Conceptualisation of a multimodal, non-intrusive, generative AI-based assistive system for assembly

CIRP Annals - Manufacturing Technology 00 (2025) 1–5

Contents lists available at ScienceDirect

CIRP Annals - Manufacturing Technology

journal homepage: https://www.editorialmanager.com/CIRP/default.aspx

Conceptualisation of a multimodal, non-intrusive, generative AI-based assistive system for assembly

Alessandro Simeone*, Yuchen Fan, Dario Antonelli, Paolo C. Priarone (2), Luca Settineri (1)

Department of Management and Production Engineering, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy

ARTICLE INFO

Article history:
Available online xxx

Keywords:
Generative artificial intelligence
Assembly
Multimodal assistance

ABSTRACT

The transition to Industry 5.0 highlights the necessity for human-centric and adaptive manufacturing systems. This study conceptualises a multimodal, generative AI-based assistive system for assembly designed to deliver real-time error detection and adaptive guidance tailored to diverse operator profiles. The system improves human-machine interaction by issuing preventive warnings to the operator prior to critical tasks, detecting assembly errors, providing multimodal corrective instructions during operations, and deploying robotic interventions when operator-driven corrections prove inadequate. Preliminary laboratory-scale implementation results show the system capability in mitigating assembly errors through dynamic assistive technology selection and iterative feedback learning.

1. Introduction

The manufacturing sector is transforming under Industry 5.0, which integrates advanced technologies with human skills [1,2]. This human-centric paradigm prioritises collaboration between humans and machines to enhance productivity, adaptability, and sustainability [3]. In this context, generative AI (genAI), including large language models (LLMs), drives advancements in manufacturing, such as multimodal guidance, real-time error detection, and task optimisation [4–6]. GenAI is currently showing significant potential within industrial applications and assistive technologies, addressing challenges such as user-specific preferences, error management, and adaptability [7]. LLMs are leading this transformation, enabling natural language-based instructions and adaptive guidance tailored to complex tasks [8–10]. The integration of LLMs with vision-language models could enhance human-robot collaboration (HRC) in manufacturing environments by enabling more adaptive task planning and improving operator efficiency through the alignment of visual and linguistic cues [11,12].

Augmented reality (AR) could complement these capabilities by improving task efficiency, and providing context-aware solutions through AR-assisted digital twins that integrate virtual and physical environments [3,13,14]. Together, genAI and complementary tools like AR could provide a robust framework for adaptive, human-centric manufacturing processes [15].

Despite these advances, state-of-the-art systems often rely on intrusive or immersive interfaces, such as head-mounted displays (HMDs), which could disrupt workflows with negative effects on cognitive load [11,16], and challenges remain in scaling these systems to dynamic and complex industrial contexts [12]. In addition, there is a lack of solutions tailored to operators with diverse cognitive needs [1,5]. In response, this research proposes a non-intrusive, genAI-based multimodal assistive system for assembly tasks. The system is designed to provide real-time, adaptive guidance through integrated error detection, multimodal feedback (including voice, visual cues, and robotic support), and communication channel prioritisation tailored to individual operator needs. Such a system is intended to help reduce assembly errors, promote human-machine collaboration, and improve overall operational efficiency.

2. Methodology

The methodological framework represented in Fig. 1 is structured across interconnected layers, integrating hardware, software, digital systems, and human-machine interaction.

2.1. Data management layer

The Data Management Layer (DML) handles various types of data, including personal user information (such as preferences and cognitive profiles), assembly task data (such as instructions), and error history logs covering operator- and workpiece-related issues, all stored in a centralised database. Statistical error analysis within this layer derives key metrics to identify critical assembly steps. These metrics are evaluated against predefined criteria to determine if proactive communication with the operator is required [17]. If the criteria are satisfied, the relevant information is automatically integrated into a warning prompt, which is transmitted to the genAI Model. This model generates a tailored warning message for the specific situation by processing multimodal data and dynamically adapting outputs [6].

