A framework for digital assembly instructions as a step towards manufacturing inclusiveness

Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 132 (2025) 116–121

www.elsevier.com/locate/procedia

12th CIRP Global Web Conference (CIRPe 2024)

A framework for digital assembly instructions as a step towards manufacturing inclusiveness

Yuchen Fan [a], Alessandro Simeone [a,*], Dario Antonelli [a], Alessandra Caggiano [b,c], Paolo C. Priarone [a], Luca Settineri [a]

[a]Department of Management and Production Engineering, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129, Turin, Italy
[b]Center for Advanced Metrological and Technological Services (CESMA), University of Naples Federico II, Piazzale Tecchio 80, 80125, Naples, Italy
[c]Fraunhofer Joint Laboratory of Excellence on Advanced Production Technology (Fh-J_LEAPT UniNaples), Naples, Italy

* Corresponding author. Tel.: +390110907689. *E-mail address:* alessandro.simeone@polito.it

Abstract

Industrial manufacturing processes require precise understanding of instructions, which can be challenging for neurodiverse operators with reading difficulties. To bridge this gap, a digital instruction framework using object detection and natural language processing is proposed in this research. The framework uses an intelligent vision system to monitor task execution, coupled with the automatic generation of personalised voice instructions via large language models. This approach aims to improve accessibility and inclusivity in assembly lines. A case study on the assembly of a horizontal bare-shaft centrifugal pump demonstrates the effectiveness of the framework in reducing assembly errors and improving operational efficiency, making it particularly beneficial for neurodiverse individuals and promoting an inclusive work environment.

Keywords: Digital work instructions; Inclusive manufacturing; Assembly; Industry 5.0; Artificial Intelligence

1. Introduction

The manufacturing industry has historically faced challenges in creating inclusive working environments, particularly in assembly processes that often require high levels of skill and precision. Traditional work instruction methods can be insufficiently adaptable, creating barriers for a diverse workforce, including those with varying levels of experience, physical/cognitive abilities and language proficiencies. Simplifying information presentation and developing effective information assistance systems that act as ‘cognitive amplifiers’ can notably enhance operator performance and reduce errors [1, 2]. As traditional text-based instructions often fail to consider operators’ cognitive processes, leading to inefficiencies and errors, there is a need to improve the way instructions are given to properly support operators, including not only common product/process-related factors but also adaptable, human-centred features [1]. Recent literature has attempted to define the functionalities that future assistive systems should provide to reduce mental load in manual assembly processes, and has presented some approaches for their implementation [2]. Some methodologies focus on providing informational guidance, such as visual guidance technologies, while others concentrate on delivering stepwise and individually tailored assembly instructions [3]. The adoption of design principles for information presentation to support operators in assembly tasks by considering cognitive limitations, individual differences and needs has also been proposed [4]. In this context, the use of digital assistance systems can offer sensory and cognitive support to employees, by supporting perception (data acquisition) and providing workers with instructions and analyses (data processing) [5].

The digitalisation of work instructions serves as a valuable tool to improve performance, reduce cognitive load, and offer

