Intelligent robot assistants for the integration of neurodiverse operators in manufacturing industry

Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 126 (2024) 236–241

www.elsevier.com/locate/procedia

17th CIRP Conference on Intelligent Computation in Manufacturing Engineering (CIRP ICME '23)

Intelligent robot assistants for the integration of neurodiverse operators in manufacturing industry

Yuchen Fan[a,b], Dario Antonelli[b], Alessandro Simeone[b]*, Nengsheng Bao[a]

[a]Intelligent Manufacturing Key Laboratory of Ministry of Education, Shantou University, Shantou, China
[b]Department of Management and Production Engineering, Politecnico di Torino, Corso Duca degli Abruzzi 24, Torino, Italy

* Corresponding author. Tel.: +39-011-0907689. E-mail address: alessandro.simeone@polito.it

Abstract

The integration of neurodiverse operators in industry poses interesting challenges highlighting the potential for human-robot collaboration(HRC) to provide the necessary support and assistance to complete manufacturing tasks. This study proposes a reciprocal learning-based framework to support the operator to carry out assembly tasks via collaborative robot (Cobot) assistance. Such framework includes image acquisition and processing, machine learning (ML)-based classification of workpieces and Cobot operational tasks.

An experimental case study is proposed to verify the framework applicability to a number of likely scenarios. Preliminary results show the suitability of HRC systems to effectively aid neurodiverse operators with memory issues in performing assembly tasks correctly, while at the same time improving the operator ability to learn the assembly sequence and iteratively improving ML classification accuracy. The potential for intelligent robotics-based increased inclusiveness in manufacturing industry is discussed in terms of benefits to support individuals with cognitive disabilities.

Keywords: Reciprocal learning, Inclusive manufacturing, Human-Robot Collaboration, Neurodiversity.

1. Introduction

Neurodiversity is a concept that encompasses a wide range of neurodevelopmental and neurofunctional differences, including but not limited to dyslexia, dyspraxia, autism, attention deficit hyperactivity disorder (ADHD), learning disabilities and anxiety disorders [1, 2].

Promoting neurodiversity inclusion in manufacturing industry aligns with the goal of sustainable economic growth outlined in the United Nations' 17 Sustainable Development Goals (SDGs, 2015) [3], which is to promote sustained, inclusive, and sustainable economic development, along with ensuring full and productive employment opportunities, and promoting decent work conditions for all. In line with this, enabling neurodiverse individuals to integrate into the manufacturing scenario is crucial to achieving the highest possible level of satisfaction, personal fulfillment, social integration, and autonomy through their professional work [4].

In this context, with reference to manufacturing environments, significant complexities are posed by assembly tasks, which are of crucial importance, as they directly affect quality, efficiency and productivity [5]. When considering the integration of neurodiverse operators into these tasks, it becomes essential to acknowledge the distinct challenges they may face [6], such as:

- Attention: keeping focused on the current task;
- Short-term memory: storing information temporarily;
- Long-term memory: memory retrieval for future use;
- Logic and reasoning: prioritizing, planning, and problem-solving;

10.1016/j.procir.2024.08.332

- Language processing: recognizing letters and words and understanding written or spoken language;
- Mathematical processing: recognizing numbers and symbols and understanding and calculating simple math.

The ability to memorize and self-attention [7] are essential for operators to perform assembly tasks, where operators must remember the sequence in which components are to be assembled, the specific tools required for a task and the sequence of steps needed to complete a task. Differences in memory and self-attention ability can directly affect the accuracy and efficiency with which an operator performs a manually assembly task [8, 9].

However, neurodiverse operators may encounter unique challenges related to memory. For instance, individuals with ADHD may struggle with short-term memory, making it difficult for them to remember detailed information, while those with dyslexia may experience difficulties with working memory, which belongs to short-term memory, making it harder for them to remember multi-step procedures [10, 11]. Higher levels of dyslexia symptoms were found to be positively correlated with daily attention deficits and the frequency of memory failures.[11] Therefore, recognizing memory and self-attention significance in assembly tasks and supporting neurodivergent operators is vital for inclusivity.

Collaborative robots, or Cobots, are intelligent robots that can work alongside humans [12]. Through a sensor network, Cobots can sense the presence of humans and adjust their actions accordingly to work together with them to complete tasks [12].