2.2. Human layer

The Human Layer represents the operator active role in the assembly process. This layer is central to the system, as operators

* Corresponding author.
E-mail address: alessandro.simeone@polito.it (A. Simeone).

https://doi.org/10.1016/j.cirp.2025.04.061
0007-8506/

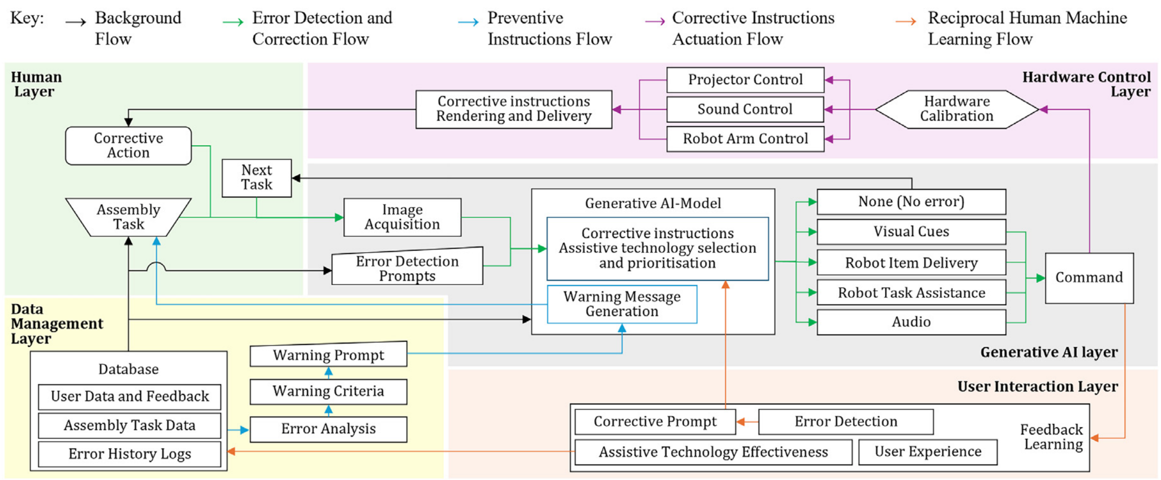

Fig. 1. Methodological framework.

perform assembly tasks and engage with the assistive technologies provided. Operators receive task-specific instructions, error warnings, and corrective guidance through multimodal communication channels, as discussed in the following sections.

2.3. Generative AI layer

During assembly, multiple cameras capture top and side views of the workbench. Digital images and video streams are transmitted via a Python-based Application Programming Interface (API) to pre-trained genAI models (e.g., ChatGPT) accessed through standard APIs. The system employs reference-based image analysis against predefined reference images, with updatable prompts tailored to task-specific conditions. No additional training or fine tuning is required; each error detection session is initiated independently without carry-over from previous interactions.

Within this layer, structured prompts verify part selection, assess tool suitability and evaluate assembly accuracy using predefined instructions and reference images [18]. The LLM exploits its dual object detection and generative capabilities by analysing visual data against these references [19]. In addition, it performs generative functions by providing detailed error descriptions and creating tailored corrective and preventive instructions to guide the operator when an error is detected or when a step is classified as critical. This diagnosis is relayed back through the API to initiate tailored corrective instructions; if no error is identified, a green light is displayed, authorising progression to the subsequent assembly phase [20]. In addition to the tasks mentioned above, assistive technologies are selected and prioritised by analysing operator-specific profiles and contextual data. The analysis considers key inputs such as task requirements, communication preferences, and neurodiversity-related details (if applicable) obtained from operator profiles or medical certifications. Based on this information, the system evaluates and ranks assistive options, according to their effectiveness in addressing the operator specific needs. After finalising the prioritisation, the Generative AI Layer encodes these decisions into commands and transmits them to the Hardware Control Layer.