10.1016/j.procir.2025.01.020

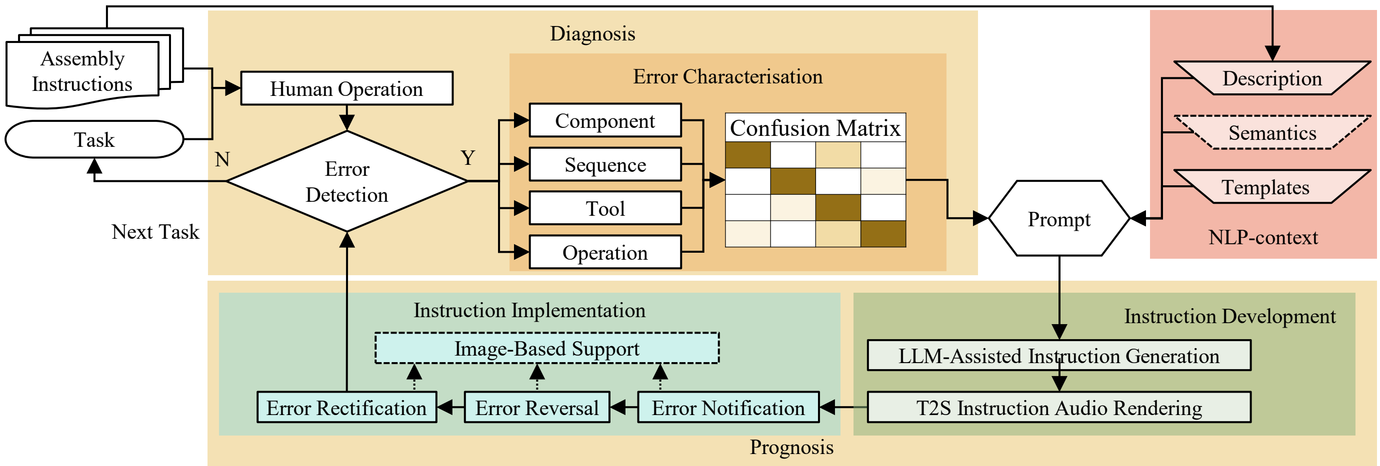

Fig. 1. General framework flowchart.

superior cognitive support compared to traditional text or paper-based instructions, aiming to enhance efficiency, quality and operator well-being [1]. Most modern assistance systems are based on Pick-by-Light or Pick-by-Vision methods [6], typically supported via image-text combinations, e.g. based on AR technologies allowing visualization on various devices like tablets or AR glasses [7–9]. Other approaches provide visual feedback via app to the operator and employ motion capture technology during manual assembly [5]. Instead, with the aim of fostering inclusivity in manufacturing. this research work proposes a digital instruction framework design to provide information support via different communication channels, including voice. The proposed approach leverages technologies such as intelligent vision, natural language processing (NLP) and large language models (LLM) technologies for developing an interactive user interface capable to provide customised voice instructions to operators. In manufacturing, the use of LLM, implemented in GPT-4, was explored in [10] to develop an AI-based approach to human-robot collaborative assembly. In [11], NL-based task prompt instructions were used for human-robot collaboration and fed to LLMs for adaptive task planning. In the proposed framework, personalised, accessible, and efficient guidance for assembly workers is realised through the implementation of intelligent task execution monitoring and diagnosis based on deep learning for object detection, NLP-based interpretation of assembly instructions and error narratives, and the automatic generation of personalised voice instructions via LLMs. The implementation of the system, which enables multiple digital instruction communication channels, not only improves operational efficiency, but also promotes a more inclusive workplace by accommodating the diverse needs of all employees. To verify the potential of this framework to reduce assembly errors and improve operational efficiency, a case study of the assembly of a horizontal bare-shaft centrifugal pump is considered.

2. Framework and methodology

The proposed digital instruction framework, shown in Fig. 1, integrates an intelligent vision system with NLP and LLMs to provide inclusive, real-time support for neurodiverse operators in industrial manufacturing environments. Tailored human-machine communication channels can be implemented to address various neurodiversity and cognitive challenges. Operators with memory issues benefit from a graphical user interface designed with intuitive visual elements such as icons and color-coded sections. Those with writing difficulties use vocal synthesis, which converts spoken instructions into machine-readable commands via speech recognition technologies. For those with reading difficulties, audio messages are transmitted through earphones. Operators with attention issues receive information via a mix of visual cues and auditory signals. For language difficulties, the user interface employs simple, possibly visual language and offers multilingual support. Finally, for those with logic difficulties, the interface provides clear, step-by-step instructions and visual process maps. In this research, the proposed system is designed to detect and correct operational errors, streamline assembly processes, and facilitate task comprehension through personalised voice instructions. The framework operates through the modules illustrated in the following sections.