Human-Robot Collaboration (HRC) systems combine Cobots, ML and computer vision (CV), and can potentially accommodate the abilities, preferences, and limitations of neurodiverse operators, providing enhanced support for assembly tasks. Utilizing ML and CV technologies, the system can dynamically perceive the actions of operators in real-time, offering accurate assistance and guidance to improve the efficiency and precision of assembly tasks. Cobots can shoulder repetitive, tedious, or hazardous tasks, relieving operators of burdensome responsibilities. This promotes an inclusive manufacturing environment.

Therefore, this paper aims at developing a HRC system that integrates ML, CV and Cobots to assist neurodiverse operators in performing assembly tasks. A framework is therefore proposed to illustrate the specific technical methodology, and a laboratory-scale case study is reported to demonstrate the effectiveness of the proposed approach.

2. Framework and methodology

The methodological framework proposed in this paper comprises three key technical components: ML Model Training, Real-time Object Detection, and Collaborative Assembly, as illustrated in Fig. 1.

The methodology also incorporates the involvement of humans and machines in a loop, highlighting the collaborative nature of the approach. Below are the detailed explanations of each component:

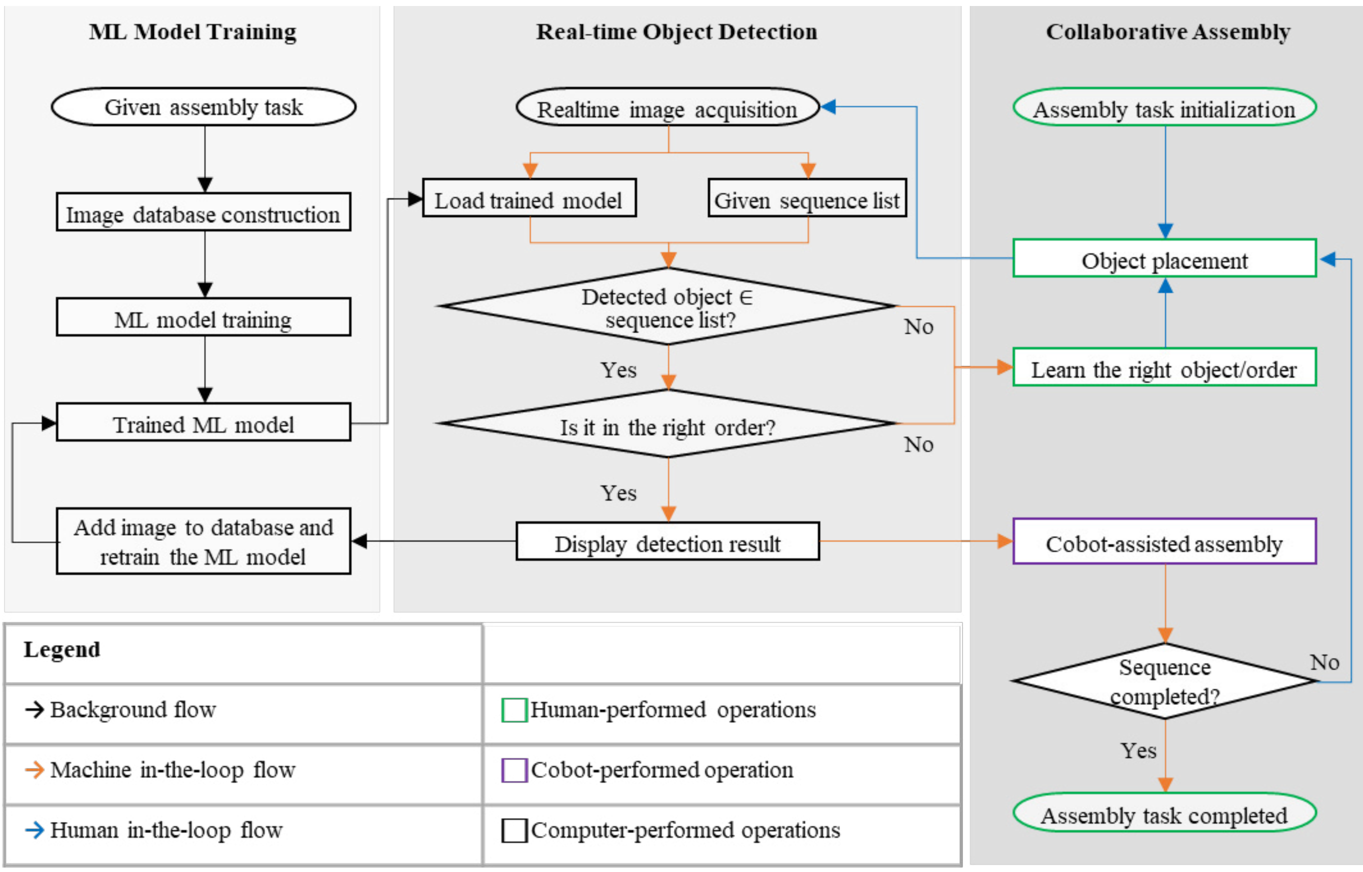


Fig. 1. Overall framework flowchart

First, the neurodiverse operator selects an object to be placed on the detection platform, then the camera captures the real-time image in the fixed work area containing the object, and loads the pre-trained ML model and pre-saved sequence list for object detection. Object detection is divided into two steps: 1) detecting whether the object is in the sequence list; 2) detecting whether objects are in the correct order. Whenever there is an error in the detection, it means that the object has been placed incorrectly and the operator needs to learn the correct object or sequence and re-place it to be detected again.