2.4. Hardware control layer

This layer actuates the corresponding hardware components, such as earplug audio systems, mobile projectors, and robotic arms, to deliver the chosen assistive support. Commands from the Generative AI Layer are executed via a structured API, enabling precise control of each device [21]. The projector displays visual cues directly on the workspace or on the workpiece, the audio system delivers context-specific auditory instructions, and the robotic arm performs targeted actions, including item delivery or task execution [22].

2.5. User interaction layer

This layer supports feedback learning via reciprocal human-machine learning [1], enabling the system to iteratively improve error detection and assistive technology selection. In the event of error misclassification, the operator interacts with the genAI Model via text or voice, generating an updated prompt to refine error detection performance. For assistive technology ineffectiveness, feedback is recorded and sent to the DML to update the user profile.

3. System architecture and implementation

Fig. 2 illustrates the laboratory-scale setup integrating error detection, assistive technology selection, calibration, and delivery for assembly tasks. Task instructions are displayed to operators under normal conditions to guide assembly execution. Error detection uses two Logitech C505 HD webcams: one overhead for a top-down view and one lateral for monitoring operator actions. Real-time image streams are processed in the genAI Layer via a Python-based API,

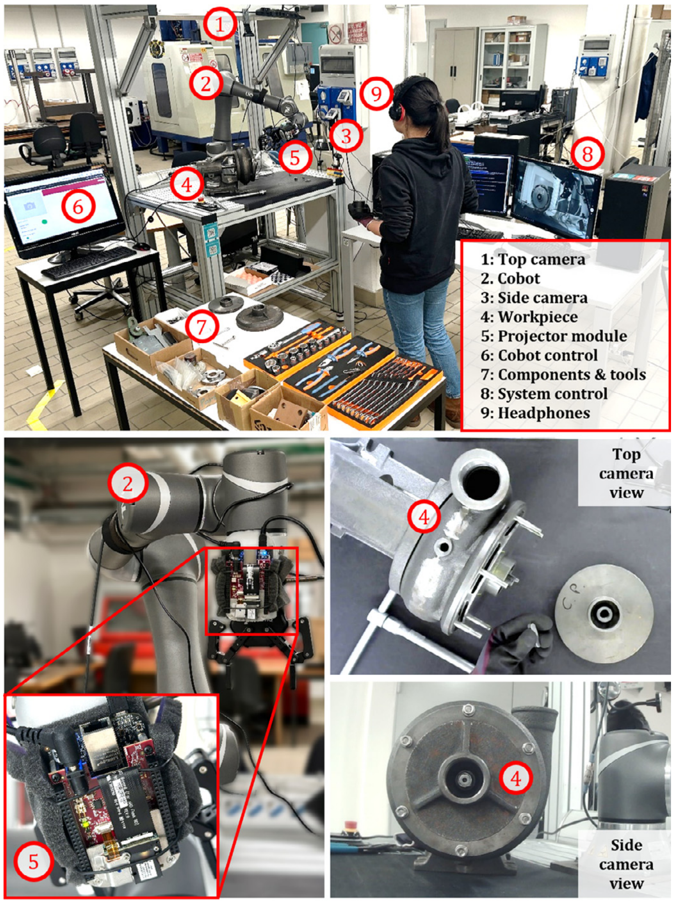

Fig. 2. Experimental setup.

Please cite this article as: A. Simeone et al., Conceptualisation of a multimodal, non-intrusive, generative AI-based assistive system for assembly, CIRP Annals - Manufacturing Technology (2025), https://doi.org/10.1016/j.cirp.2025.04.061

where structured prompts compare captured images with reference data to verify components, tools, and outcomes. Preloaded textual and photographic assembly instructions guide this process. Images (1280 × 720 px) are optimised with OpenCV calibration routines for exposure, white balance, and focus, ensuring reliability under varying lighting. As shown in Fig. 1, discrepancies trigger the genAI Model to process prompts to determine suitable assistive feedback.