2.1. Diagnosis module

The 'diagnosis' module is triggered when an operator initiates a task. The intelligent vision system scans the work area, identifying components, tools and operator actions to ensure that the predefined assembly sequence is followed. If a discrepancy is observed, signalling a potential error (i.e., the 'Yes' path in Fig. 1), the system immediately flags the incident. Once an error is detected, the diagnostic module performs an error characterisation process. This module assesses the type of error, which in this research may involve:

- *Component*: the given part/assembly involved in the error;
- *Sequence*: the order of operations undertaken by the operator and where the deviation occurred;
- *Tool*: the equipment being used at the time of the error, determining if incorrect tool use contributed to the issue;
- *Operation*: the particular action being performed incorrectly by the operator.

The output of this module is a confusion matrix. This matrix characterises assembly errors by categorising them as true positives, false positives, true negatives, and false negatives.

2.2. Instruction development

Within the digital instruction framework, the NLP subsystem is constructed to provide a precise interpretation of both assembly instructions and error narratives. A complete understanding of the product and process context is crucial in order to adapt the NLP models to the domain-specific language used in manufacturing [11]. In this respect, manually-prepared Context Description derived from the assembly instructions provides the background on the assembly process, detailing specific tools, components and operations.

This helps NLP tools in accurately interpreting terminology and contexts within textual error descriptions, ensuring optimal performance and precision [12]. Along with the context description, the proposed framework includes customised error log templates that standardise the documentation of errors by organising information extracted from the NLP analysis and insights from the confusion matrix. Each template should include date, time, operator name, location, error description, error details and corrective actions to facilitate the rapid understanding and implementation of corrections [13].

At this point, the framework initiates a procedure to create an error log by populating the templates with textual data from the assembly instructions and the descriptions extracted from the confusion matrix. These descriptions detail instances of misclassification, such as incorrect component or tool usage, and provide the raw data necessary for further analysis. Using advanced NLP techniques, the framework then applies semantic analysis to this textual data [14].

This involves syntactic structure analysis, parsing sentences to understand their grammatical composition and the roles of various parts of speech, as well as semantic role labelling, which identifies relationships and interactions within error descriptions. The insights from the semantic analysis are subsequently integrated with quantitative data from the confusion matrix and contextual information from the assembly process [15]. This is done by correlating the detailed error insights derived from the semantic analysis with specific error instances identified in the confusion matrix, providing an explanation of error occurrence. Context descriptions complement this step by providing background information that helps to contextualise each error, but also enhances the detail and accuracy of the information captured in the error logs. The error logs themselves are created at this stage by automatically filling in the templates mentioned above. These templates, designed to ensure consistency, are populated with all the necessary details such as the specific tool, component, and the nature of the error, as soon as errors are detected and logged.

These enhanced error logs, enriched with semantic insights, are then used along with a detailed prompt, to feed an LLM with the aim of generating precise and accurate instructions for human operators [16, 17]. The prompt should contain instructions to generate a dyslexic-friendly and voice-enabled simple explanation of the error, with numbered correction steps, in a supportive tone, based on the above inputs. The additional level of understanding provided by semantic analysis ensures that these applications are more accurate and effective in addressing the specific needs highlighted by the error logs. The LLM, through an interactive process, generates task-specific instructions which are then rendered into audio format through text-to-speech (T2S) technology, providing a hands-free, attention-getting modality for operators to follow. This component is particularly beneficial for neurodiverse operators who may find visual instructions difficult to interpret.

2.3. Prognosis module / Instructions implementation

The voice instruction message, once produced, is transmitted directly to the operator via earphones to ensure that the guidance is delivered promptly and clearly [18].

In cases where an operation has been performed incorrectly, the voice instructions provide a detailed guidance on how to reverse the error, facilitating the operator to restore the affected component, tool, or sequence to its previous correct state.

This consists in providing verbal instructions to help the operator perform the task correctly, fostering compliance to the assembly procedures.

The whole process is augmented by visual support presented on a video terminal. The imagery parallels the vocal guidance, providing an integrated, multi-sensory instructional experience [18]. This not only assists in the immediate correction of the error, but also serves as a visual reinforcement, increasing the operator's understanding and mastery of the task. This cohesive method can ensure that corrective measures are not only implemented effectively, but also reinforced for future operations.