When both detection steps are correct, the image of the object is transferred to the ML model database to help train the model again to improve precision, while Cobot grabs the object and moves it to the assembly area. The assembly task is completed when all objects in the pre-set sequence list have been successfully placed.

3. Case study

In order to validate the framework developed in this paper, a case study is reported to support assembly task. In this case study the HRC system will perform object detection, sequence assessment and object manipulation, while the neurodiverse operator will perform workpiece picking and placement and completing the assembly task. The essence of the case study lies in the reciprocal learning process between the Cobot and the neurodiverse operator. As the collaboration progresses, the Cobot enhances its ability to recognize and classify a diverse set of objects with greater accuracy. Simultaneously, the human operator acquires valuable knowledge on how to correctly execute the assembly sequence. The mutual learning between the Cobot and the neurodiverse operator is crucial to the framework, demonstrating effective cooperation and skill enhancement in human-robot collaboration.

3.1. Experimental setup

An OMRON TM5-900 Cobot is used to perform the collaborative assembly with the human operator. The robotic arm is endowed with an on-wrist 5 Mpx camera for localization purposes. An additional webcam is placed on the top of the platform which is used for the actual object detection and classification, as shown in Fig. 2.

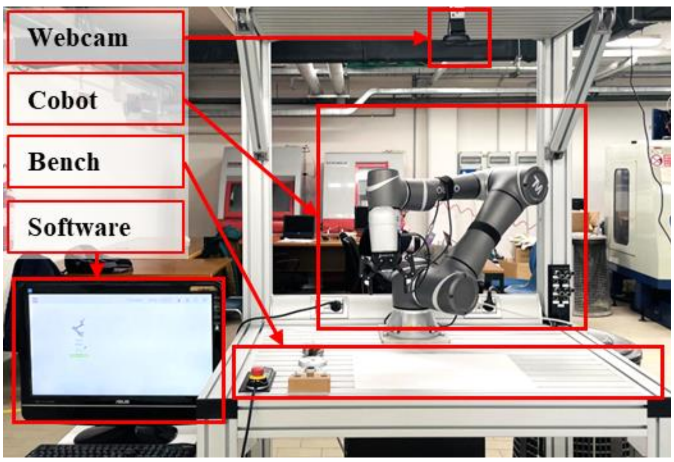
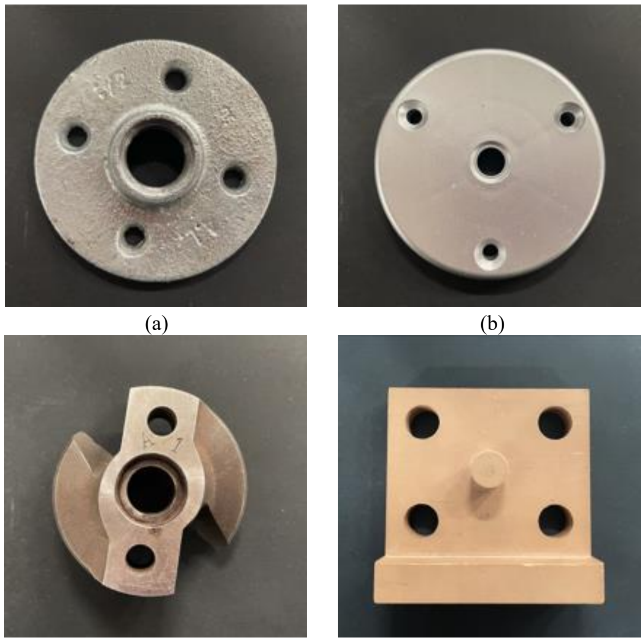


Fig. 2. Experimental setup.



