ethernet connection. The Python application programming interface on RoboDK was utilized in order to retrieve this information from the server and position three-dimensional virtual models of human workers.

At this point, all needed information of robotics laboratory entities was bridged to the virtual world. The DT implementation described here provides possibility for visual inspection and interaction by means of traditional human interface devices such as mouse or touchpad. However, it does not provide immersive or intuitive way of stepping into the virtual environment.

In order to provide a more realistic and intuitive way for safety training and analysis, a connection to the VR environment was added. SteamVR is a VR platform enabling immersive, interactive and intuitive user experience. RoboDK supports publishing a virtual twin on the SteamVR platform. Once published, users with VR-headsets and hand held controllers can immerse to the virtual twin in a very realistic way. Sensors inside headsets and controllers provide user interactions with the virtual twin.

4. EXPERIMENTAL RESULTS

Fig. 5 shows training and validation losses together with performance metrics calculated on the validation data. Based on the learning curves, the training could have been stopped earlier. On the other hand, there is still a minor improvement in the precision after 400 epochs as shown in Fig. 5b. Overall, there are no obvious signs of overfitting and the model should generalize well.

Fig. 6 shows inference results where people inside the robotics laboratory are detected using tangent images extracted from spherical images obtained from two different camera installation locations. It can be observed that the trained model performs well within this challenging environment. The model scales well with different amounts of people and learned to detect individual people inside groups even if they are partially blocked by other people or obstacles.

Fig. 7 presents conduction of risk-analysis with the DT of a robotic production cell. Another use case for using the DT is related to training. Before entering the robotics laboratory at Centria, safety training is required. The DT implemented for Centria robotics laboratory enables a risk-free method of safety training for visitors. The first step to start the training is to wear VR-headset and handheld controllers. During the training, the instructor presents the safety zones and functionalities of each safety device to trainees. Trainees can examine all robot cells and study safety features without the risk of physical injury. Trainees can communicate with the instructor and other

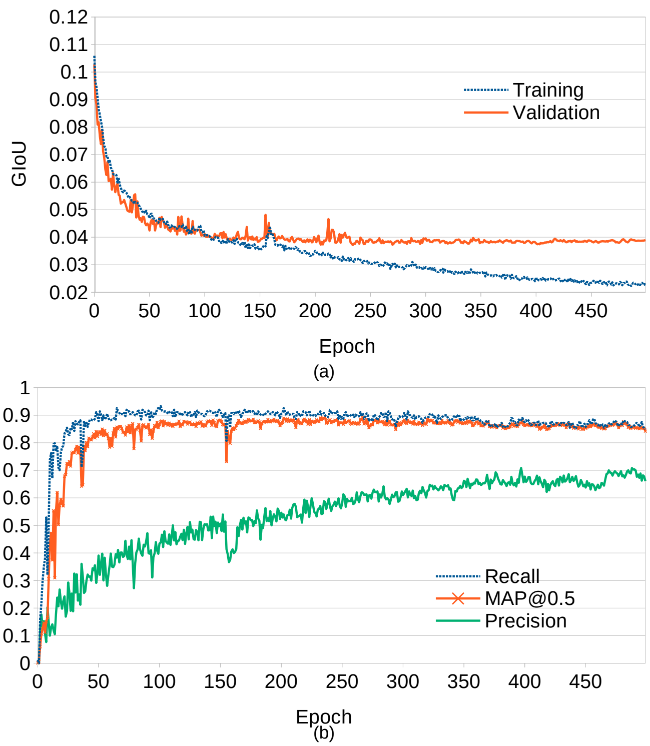

Fig. 5. Training result. (a) Training and validation loss (generalized intersection over union), (b) recall, precision, and mean average precision at 0.5 confidence level (mAP@0.5).

Fig. 6. CNN inference examples. (a) original tangent image extracted from the ceiling installed camera, (b) inference result of (a), (c) original tangent image extracted from the on-desk installed camera, (b) inference result of (b).

trainees through built-in microphones and speakers on VR-headsets. After the training, participants are familiar with safe zones, the functionality of safety devices, and operation of robot cells.