The selection and prioritisation of assistive technologies in the Generative AI Layer are operationalised through prompts incorporating dynamic analysis of contextual error data, operator-specific profiles (communication preferences and/or requirements) and task complexity. Commands from this layer direct hardware actions, under DML oversight. The system selects the appropriate modality (e.g., visual cues, auditory instructions, or robotic assistance) based on the operator profile, and generates corrective outputs to address identified discrepancies. If the initial modality proves ineffective, real-time adjustments are performed by iteratively updating the operator profile stored in the DML. For example, if auditory instructions prove ineffective, the system may provide visual cues or projected messages instead. Operator profiles are iteratively updated based on task outcomes and interactions, enabling continual refinement of assistance. Figs. 3 and 4 illustrate key elements of the system workflow and robotic actions (see also the video available in the Supplementary Material), including the generation and projection of visual instructions via the cobot-mounted projector (Fig. 3) and robotic assistance for a part removal task (Fig. 4). Auditory instructions and robotic component handling are not explicitly shown in this study, as their flow has already been detailed in previous research [1,10], with this work focusing on the integration of genAI for real-time error detection and adaptive multimodal guidance.

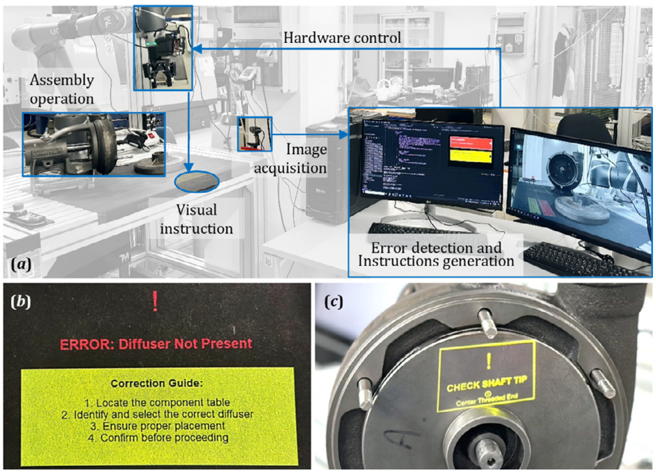

Fig. 3. Workflow for instructions generation (*a*) with detailed view of visual instructions on workbench (*b*) and workpiece (*c*).

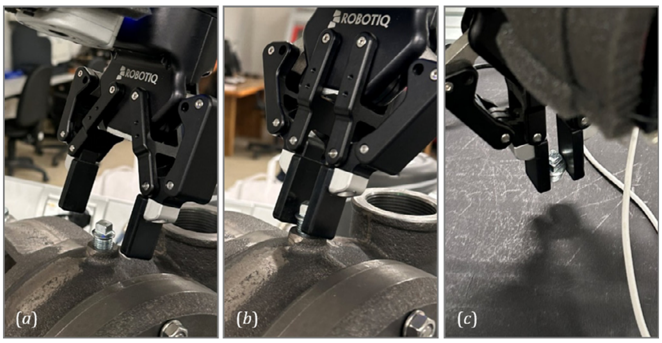

Fig. 4. Cobot assisted task (part removal): alignment (*a*), gripping and extraction (*b*), clearance (*c*).

3.1. System calibration

Calibration ensures precision and synchronisation across auditory, visual, and robotic modalities for effective task assistance [23]. The visual guidance system used a DLP® LightCrafter™ Display 2000 projector, mounted on an OMRON TM5–900 collaborative robot, to deliver task-specific visual cues dynamically aligned with the cobot movements. Corrective instructions generated by the genAI model were converted into SVG format using Python [24], then transformed by CairoSVG into JPEG or PNG formats, depending on transparency requirements, for projection compatibility [25].