(a) (b) (c) (d)

Fig. 3. Assembly set: circular flanges with: (a) 4 holes, (b) 3 holes, (c) 2 holes, (d) square flange

3.2. Experimental procedures

An assembly sequence has been set in this paper for simulation including 4 components, namely 4-hole circular flange (A), 3-hole circular flange (B), 2-hole circular flange (C) and a square flange (D) in this specific order. Such components, shown Fig. 3, are used as items to train the YOLOv7, according to their different features. The whole experimental process is divided into 3 steps: dataset creation, deep learning (DL) modelling and real-time HRC assembly, as shown in Fig.4. The object classification is configured here as a pattern recognition problem to be performed via ML.

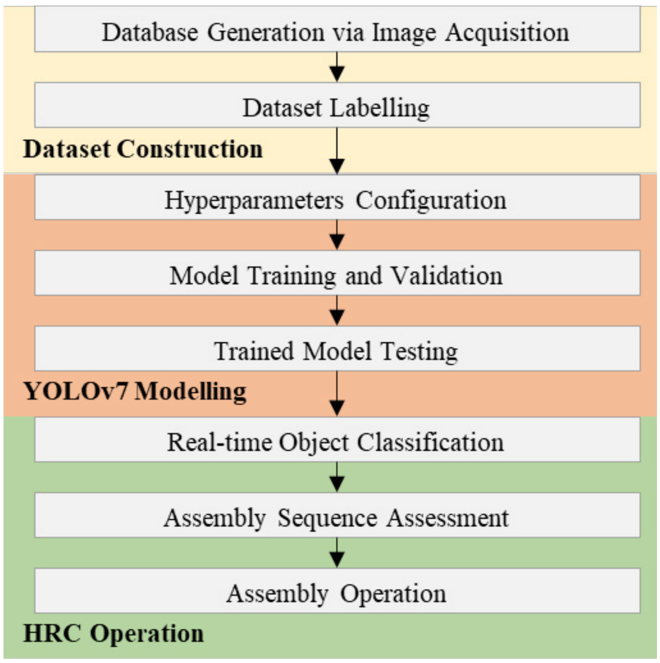


Fig. 4. Flowchart of experimental activities.

To achieve this, a DL algorithm known as YOLOv7 [13] is utilized. YOLOv7 is a highly efficient and fast neural network designed to detect and classify objects in digital images. The algorithm divides an input image into a grid, with each cell being responsible for detecting objects within its boundaries. One of the key advantages of YOLOv7 is its speed and efficiency in processing large volumes of data. This makes it particularly suitable for use in real-time applications. Additionally, the algorithm ability to accurately detect and classify objects in complex and cluttered scenes further enhances its utility in a variety of practical settings [13].

In this context, the digital images were firstly acquired using the webcam and OpenCV [14], an open-source CV and ML software library employed to automate the image acquisition and database management. All the digital images were then manually labelled using LabelImg [15], an image annotation tool, allowing forthe image labelling according to the YOLO format, containing the category id and the bounding box coordinates. Such labelling operation is necessary for the YOLOv7 training.

The image dataset consists in 80 images for each part, a total of 320 images. Then the YOLOv7 configuration was carried out in terms of number of epochs (300), image size (640*640 pixels), batch size (16), learning rate (0.1%), training, validation and testing set consisting in respectively 240 images (75%), 40 images (12.5%) and 40 images (12.5%). In accordance with the above mentioned parameters, the YOLOv7 training was conducted, and the resulting loss values are used as indicators of the model performance throughout the training process. The loss values obtained during training reflect how well the model performed. Loss values measure the difference between what the model predicted and the actual labels. Smaller loss values mean that the model predictions were closer to the actual labels [14].