3. Case study

To demonstrate the applicability of the proposed framework, an industrial case study focusing on the assembly of a cast iron horizontal bare-shaft centrifugal pump is reported here. The bill of materials (BoM), illustrated in Fig. 2 and detailed in the supplementary material, includes 26 different components to be assembled using 7 different tools (a detailed description of the materials, tools and assembly sequence is reported in the supplementary material).

To provide a proof-of-concept, the object-recognition-based error detection model was trained on 5 components, 2 tools and 3 operations. The image dataset was augmented with horizontal flips, rotations, and adjustments to saturation and brightness to enhance model robustness to various changes.

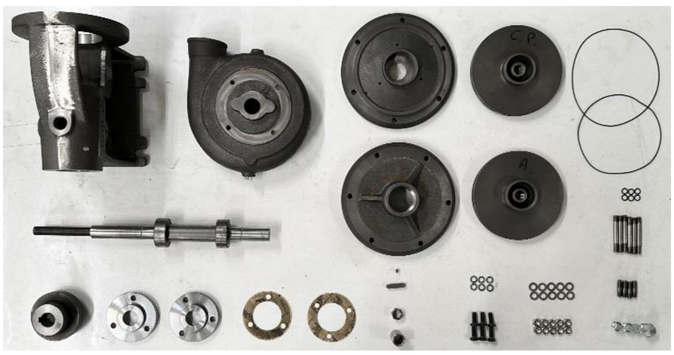

Fig. 2. Assembly components. Courtesy of ALDO BARBERA srl, Brandizzo, Torino, Italy.

Table 1. Instructions

Step	Component(s)	Tool	Instructions
S6	Hex cap screw (big) (C26)	N/A	Thread the big hex cap screw (C26) by hand into the threaded hole above the support (PS5).
S9	Drive shaft companion flange (C8)	Allen key (T5)	Attach the drive shaft companion flange (C8) onto the shaft (C5), ensuring alignment of the front surfaces of both the shaft (C5) and the flange (C8). Secure it in place using the Allen key (T5) with half a turn.

Table 2. Errors

Error ID	Error class	Error description
E1	Component	Operator picks the drive shaft companion flange (C8) instead of Hex cap screw (big) (C26) in step 6
E2	Sequence	Operator selects wrench size 13 (T6) which is not required at this step instead of Allen key (T5) in step 9

This training was carried out using YOLOv7, a convolutional neural network (CNN)-based real-time object detector, that outperforms others in speed and accuracy, requires less hardware, and trains faster on small datasets without pre-trained weights [19]. The model was trained on the Colab cloud platform with an L4 GPU using a learning rate of 0.01, batch size of 16, over 600 epochs, and completed in 2.3 hours. The training performance was satisfactory, with perfect classification in most categories and only minor misclassification for the 'Nut' and 'Washer' classes, likely due to their similar geometries. The scope of this case study is narrowed down by focusing on two assembly steps, namely S6 and S9, for which the assembly instructions are partially given in Table 1, together with the required components and tools. To speed-up the error detection, each step has been divided into three subtasks, i.e., component selection, tool selection and assembly operation. In order to simulate the conditions of neurodiversity, two different types of errors are considered and deliberately induced in this case study, all attributable to problems in reading the assembly instructions, namely a component-related error (E1) and a sequence-related error (E2). The specific error descriptions are reported in Table 2. This case study evaluates two scenarios to assess the effectiveness of a digital instruction system over traditional text-based methods. The baseline scenario involves an operator with simulated dyslexia in terms of reading problems, i.e., reading and understanding written instructions. The digital instruction scenario uses audio instructions designed and generated for neurodiverse individuals. Semantic analysis and image-based support were excluded from this case study to simplify the scope.