Auditory instructions were delivered via high-fidelity earphones using a text-to-speech system that adapted voice parameters (e.g., pitch, tone, speed) to suit environmental conditions [10]. The system processes textual commands through a command mapping module that translates them into TMflow Project ID numbers, transmitted as Transmission Control Protocol (TCP) messages to trigger pre-programmed motion sequences in the robot controller. Execution is verified through status feedback from the TMflow system, confirming completion of each action.

The cobot dynamically adjusts the projector alignment, with calibration performed via the Inter-Integrated Circuit (I2C) protocol. TCP/IP protocols are used to guarantee reliable communication and control of robotic tasks, including component delivery and misalignment correction, while Secure Shell (SSH) connections managed configuration and data exchange [26]. Python scripts handled an event queue to manage and execute predefined actions such as recalibrating the projector, updating auditory instructions, and adjusting robotic tasks, warranting synchronisation and adaptive assistance.

3.2. Feedback learning

Concerning the error detection aspects of feedback learning, at this stage of research, to address AI erroneous outputs in (*i*) error detection, (*ii*) error description, and (*iii*) instruction generation, a feedback-driven corrective mechanism was implemented. Human operators provided corrective prompts, including the input, erroneous output, corrective instruction, corrected output, and metadata such as timestamps, error categories, and operator identifiers. These human-machine interactions were stored in the DLM database, potentially enabling systematic retrieval of past corrections. The database served as an external memory by updating the corrective prompts for each new session with the operator feedback [27]. Importantly, each error detection check is conducted in a new dialogue session, ensuring that although operator feedback is incorporated into the corrective prompt, the outcomes remain independent. User feedback is appended to the corrective prompt, thereby enabling immediate correction of GenAI errors. Concurrently, the same feedback is integrated into the DML to inform subsequent error-detection prompts.

Beyond addressing error detection, the framework also handled technological ineffectiveness and user-specific preferences. When assistive technologies failed, the system dynamically switched modalities based on operator preferences stored in the database. Feedback, logged via a Command-Line Interface (CLI), triggered automatic selection of the next-highest ranked option. Python scripts using the pandas library managed real-time preference ranking and updates, while an event-driven architecture activated by feedback or modality failures ensured responsive coordination and adaptive adjustments [28,29].

4. Preliminary results

The system was tested on the assembly of a cast iron horizontal bare-shaft centrifugal pump, focusing on 14 assembly steps, including 4 critical steps identified for their high error rates and impact on quality and time. Three operators with varying experience levels were monitored during implementation. The following subsections examine the use of genAI for error detection and the effects of the various assistive technologies.

4.1. Generative AI-based error detection

The error detection capabilities of genAI models were assessed using an optimised prompt (excerpt shown in Fig. 5, full version in

Please cite this article as: A. Simeone et al., Conceptualisation of a multimodal, non-intrusive, generative AI-based assistive system for assembly, CIRP Annals - Manufacturing Technology (2025), https://doi.org/10.1016/j.cirp.2025.04.061

```
The type of this task is "{task_column['operation name']}". The
first image shows the correct assembly reference, and the second
image shows the real-time operation result. Based on the two
images and the instruction "{task_column['work instructions']}",
please analyse whether the operator correctly completed this
step.
Then, output a Detection Log, which includes:
# "Error" or "Correct"
# Specific error description details
# Clear and concise correction instructions, if an error is
present.
```

Fig. 5. GenAI prompt excerpt for error detection and correction.