The trained model showed great effectiveness during training by achieving low average losses for the three considered indicators [16], i.e., bounding box, classification, and object detection (0.01145, 0.00175, and 0.00236, respectively). The specific YOLOv7 training results is shown in Table 1 and Fig. 5 for four image instances.

With reference to Table 1, precision indicates the model accuracy in predicting positive samples, while recall is the proportion of true positive samples identified by the model. The mAP@.5 indicator refers to the average precision at an intersection over union(IoU) threshold of 0.5, and mAP@.5:.95 refers to thresholds ranging from 0.5 to 0.95. This pre-trained YOLOv7 model will be put to use in real-time object detection and will be intensively trained to improve accuracy as it is used. The proposed HRC system is endowed with a graphical user interface (GUI) to facilitate the neurodiverse operator with assembly tasks.

Table 1. YOLOv7 training results.

Component	Precision	Recall	mAp@.5	mAp@.5:.95
A	0.995	1	0.995	0.952
B	0.997	1	0.995	0.979
C	1	1	0.995	0.92
D	0.993	1	0.995	0.963
Total	0.996	1	0.995	0.954

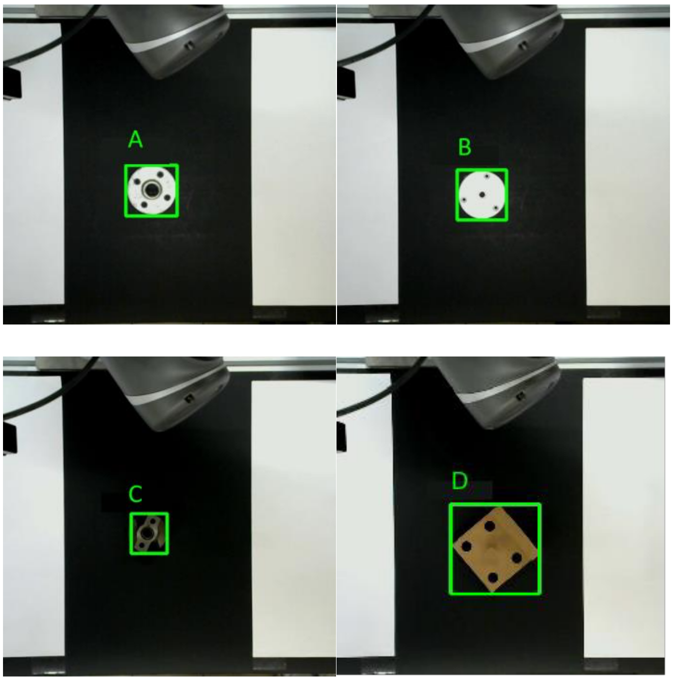


Fig. 5. Example of classification test results.

This is shown in Fig. 6, including the real-time image acquisition view along with the YOLOv7 prompt and dialog box. The interaction between the operator and the cobot is carried out via GUI. The operator first selects a component and put it in a predetermined assembly area. The webcam captures a real-time image and then detects this image by the pre-trained YOLOv7 model, which, upon classification, can show three output messages:

- Correct_ the component is correct and placed in the right assembly sequence step;
- (2) Object is not in the sequence list: the component does not belong to the sequence list;
- (3) Object is in the sequence list but placed in a wrong order.

Only when the objects in the list are placed in the correct order in the assembly area, the HRC system establishes a connection with the Cobot and transmits the coordinates of the recognised objects to the Cobot for its initial positioning. The Cobot receives the coordinates and automatically moves to the centre of the coordinates and then switches on its on-wrist camera for further precise positioning, as shown in Fig. 7(b).

Once the Cobot has calibrated the exact location of the object, it carries out a series of pre-set actions: first it opens the gripper and moves down a certain distance according to the pre-set spatial coordinates, then it grabs the object and lifts it to a certain height, moves again to directly above the assembly box, the Cobot drops in height and opens the gripper, successfully placing the object in the assembly box, and then returns to the origin as shown in Fig. 7.

When Cobot operations are completed and the assembly task is finished, the system is ready for the final step, i.e., the YOLOv7 model retraining. During this process, the image is automatically added to the YOLOv7 training set. The related object annotation information, including the category ID and the coordinates of the bounding box, are also automatically stored.