4. Results and discussion

The results are presented and discussed in terms of the error log display for both errors E1 and E2, a transcription of the voice-rendered digital instructions for both errors, and a time analysis to demonstrate the effectiveness of the proposed system. Figs 6-7 show the error logs for E1 and E2, respectively, consisting of the templates which are filled with data from the system log in terms of date, time, operator and working unit. Step and task data are taken from the assembly instructions, while components, tools and other diagnostic data are automatically retrieved from the confusion matrix. The error logs were then fed into OpenAI ChatGPT 4.0 via a custom interface developed in Python 3.10.7.

Fig. 3. E1 error detection.

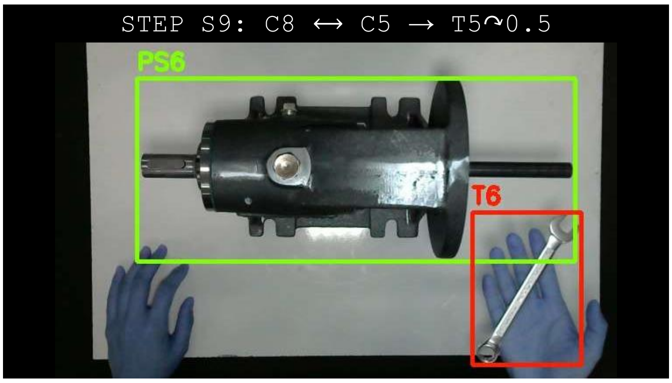

Fig. 4. E2 error detection.

```
 You are given two attached documents: the assembly instructions for a centrifugal pump and an error log. Your task is to generate clear, concise, and dyslexic-friendly instructions for the given error, suitable for voice rendering.
   ### Instructions:
   1. Use the assembly instructions document to understand the correct assembly process.
   2. Use the error log document to identify the error and the required corrective actions.
   3. Generate instructions that follow this format:
      - Greet the operator.
      - Explain the error in simple terms.
      - Provide numbered steps to correct the error.
      - Include a step to verify the correction.
      - Maintain a supportive and friendly tone.
   ### Example Output:
   Hey [Operator's Name],
   It looks like there was a mistake in the assembly process. In step [Step Number], you [describe the error in simple terms].
   Here's what to do to fix it:
   1. Stop the current task.
   2. [First corrective action]
   3. [Second corrective action] etc.
```

Fig. 5. Prompt.

```
    # Error Log
    Date and Time: 2024-06-10 10:13
    Reported by: Yuchen
    Location: Assembly Line 3, Workstation 7
    ## Error Description
    Step: S6.1
    Subtask: Prepare the following component: Hex
cap screw (big) (C26)
    Expected Component: ['Hex cap screw big (C26)']
    Actual Component: ['Drive shaft companion
flange']
    ## Error Details
    Description:
    The operator mistakenly picked the ['Drive shaft
companion flange'] instead of the ['Hex cap screw
(big) (C26)'] during step S6.1 of the assembly
process.
    ## Corrective Actions
    Immediate Action Taken:
    - Replace the ['Drive shaft companion flange']
with the ['Hex cap screw (big) (C26)']
```

Fig. 6. Error log for E1. White text: template; yellow text retrieved from confusion matrix, cyan text from the assembly instructions, magenta text from the system log.

```
    # Error Log
    Date and Time: 2024-05-31 09:53
    Reported by: Yuchen
    Location: Assembly Line 3, Workstation 7
    ## Error Description
    Step: S9.2
    Subtask: Please prepare the following tool:
Allen key (T5)
    Expected Tool: Allen key (T5)
    Actual Tool: Wrench size 13 (T6)
    ## Error Details
    Description:
    The operator mistakenly picked the ['Wrench size
13 (T6)'] instead of the ['Allen key (T5)'] during
step S9 of the assembly process.
    ## Corrective Actions
    Immediate Action Taken:
- Replace the ['Wrench size 13 (T6)'] with the
['Allen key (T5)']
```

Fig. 7. Error log for E2. White text: template; yellow text retrieved from confusion matrix, cyan text from the assembly instructions, magenta text from the system log.