the Supplementary Material) designed for recognition tasks such as selecting tools and components (e.g., T-type wrench and diffuser) and comparing assemblies. Success rate was defined as the proportion of tasks where the model correctly identified components, tools, or operations without requiring further correction on the first attempt. Accuracy in error descriptions and corrective instructions was assessed based on the model ability to precisely characterise errors and generate actionable guidance. Two genAI models, ChatGPT v4.0 and Claude v3.5, were tested. ChatGPT achieved an 84 % success rate, with 16 % false positives, (non-existent errors flagged) and no false negatives. This ensured that all assembly mistakes were identified and no incorrect actions were allowed to proceed. Whenever ChatGPT produced an erroneous output, the operator supplied a corrective second prompt (see Supplementary Material) instructing the LLM of its mistake, resulting in 100 % rectification of misclassified instances in subsequent attempts. ChatGPT also achieved 84 % accuracy in describing assembly errors and generating corrective instructions on the first attempt. All the remaining issues, confined to the same instances as detection errors, were fully resolved after submitting the corrective prompts. Claude, tested under identical conditions, underperformed compared to ChatGPT despite iterative prompt adjustments. While these results are limited to the specific case considered, they highlight the critical role of prompt design in maximising error detection performance. The approach applies the feedback learning mechanism outlined in Section 3.2, incrementally improving performance by integrating operator corrections during runtime. This dynamic refinement eliminates the need for retraining, unlike static models such as YOLO [30], which require extensive retraining to adapt to new contexts.

4.2. Assistive technology effectiveness

Of the 12 critical steps (4 steps times 3 operators), warning instructions successfully mitigated errors in 11 cases (91.7 %), highlighting the system effectiveness in reducing critical mistakes. A total of seven visual warnings and five auditory warnings were delivered; one visual warning proved ineffective and was subsequently corrected through robotic intervention. Table 1 reports the average processing times ($\overline{x}$) and standard deviations (σ) for tasks attributable to genAI actions adopting ChatGPT. These processing times, on the order of seconds, demonstrate consistency and low variability, even in a laboratory-scale setup. Computation times for error detection, instruction generation, and rendering appear minimally disruptive to the nominal assembly duration (~25 min). Compared to manual assistance by a human supervisor, as per [1,10], these results suggest that the developed system can reduce assembly errors with expected positive impact on time and, in turn, time-associated costs. Nevertheless, error detection times, though optimised for this experimental configuration, remain relatively high compared with state-of-the-art deep learning technologies [30]. As shown in Table 1, the delivery time for vocal warnings and corrective instructions is not explicitly reported, as it depends on case-specific factors, particularly the length of the instructions (and thus, the duration of the voice message playback). Similarly, robot assistance, encompassing actions such as item movement and task execution, exhibited time spans ranging from 25 to 32 s during preliminary tests. Longer durations were associated with tasks executed by the robot, while shorter durations pertained to item delivery. These recorded times, obtained under laboratory experimental conditions, reflect the cobot characteristics, including payload capacity, speed, and safety features (e.g., compliance with interaction protocols).

Table 1
Processing times for generative AI tasks

Phase	Task	$\overline{x}$ (s)	σ (s)
Pre-assembly	Critical task identification	<0.01	<0.01
	Warning mode selection	0.53	0.27
	Vocal warning generation	2.68	1.21
	Visual warning generation	2.38	0.39
Post-assembly	Error detection	5.89	1.28
	Corrective instruction mode selection	0.97	0.35
	Vocal corrective instructions generation	3.45	0.28
	Visual corrective instructions generation	4.01	0.47
	Robot movement instructions generation	0.90	0.18
	Robot task instructions generation	0.71	0.11
Follow-up check	Error detection	4.57	1.11