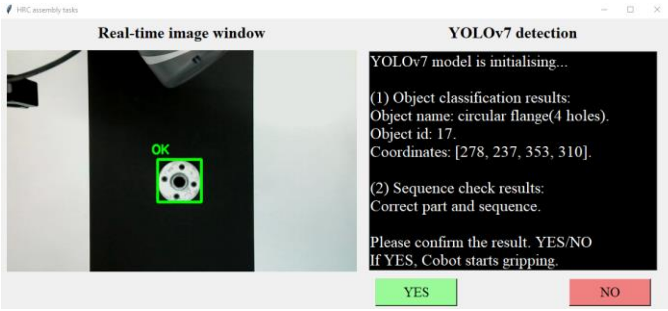


Fig. 6. GUI screenshot during YOLOv7 detection.

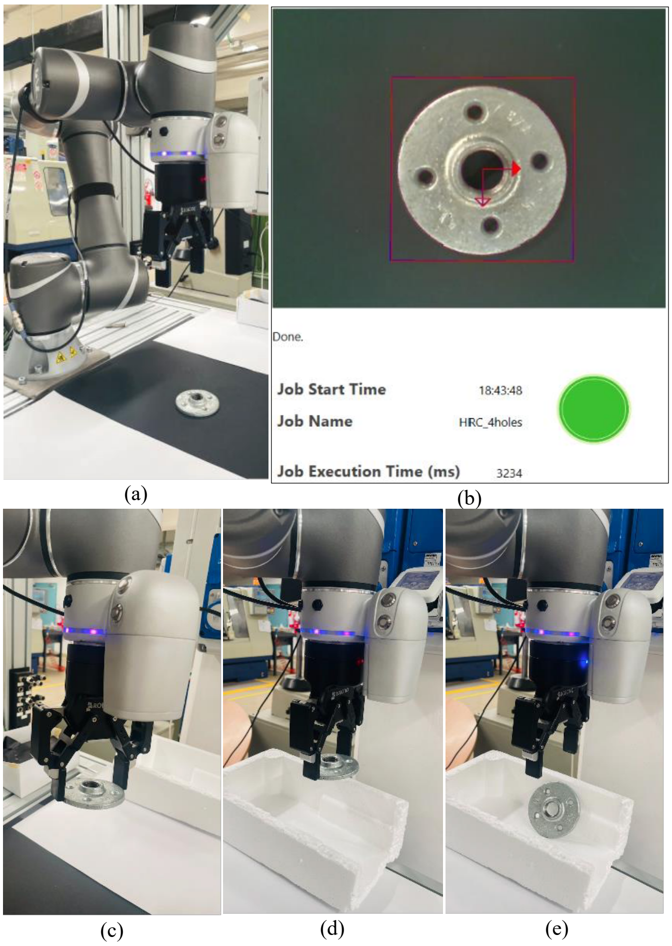


Fig. 7. Cobot operation: (a) Cobot localization; (b) View of on-wrist camera; (c) Cobot gripping; (d) Cobot placing; (e) Cobot releasing.

Once the new image has been appended, the YOLOv7 model is retrained. To validate the experimental procedure based on the proposed framework, three series of experimental tests have been carried out simulating three possible scenarios, i.e. the correct assembly sequence, the use of a component that does not belong to the assembly set and the wrong assembly sequence.

The correct sequence was simulated by placing the components as described above, i.e., $[A \rightarrow B \rightarrow C \rightarrow D]$; the use of components not belonging to the assembly set has been simulated by placing in random order unknown components with similar size and shape of the assembly components . The wrong assembly order has been simulated by randomly swapping the components, one per each test.