This process used the prompt reported in Fig. 5 with the aim of generating the digital instructions, whose transcriptions are reported in Figs. 9-10 for E1 and E2, respectively. These instructions have been subsequently voice-rendered to obtained audio files (available in the supplementary material in mp3 format). To evaluate time efficiency, reading times were simulated based on literature regarding reading speeds (words per minute) for individuals with dyslexia [20]. The experimental tests, conducted with a single operator over 10 repetitions, comparing this baseline scenario with the digital instruction system, reveal several significant benefits. The values in Figs 10-11 represent averages of human operation measurements, while the digital instruction times are fixed. For E1 (Fig. 10), in the baseline scenario, which involves reading and understanding text instructions, instruction input takes 27.8 s. In contrast, the digital instruction system, which involves listening to and understanding instructions, reduces this time to 14.8 s. Both scenarios require 5 s for component selection and 3 s for handling. The most significant advantage of the digital instruction system is in error detection, where the baseline scenario takes 45 s due to multiple steps, including calling a supervisor and checking for errors. The digital instruction system streamlines this process, reducing the time to just 2 s. Error communication times are similar for both scenarios, taking approximately 20.8 s each. However, the digital instruction system improves error correction, reducing the time from 33 s in the baseline scenario to only 3 s. In the subsequent component selection and handling, both scenarios again take 5 s and 3 s, respectively. In the second phase of error detection, the baseline scenario takes 10 s compared to just 2 s with the digital instruction system. In summary, the total time saving for S6 results to be 94 s, which represents a 61% reduction in the step duration. For E2 (Fig. 11), instruction input in the baseline scenario involves reading and understanding text instructions and takes 14.8 s. The digital instruction system reduces this time to 9.8 s. Both scenarios require 5 s for tool selection and 3 s for use.

```
    Hey Yuchen,
    It looks like there was a small mistake in the
assembly process. In step S6, you picked up the
wrong component.
    Here's what to do to fix it:
    1. Stop the current task.
    2. Remove the 'Drive shaft companion flange' from
the assembly area.
    3. Find and pick up the correct component: 'Hex
cap screw (big), C26'.
    4. Thread the 'Hex cap screw (big), C26' by hand
into the threaded hole above the support (C1).
    After you're done, double-check your work to make
sure that the 'Hex cap screw (big), C26' is properly
installed.
    If you need any more help, just let us know.
```

Fig. 8. Digital instructions transcription for E1.

```
    Hey Yuchen,
    It looks like there was a mistake in the assembly
process. In step S9, you used a Wrench size 13
instead of the Allen key T5.
    Here's what to do to fix it:
    1. Stop the current task.
    2. Remove the Wrench size 13.
    3. Get the correct tool, which is the Allen key T5.
    4. Re-attach the drive shaft companion flange
(C8) onto the shaft (C5), ensuring the front
surfaces of both the shaft and the flange are
aligned.
    5. Secure it in place using the Allen key T5 with
half a turn.
    After you're done, double-check your work to make
sure everything is correct.
    If you need any more help, just let us know.
```

Fig. 9. Digital instructions transcription for E2.