5. Conclusions

This study proposed a generative AI-based multimodal assistive system for assembly that integrates real-time error detection, adaptive guidance, and personalised operator support. The projection system delivers customisable visual instructions, which can range from static images to animations and short videos. Unlike approaches relying on intrusive or immersive technologies such as virtual or AR headsets, this study focuses on non-intrusive solutions and fully integrates genAI into all phases of operator support. In-situ projection could offer performance comparable to AR HMDs for skill-based tasks while reducing cognitive load [31]. Moreover, prolonged use of HMDs can increase cognitive strain, reinforcing the advantages of projection-based systems [32]. The system maximises human-machine interaction by delivering preventive warning messages before assembly, corrective instructions during assembly, and automatic robotic interventions when these solutions prove ineffective. The study is subject to limitations regarding current genAI models, hardware and software, which are not yet optimised for industrial applications. Despite the limitations of the laboratory scale, the modular architecture of the system is expected to facilitate integration into industrial workstations and to enable adaptation to various assembly tasks. In addition, this multimodal system holds promise for enhancing the inclusion of neurodiverse operators, leveraging its ability to adapt to diverse cognitive profiles and communication preferences. At this stage, robotic intervention operates in a reactive mode. Future research could develop a proactive approach based on monitoring operator perception and cognition. However, this shift raises ethical issues regarding privacy, consent and the balance between user empowerment and misuse, which may necessitate rethinking assistive systems from a person-centred to a task-centred paradigm.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Alessandro Simeone: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Yuchen Fan:** Writing – original draft, Visualization, Software, Investigation, Data curation. **Dario Antonelli:** Validation, Investigation, Conceptualization. **Paolo C. Priarone:** Writing – review & editing, Visualization, Methodology, Formal analysis. **Luca Settineri:** Writing – review & editing, Supervision, Resources, Methodology.

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.cirp.2025.04.061.