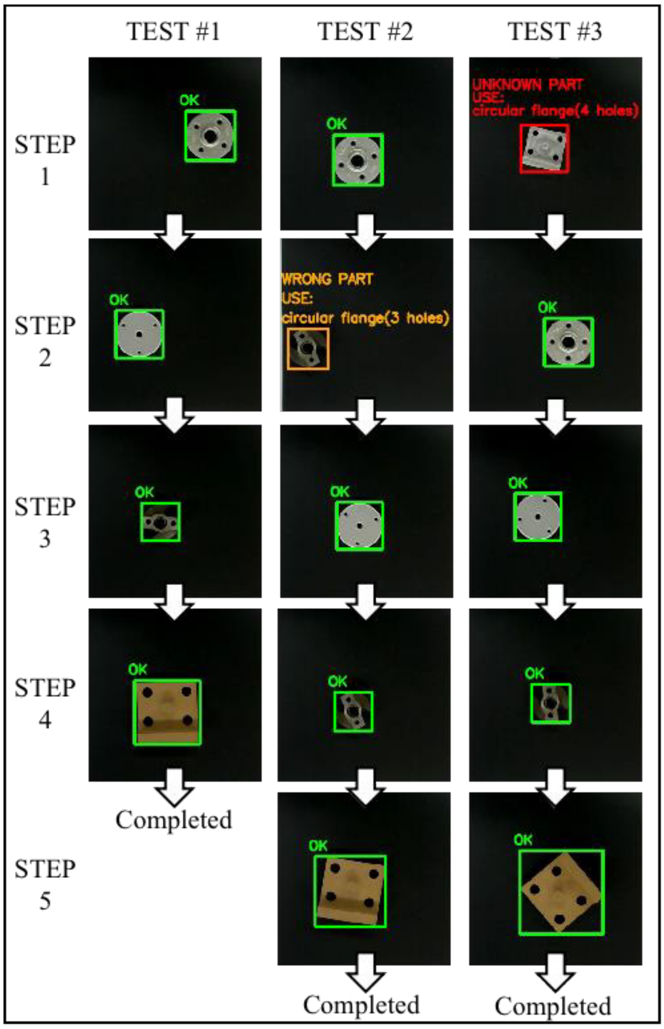


Fig. 8. Example of experimental processes.

An experimental test is considered successful if no misclassification instance occurs throughout the whole assembly sequence, conversely, if a misclassification occurs, the test is interrupted and considered unsuccessful. Each test has been repeated at least 20 times to increase the statistical reliability, and the threshold value for real-time object detection was set to 0.9. An example of each experimental test series is shown in Fig. 8.

4. Results and discussion

In Test #1 series, the assembly sequence was completed correctly 20 times with no misclassification reported, with a success rate of 100%. In Test #2 series, each component was purposedly misplaced 20 times each, and a all the 80 instances were successfully classified, for a success rate of 100%. In the Test #3 series, the experimental tests included objects that exhibited similar visual characteristics to the objects in the predefined sequence list. During this testing phase, the identification process successfully classified 17 instances as "unknown" objects, while 3 instances were misclassified. As a result, the overall success rate reached 85%. These results show a good potential of the YOLOv7 classification system support HRC while highlighting the criticality of overfitting issues.

With the assistance of ML, operators, whether they have normal cognitive abilities or memory deficiencies, can benefit from a system that ensures they do not deviate from the correct assembly sequence. ML algorithms can provide continuous reminders of the correct procedure, helping to minimize errors and improve overall productivity. This feature is particularly valuable for individuals with memory deficiencies who may struggle to remember complex sequences.

Conversely, the robot functionality should not be hindered by failures in the image detection algorithm. Even if the algorithm fails to recognize a part on the workbench, the robot is instructed to continue its operation without unnecessary interruptions and the new sample is added to its training set. This ensures that the workflow remains smooth and efficient, without excessive reliance on flawless image detection.

However, there are still some issues that need to be addressed in the system. The image detection algorithm may not perform optimally on standard metal workbenches, as they lack a uniform and contrasting background. The algorithm accuracy can be compromised due to the absence of distinct visual cues. Additionally, the algorithm performance is highly sensitive to lighting conditions and can be further affected by the presence of shadows. These challenges need to be overcome through advanced image processing techniques or alternative sensing mechanisms to ensure reliable and robust part recognition. Moreover, the current human-machine interface (HMI) used to communicate with the operator is limited to text and images. This interface can be slow and may add cognitive load to the operator, potentially affecting their efficiency and task performance. To address this, alternative interfaces such as voice commands or augmented reality (AR) interfaces could be explored. Voice-based instructions can provide real-time guidance, reducing the reliance on visual cues and allowing for faster and more intuitive communication. AR interfaces can overlay digital information onto the physical environment, providing contextual information and step-by-step instructions, thereby reducing cognitive load and enhancing operator understanding and performance.