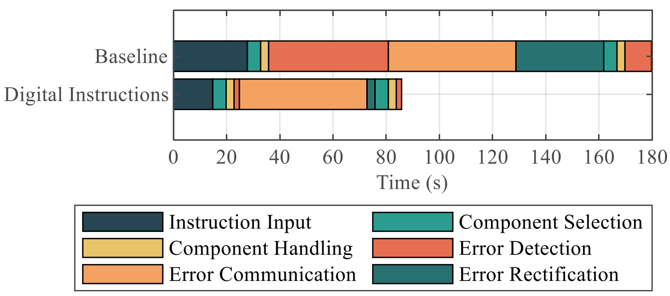

Fig. 10. Error E1 timeline bar chart

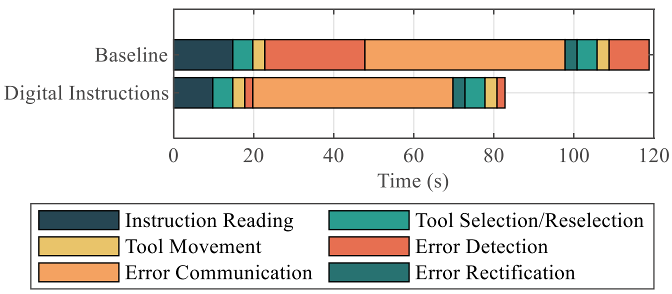

Fig. 11. Error E2 timeline bar chart

For error detection, the baseline scenario takes 25 s due to multiple steps, including calling a supervisor and checking for errors. The digital instruction system streamlines this process, reducing the time to just 2 s. Error communication times are similar for both scenarios, taking approximately 18.4 s each. However, the digital instruction system improves error correction, reducing the time from 3 s in both scenarios to return the tool to the toolbox. In subsequent tool selection and utilisation, both scenarios again take 5 s and 3 s, respectively. In the second phase of error detection, the baseline scenario takes 10 s compared to just 2 s with the digital instruction system. In this case, the overall time saving for S9 results to be 36 s corresponding to a 41% reduction of the step duration.

The proposed digital instruction system is particularly suitable for industries engaged in complex mechanical assembly of customised products. It improves operator efficiency by reducing the time required for instruction input, error detection, and error correction. This streamlined approach demonstrates the benefits of implementing the digital system over the traditional text-based method for operators with reading difficulties. By providing customised work instruction through a communication channel compatible with neurodiversity, the system is likely to reduce the excessive cognitive load resulting from task errors [21]. This reduction in cognitive load could help operators to process information more effectively, resulting in fewer errors and improved overall performance.

The proposed system has however some limitations: it relies on ideal lighting and camera positioning and extensive deep learning training. The automatic generation of instructions, whilst customised, offers limited semantics, potentially affecting complex tasks. Additionally, hardware and software requirements may be challenging in resource-limited settings. As concerns error handling features, to prevent endless loops, three consecutive identical errors trigger an alarm for supervisor intervention. Future enhancements will rephrase instructions after two identical errors.

5. Conclusions

This research presents a novel digital work instruction framework aimed at improving inclusivity in manufacturing assembly processes. By integrating object detection, NLP and LLMs technologies, the framework provides personalised, voice-based instructions to support neurodiverse operators. The case study of the assembly of a horizontal bare-shaft centrifugal pump demonstrates the potential of the framework to reduce errors, boost efficiency, and promote inclusive manufacturing through Industry 5.0 technologies. Further experimental work is required to characterise each type of error and provide the related instructions. The next step should include an estimation of the cognitive load with and without the application of the proposed framework to assess the extent of the expected benefits [21]. In addition, Incorporating this system with a collaborative robot can further enhance task performance and reduce error rates. Collaborative robots can handle repetitive tasks, allowing operators to focus on complex tasks requiring human judgment. The digital instructions guide both the operator and the robot, ensuring seamless coordination and increased productivity. This integration also supports real-time error detection and correction, enabling swift corrective actions.

Acknowledgements

The authors gratefully acknowledge ALDO BARBERA srl of Brandizzo, Torino, Italy, for providing the resources for the case study.