Besides, the DT enables the examination of safe distances and possible hazards for educational purposes. As the digital information can be stored, another thought scenario is to use the DT for analysis of accidents or other events

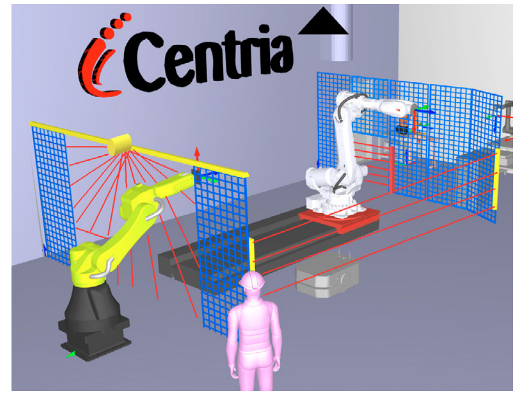

Fig. 7. DT as a risk-analysis tool

that need to be studied afterward. With these features, a DT is a powerful tool for conducting risk-analysis and examining the safety qualities of a robot cell.

5. CONCLUSION

The current trend in industrial robotics applications is to move robots outside fences. That is indeed possible by following existing safety standards. However, more flexible systems are still required. In this paper, an enhanced robotic safety system consisting of additional sensors and artificial intelligence was presented. The proposed system improves the flexibility of the safety-approved systems and allows monitoring of larger regions in industrial environments.

Implementation of a DT can be an effective method in safety training, risk analysis, and safety planning. The DT is capable of providing a more realistic safety training experience compared to VR. Furthermore utilizing DTs eliminates the time, skill, and labor for creating functionality and movement of entities with game programming tools.

ACKNOWLEDGEMENTS

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement n° 825196/.

REFERENCES

Bellalouna, F. (2019). Virtual-reality-based approach for cognitive design-review and fmea in the industrial and manufacturing engineering. In *2019 10th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, 41–46. IEEE.

Bi, Z., Luo, M., Miao, Z., Zhang, B., Zhang, W., and Wang, L. (2020). Safety assurance mechanisms of collaborative robotic systems in manufacturing. *Robotics and Computer-Integrated Manufacturing*, 67, 102022.

Bochkovskiy, A., Wang, C.Y., and Liao, H.Y.M. (2020). YOLOv4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*.

Chen, L., Zhang, H., Xiao, J., Nie, L., Shao, J., Liu, W., and Chua, T.S. (2017). Sca-cnn: Spatial and channel-wise attention in convolutional networks for image captioning. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 5659–5667.

Eder, M., Shvets, M., Lim, J., and Frahm, J.M. (2020). Tangent images for mitigating spherical distortion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 12426–12434.

Girshick, R., Donahue, J., Darrell, T., and Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 580–587.

Grieves, M. (2014). Digital twin: manufacturing excellence through virtual factory replication. *White paper*, 1, 1–7.

Halme, R.J., Lanz, M., Kämäräinen, J., Pieters, R., Latokartano, J., and Hietanen, A. (2018). Review of vision-based safety systems for human-robot collaboration. *Procedia CIRP*, 72, 111–116.

Hewlett-Packard (2015). LEARN MORE : WINDOWS MIXED REALITY PLATFORM + STEAMVR. URL `https://www8.hp.com/h20195/v2/GetPDF.aspx/4AA7-5433ENW.pdf`. Accessed 2 December 2020.

IEC (2010a). Iec 61508-1 functional safety of electrical/electronic/programmable electronic safety-related systems – part 1: General requirements.

IEC (2010b). IEC 61508-2 Functional safety of electrical/electronic/programmable electronic safety-related systems – Part 1: Requirements for electrical/electronic/programmable electronic safety-related systems.

ISO (2010). ISO 12100: Safety of machinery–General principles for design–Risk assessment and risk reduction (ISO 12100: 2010).

ISO (2011). ISO 10218: Robots and Robotic Devices–Safety Requirements for Industrial Robots–Part 2: Robot Systems and Integration.