5. Conclusions

This study investigated the potential of HRC to support neurodiverse operators in a manufacturing environment. This paper has to be considered a kick-off which aims at verifying the suitability of two enabling technologies, namely the Cobot and the ML assisting the neurodiverse operator assembly tasks.

The potential benefits include increased productivity, improved accuracy, reduced error rates, as well as the ability for operators to learn from the physical and cognitive support provided by the robots. This is supposedly a driver to increase comfort and confidence at work. At the same time, the operators can help the HRC system to boost the robot learning capabilities, leading to reciprocal learning between humans and machines. While there are many benefits of ML assistance in assembly tasks, several experimental challenges remain. These include improving the performance of image recognition algorithms on metal workbenches, addressing lighting and shadow issues, and enhancing the HMI with speech or AR to create a more efficient and user-friendly system in industrial environments.

Acknowledgements

This work was partially supported by the 2020 Li Ka Shing Foundation Cross-Disciplinary Research under Grant 2020LKSFG06D.

References

[1] Walkowiak, E., 2021. Neurodiversity of the workforce and digital transformation: The case of inclusion of autistic workers at the workplace. *Technological Forecasting and Social Change*. 168 (March 2020), 120739.

[2] LeFevre-Levy, R., Melson-Silimon, A., Harmata, R., Hulett, A.L., and Carter, N.T., 2023. Neurodiversity in the workplace: Considering neuroatypicality as a form of diversity. *Industrial and Organizational Psychology*. 16 (1), 1–19.

[3] United Nations General Assembly, 2015. Transforming our world: the 2030 Agenda for Sustainable Development. .

[4] Kildal, J., Maurtua, I., Martin, M., and Ipiña, I., 2018. Towards Including Workers with Cognitive Disabilities in the Factory of the Future. in: Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility, ACM, New York, NY, USApp. 426–428.

[5] Nourmohammadi, A., Fathi, M., and Ng, A.H.C., 2022. Balancing and scheduling assembly lines with human-robot collaboration tasks. *Computers & Operations Research*. 140 105674.

[6] Walkowiak, E., 2021. Neurodiversity of the workforce and digital transformation: The case of inclusion of autistic workers at the workplace. *Technological Forecasting and Social Change*. 168 (March 2020), 120739.

[7] Zhang, R., Lv, J., Li, J., Bao, J., Zheng, P., and Peng, T., 2022. A graph-based reinforcement learning-enabled approach for adaptive human-robot collaborative assembly operations. *Journal of Manufacturing Systems*. 63 491–503.

[8] Male, J. and Martinez-Hernandez, U., 2021. Collaborative architecture for human-robot assembly tasks using multimodal sensors. in: 2021 20th International Conference on Advanced Robotics (ICAR), IEEE, pp. 1024–1029.

[9] Hou, L. and Wang, X., 2013. A study on the benefits of augmented reality in retaining working memory in assembly tasks: A focus on differences in gender. *Automation in Construction*. 32 (October 2021), 38–45.

[10] Provazza, S., Adams, A.-M., Giofrè, D., and Roberts, D.J., 2019. Double Trouble: Visual and Phonological Impairments in English Dyslexic Readers. *Frontiers in Psychology*. 10 (December),.

[11] Protopapa, C. and Smith-Spark, J.H., 2022. Self-reported symptoms of developmental dyslexia predict impairments in everyday cognition in adults. *Research in Developmental Disabilities*. 128 (March 2021), 104288.

[12] Segura, P., Lobato-Calleros, O., Ramírez-Serrano, A., and Soria, I., 2021. Human-robot collaborative systems: Structural components for current manufacturing applications. *Advances in Industrial and Manufacturing Engineering*. 3 100060.

[13] Sivkov, S., Novikov, L., Romanova, G., Romanova, A., Vaganov, D., Valitov, M., et al., 2020. The algorithm development for operation of a computer vision system via the OpenCV library. *Procedia Computer Science*. 169 (2019), 662–667.

[14] Wang, C.-Y., Bochkovskiy, A., and Liao, H.-Y.M., 2023. YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors. in: 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, pp. 7464–7475.

[15] Tzutalin, 2015. labelImg. *Git Code*.

[16] Kambhatla, A. and Ahmed, K.R., 2023. Firearm Detection Using Deep Learning. in: Intelligent Systems and Applications. IntelliSys 2022. Lecture Notes in Networks and Systems, Vol 544, pp. 200–218.