References

[1] Leder, R., Stern, H., and Freitag, M., 2022. Towards design guidance for the digitalisation of work instructions by focusing on technological possibilities and industrial requirements. *Procedia CIRP*. 109 466–471.
[2] Petzoldt, C., Keiser, D., Beinke, T., and Freitag, M., 2020. Functionalities and Implementation of Future Informational Assistance Systems for Manual Assembly. in: pp. 88–109.
[3] Bläsing, D. and Bornewasser, M., 2021. Influence of Increasing Task Complexity and Use of Informational Assistance Systems on Mental Workload. *Brain Sciences*. 11 (1), 102.
[4] Mattsson, S., Li, D., and Fast-Berglund, Å., 2018. Application of design principles for assembly instructions – evaluation of practitioner use. *Procedia CIRP*. 76 42–47.
[5] Keller, T., Bayer, C., Bausch, P., and Metternich, J., 2019. Benefit evaluation of digital assistance systems for assembly workstations. *Procedia CIRP*. 81 441–446.
[6] Faccio, M., Ferrari, E., Galizia, F.G., Gamberi, M., and Pilati, F., 2019. Real-time assistance to manual assembly through depth camera and visual feedback. *Procedia CIRP*. 81 1254–1259.
[7] Wang, Y., Zhang, S., Wan, B., He, W., and Bai, X., 2018. Point cloud and visual feature-based tracking method for an augmented reality-aided mechanical assembly system. *The International Journal of Advanced Manufacturing Technology*. 99 (9–12), 2341–2352.
[8] Mura, M.D., Dini, G., and Failli, F., 2016. An Integrated Environment Based on Augmented Reality and Sensing Device for Manual Assembly Workstations. *Procedia CIRP*. 41 340–345.
[9] Zheng, P., Li, S., Xia, L., Wang, L., and Nassehi, A., 2022. A visual reasoning-based approach for mutual-cognitive human-robot collaboration. *CIRP Annals*. 71 (1), 377–380.
[10] Liu, S., Zhang, J., Wang, L., and Gao, R.X., 2024. Vision AI-based human-robot collaborative assembly driven by autonomous robots. *CIRP Annals*.
[11] Wang, J., Gao, S., Tang, Z., Tan, D., Cao, B., and Fan, J., 2023. A context-aware recommendation system for improving manufacturing process modeling. *Journal of Intelligent Manufacturing*. 34 (3), 1347–1368.
[12] Gogineni, S., Lindow, K., Nickel, J., and Stark, R., 2020. Applying Contextualization for Data-Driven Transformation in Manufacturing. in: pp. 154–161.
[13] Abramovici, M., Gebus, P., Göbel, J.C., and Savarino, P., 2018. Semantic Quality Assurance of Heterogeneous Unstructured Repair Reports. *Procedia CIRP*. 73 265–270.
[14] May, M.C., Neidhöfer, J., Körner, T., Schäfer, L., and Lanza, G., 2022. Applying Natural Language Processing in Manufacturing. *Procedia CIRP*. 115 184–189.
[15] Madhusudanan, N., Gurumoorthy, B., and Chakrabarti, A., 2019. From natural language text to rules: knowledge acquisition from formal documents for aircraft assembly. *Journal of Engineering Design*. 30 (10–12), 417–444.
[16] Fan, H., Liu, X., Fuh, J.Y.H., Lu, W.F., and Li, B., 2024. Embodied intelligence in manufacturing: leveraging large language models for autonomous industrial robotics. *Journal of Intelligent Manufacturing*.
[17] Kernan Freire, S., Foosherian, M., Wang, C., and Niforatos, E., 2023. Harnessing Large Language Models for Cognitive Assistants in Factories. in: Proceedings of the 5th International Conference on Conversational User Interfaces, ACM, New York, NY, USApp. 1–6.
[18] Feng, S., He, X., He, W., and Billinghurst, M., 2023. Can you hear it? Stereo sound-assisted guidance in augmented reality assembly. *Virtual Reality*. 27 (2), 591–601.
[19] Samma, H., Al-Azani, S., Luqman, H., and Alfarraj, M., 2024. Contrastive-based YOLOv7 for personal protective equipment detection. *Neural Computing and Applications*. 36 (5), 2445–2457.
[20] Conforti, S., Marinelli, C.V., Zoccolotti, P., and Martelli, M., 2024. The metrics of reading speed: understanding developmental dyslexia. *Scientific Reports*. 14 (1), 4109.
[21] Simeone, A., Fan, Y., Antonelli, D., Catalano, A.R., Priarone, P.C., and Settineri, L., 2024. Inclusive manufacturing: A contribution to assembly processes with human-machine reciprocal learning. *CIRP Annals*. 73 (1), 5–8.