ISO (2012). 13849-2: Safety of machinery–Safety-related parts of control systems–Part 2: Validation.

ISO (2016a). 13849-1: Safety of machinery–Safety-related parts of control systems–Part 1: General principles for design.

ISO (2016b). Robots and robotic devices - Collaborative robots ISO/TS15066.

Kaarlela, T., Pieskä, S., and Tomi, P. (2020). Digital twin and virtual reality for safety training. In *2020 11th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, 000115–000120. doi: 10.1109/CogInfoCom50765.2020.9237812.

Kingma, D.P. and Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Kousi, N., Gkournelos, C., Aivaliotis, S., Giannoulis, C., Michalos, G., and Makris, S. (2019). Digital twin for adaptation of robots' behavior in flexible robotic assembly lines. *Procedia manufacturing*, 28, 121–126.

Krizhevsky, A., Sutskever, I., and Hinton, G.E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84–90.

Lasota, P.A., Fong, T., and Shah, J.A. (2017). A survey of methods for safe human-robot interaction. *Foundations and Trends in Robotics*, 5, 261–349.

Le, Q.T., Pedro, A., and Park, C.S. (2015). A social virtual reality based construction safety education system for experiential learning. *Journal of Intelligent & Robotic Systems*, 79(3-4), 487–506.

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. doi:https://doi.org:/10.1038/nature14539.

Long, J., Shelhamer, E., and Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 3431–3440.

Magrini, E., Ferraguti, F., Ronga, A.J., Pini, F., De Luca, A., and Leali, F. (2020). Human-robot coexistence and interaction in open industrial cells. *Robotics and Computer-Integrated Manufacturing*, 61, 101846.

Mainetti, L., Patrono, L., and Sergi, I. (2014). A survey on indoor positioning systems. In *2014 22nd international conference on software, telecommunications and computer networks (SoftCOM)*, 111–120. IEEE.

Malik, A.A. and Bilberg, A. (2018). Digital twins of human robot collaboration in a production setting. *Procedia manufacturing*, 17, 278–285.

Martin, W. (2019). Bluetooth direction finding, a technical overview.

Michalos, G., Makris, S., Tsarouchi, P., Guasch, T., Kontovrakis, D., and Chryssolouris, G. (2015). Design considerations for safe human-robot collaborative workplaces. *Procedia CIrP*, 37, 248–253.

Mihai, D. (2015). RoboDK: An Offline Programming and 3D Simulation Software for Industrial Robots. Accessed 2 December 2020.

Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., Lin, Z., Desmaison, A., Antiga, L., and Lerer, A. (2017). Automatic differentiation in pytorch.

Pieskä, S., Pitkäaho, T., and Kaarlela, T. (2020). Multilayered dynamic safety for high-payload collaborative robotic applications. In *2020 3rd International Symposium on Small-scale Intelligent Manufacturing Systems (SIMS)*, 1–6. IEEE.

Rashid, A., Peesapati, K., Bdiwi, M., Krusche, S., Hardt, W., and Putz, M. (2020). Local and global sensors for collision avoidance. In *2020 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, 354–359. IEEE.

Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 779–788.

Robla-Gómez, S., Becerra, V.M., Llata, J.R., Gonzalez-Sarabia, E., Torre-Ferrero, C., and Perez-Oria, J. (2017). Working together: A review on safe human-robot collaboration in industrial environments. *IEEE Access*, 5, 26754–26773.

Saenz, J., Behrens, R., Schulenburg, E., Petersen, H., Gibaru, O., Neto, P., and Elkmann, N. (2020). Methods for considering safety in design of robotics applications featuring human-robot collaboration. *The International Journal of Advanced Manufacturing Technology*, 1–19.

Sim, Z.H., Chook, Y., Hakim, M.A., Lim, W.N., and Yap, K.M. (2019). Design of virtual reality simulation-based safety training workshop. In *2019 IEEE International Symposium on Haptic, Audio and Visual Environments and Games (HAVE)*, 1–6. IEEE.

Unity (2020). Game engines - how do they work? URL `https://unity3d.com/what-is-a-game-engine`. Accessed 20 May 2020.