References

[1] Simeone A, Fan Y, Antonelli D, Catalano AR, Priarone PC, Settineri L (2024) Inclusive manufacturing: a contribution to assembly processes with human-machine reciprocal learning. *CIRP Annals* 73(1):5–8.
[2] Zheng P, Li S, Fan J, Li C, Wang L (2023) A collaborative intelligence-based approach for handling human-robot collaboration uncertainties. *CIRP Annals* 72 (1):1–4.
[3] Yin Y, Zheng P, Li C, Wang L (2023) *A state-of-the-art survey on Augmented reality-assisted Digital Twin for futuristic human-centric industry transformation*. *Robotics and Computer-Integrated Manufacturing* 81:102515.
[4] Lanighan MW, Youngquist O (2024) Online fault detection in manipulation tasks via generative models. In: 2024 IEEE International Conference on Robotics and Automation (ICRA), *IEEE, 802–808*.
[5] Aristeidou C, Dimitropoulos N, Michalos G (2024) Generative AI and neural networks towards advanced robot cognition. *CIRP Annals* 73(1):21–24.
[6] Gao RX, Krüger J, Merklein M, Möhring H-C, Váncza J (2024) Artificial Intelligence in manufacturing: state of the art, perspectives, and future directions. *CIRP Annals* 73(2):723–749.
[7] Fu B, Hadid A, Damer N (2025) Generative AI in the context of assistive technologies: trends, limitations and future directions. *Image and Vision Computing* 154:105347.
[8] Jonek M, Bast M, Manns M (2024) Manual assembly planning with AI image generators. *Procedia CIRP* 130:139–144.
[9] Wiemann R, Terei N, Raatz A (2024) Large language model for assisted robot programming in micro-assembly. *Procedia CIRP* 130:244–249.
[10] Fan Y, Simeone A, Antonelli D, Caggiano A, Priarone PC, Settineri L (2025) A framework for digital assembly instructions as a step towards manufacturing inclusiveness. *Procedia CIRP* 132:116–121.
[11] Fang W, Chen L, Zhang T, Chen C, Teng Z, Wang L (2023) Head-mounted display augmented reality in manufacturing: a systematic review. *Robotics and Computer-Integrated Manufacturing* 83:102567.
[12] Liu S, Zhang J, Wang L, Gao R X (2024) Vision AI-based human-robot collaborative assembly driven by autonomous robots. *CIRP Annals* 73(1):13–16.
[13] Eversberg L, Lambrecht J (2023) Evaluating digital work instructions with augmented reality versus paper-based documents for manual, object-specific repair tasks in a case study with experienced workers. *The International Journal of Advanced Manufacturing Technology* 127(3–4):1859–1871.
[14] Pang J, Zheng P, Fan J, Liu T (2025) Towards cognition-augmented human-centric assembly: a visual computation perspective. *Robotics and Computer-Integrated Manufacturing* 91:102852.
[15] Gkournelos C, Konstantinou C, Makris S (2024) An LLM-based approach for enabling seamless Human-Robot collaboration in assembly. *CIRP Annals* 73(1):9–12.
[16] Potthoff L, Naussedat R, Gunnemann L (2024) Exploring Generative AI's role in manual assembly: application potentials and use concepts. *Procedia CIRP* 130:194–199.
[17] Wang Z, Kiyokawa T, Sera I, Yamanobe N, Wan W, Harada K (2023) Error correction in robotic assembly planning from graphical instruction manuals. *IEEE Access* 11:107276–107286.
[18] Kim Y, Kim D, Choi J, Park J, Oh N, Park D (2024) A survey on integration of large language models with intelligent robots. *Intelligent Service Robotics* 17(5):1091–1107.
[19] Long Z, Killick G, McCreadie R, Aragon-Camarasa G (2024) RoboLLM: robotic vision tasks grounded on multimodal large language models. In: 2024 IEEE International Conference on Robotics and Automation (ICRA), *IEEE, 12428–12435*.
[20] Calaon M, Chen T, Tosello G (2024) Integration of multimodal data and explainable artificial intelligence for root cause analysis in manufacturing processes. *CIRP Annals* 73(1):365–368.
[21] Blankemeyer S, Wendorff D, Raatz A (2024) A hand-interaction model for augmented reality enhanced human-robot collaboration. *CIRP Annals* 73(1):17–20.
[22] Konstantinou C, Antonarakos D, Angelakis P, Gkournelos C, Michalos G, Makris S (2024) Leveraging generative AI prompt programming for Human-robot Collaborative Assembly. *Procedia CIRP* 128:621–626.
[23] Schmitt RH, Wolfschläger D, Woltersmann J-H, Stohrer L (2024) Measurability of quality characteristics identified in latent spaces of generative AI models. *CIRP Annals* 73(1):389–392.
[24] Ye Y, Hao J, Hou Y, Wang Z, Xiao S, Luo Y, Zeng W (2024) Generative AI for visualization: state of the art and future directions. *Visual Informatics* 8(2):43–66.
[25] CairoSVG documentation. https://cairosvg.org/. Accessed January 2025.
[26] Zafir EI, Akter A, Islam MN, Hasib SA, Islam T, Sarker SK, Muyeen SM (2024) Enhancing security of Internet of Robotic Things: a review of recent trends, practices, and recommendations with encryption and blockchain techniques. *Internet of Things* 28:101357.
[27] Mustapha KB (2025) A survey of emerging applications of large language models for problems in mechanics, product design, and manufacturing. *Advanced Engineering Informatics* 64:103066.
[28] Antonelli MG, Beomonte Zobel P, Manes C, Mattei E, Stampone N (2024) Emotional intelligence for the decision-making process of trajectories in collaborative robotics. *Machines* 12(2):113.
[29] Zhao F, Gao W, Liu T, Jiang Z-P (2023) Event-triggered robust adaptive dynamic programming with output feedback for large-scale systems. *IEEE Transactions on Control of Network Systems* 10(1):63–74.
[30] Hussain M, Khanam R (2024) In-depth review of YOLOv1 to YOLOv10 variants for enhanced photovoltaic defect detection. *Solar* 4(3):351–386.
[31] Dhiman H, Röcker C (2024) Does the medium matter? A comparison of augmented reality Media in instructing novices to perform complex, skill-based manual tasks. *Proceedings of the ACM on Human-Computer Interaction* 8(EICS):1–28.
[32] Van der Burg E, Ledegang WD, Kooi FL, Houben MMJ, Groen EL (2025) Attentional tunneling in pilots during a visual tracking task with a head mounted display. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 67(1):63–78.

Please cite this article as: A. Simeone et al., Conceptualisation of a multimodal, non-intrusive, generative AI-based assistive system for assembly, CIRP Annals - Manufacturing Technology (2025), https://doi.org/10.1016/j.cirp.2025.04.061