SCIENTIFIC OASIS

Decision Making: Applications in Management and Engineering

Journal homepage: www.dmame-journal.org
ISSN: 2560-6018, eISSN: 2620-0104

A Game-Theoretic and Multimodal Interaction Framework for Collaborative Robots in Smart Manufacturing

Dongdong Chu[1], Shulan Yu[2,*], Yuhang Ling[3], Yang Zhao[4], Jianhong Zhang[5]

1 Master, School of Mechanical and Electronic Engineering, Nanjing Forestry University, China, 210037. Email: xupengcheng@njfu.edu.cn
2 Professor, School of Home and Industrial Design, Nanjing Forestry University, China, 210037. Email: 372468296@qq.com
3 Master, School of Mechanical and Electronic Engineering, Nanjing Forestry University, China, 210037
4 Master, School of Mechanical and Electronic Engineering, Nanjing Forestry University, China, 210037
5 BA, Professor, Department of Mental Health Education, Nanjing Audit University, China, 211815

ARTICLE INFO

Article history:
Received 5 March 2024
Received in revised form 30 May 2024
Accepted 20 June 2024
Available online 29 June 2024

Keywords:
Collaborative Robots, Human-Robot Interaction, Multimodal Perception, Game-Theoretic Optimization, Smart Manufacturing Systems.

ABSTRACT

Industrial assembly settings have encountered novel challenges following the integration of collaborative robots (cobots), particularly in maintaining a balance between effective human-robot interaction and operational efficiency. Interpreting and implementing complex human behaviours remains a significant difficulty, especially in conventional operations conducted under dynamic and unpredictable conditions. A critical requirement has emerged for the development of intelligent interfaces capable of autonomously regulating systems while facilitating seamless collaboration between human operators and robotic agents. This study focuses on a multimodal perception approach enhanced by game-theoretic optimisation, which has been shown to improve cobots' responsiveness and strategic flexibility. Researchers have chosen this method due to its capacity to model interactive behaviours in variable environments. By integrating vision, auditory, and tactile sensing technologies, cobots are equipped to accurately interpret human communicative cues, even when operating with limited capabilities. The proposed framework utilises game theory to model the strategic and dynamic interactions between humans and robots, thereby addressing the dual challenges of task allocation and decision-making. This system supports optimal coordination and prompt conflict resolution, while enabling real-time behavioural adjustments based on utility-driven outcomes. Its efficacy has been validated through both simulated and practical implementations in industrial assembly scenarios, confirming its potential to enhance collaborative efficiency and safety while maintaining adaptability. Overall, this research presents a strategic design framework for cobot interaction that emphasises perceptual processing, contributing to substantial advancements in the evolution of intelligent manufacturing systems.

1. Introduction

Contemporary industrial assembly processes have significantly advanced through the integration of cobots, marking a vital step forward in the evolution of intelligent manufacturing systems. Cobots

* *Corresponding author.*
E-mail address: 372468296@qq.com

https://doi.org/10.31181/dmame7120241433

are specifically designed with integrated safety mechanisms that enable their operation in close proximity to human workers, fostering the emergence of shared workspaces [29]. This progression is largely driven by growing demands for greater production flexibility, responsiveness, and adaptability, particularly in response to trends such as mass customisation and frequent product redesigns [7]. While cobots contribute consistent precision and physical endurance, human operators remain essential for managing variability and addressing exceptions in these dynamic settings [16]. However, successful human-robot collaboration entails more than mere spatial coexistence; it requires the establishment of dependable interaction frameworks that support real-time communication, coordination, and mutual comprehension between human and robotic agents [12]. This requirement becomes especially critical in scenarios where even minor deviations in timing, perception, or intent may result in productivity loss or pose safety risks [23].

A fundamental enabler of such sophisticated collaboration is multimodal perception—the robot's capacity to interpret human actions and environmental cues through various sensory channels. These modalities include visual input, speech recognition, tactile and force feedback, and physiological indicators such as muscle activity and gaze direction [22]. Although each sensory input provides limited information individually, their integration yields a comprehensive, context-rich understanding of the environment [26]. For instance, a cobot equipped with computer vision and depth sensing technologies can identify the position and orientation of components, while responding to human gestures or verbal prompts to initiate cooperative actions [9]. Such non-intrusive, intuitive interaction reduces the cognitive burden on human operators and enhances the naturalness of collaboration [28]. Nevertheless, the fusion of heterogeneous sensor data presents significant technical challenges due to factors such as temporal misalignment, signal noise, ambiguity, and sensor reliability [10]. Furthermore, as human intentions are often probabilistic, cobots must be capable of reasoning under uncertainty and adapting to dynamic conditions in real time.

To support such advanced collaboration beyond perceptual capabilities, it is imperative to establish strategic decision-making frameworks and operational protocols that dictate task allocation, resource sharing, and conflict resolution [30]. Traditional planning and control methodologies often prove inadequate in dynamic, multi-agent environments involving human participants [2]. Game theory offers a robust mathematical foundation for modelling interactions among rational agents with potentially cooperative or conflicting objectives [4]. Within the context of industrial assembly, game-theoretic models enable optimal task distribution, anticipation of human responses, and negotiation between competing goals such as speed and accuracy [21]. For example, in a Stackelberg game framework, the human acts as the leader and the cobot as the follower, allowing the robot's strategy to adapt based on predicted human actions. Alternatively, cooperative game theory supports joint decision-making where both agents aim to maximise shared utility [15]. Integrating such models within the control architecture redefines the role of the cobot from a passive tool to an adaptive partner—one capable of negotiating roles, reallocating tasks dynamically, and aligning its operational strategy with the evolving goals of its human counterpart.

This study proposes a comprehensive framework that combines multimodal perception with game-theoretic configuration to advance cobot interaction design in industrial assembly environments. By leveraging advanced sensory processing alongside strategic, model-based decision-making, the framework enhances cobot autonomy, safety, and cooperative performance. The system undergoes validation through experimental trials and simulation-based assessments conducted in industrial facilities, with performance metrics including task completion time, workflow smoothness, and user satisfaction. The overarching aim of this research is to inform the design of next-generation collaborative robots that can emulate human-like capabilities, enabling them to perform precise assembly tasks in an intuitive and user-friendly manner.

2. Related Works

To develop a suitable interaction framework for collaborative robots within industrial assembly settings, existing research in relevant domains has been systematically reviewed. These investigations primarily concentrate on multimodal perception technologies, human-robot interaction (HRI) design, and game-theoretic optimisation approaches. The exploration of perception techniques includes the utilisation of computer vision, speech recognition, force sensing, and more recent advancements such as deep learning-based sensor fusion and probabilistic modelling methods [19]. These approaches aim to enhance the robot's capacity to interpret human behaviour and environmental cues more accurately.

In parallel, numerous studies address the challenges of task distribution, conflict mitigation, and strategy formulation through the application of game-theoretical models. These models encompass cooperative and non-cooperative games as well as Stackelberg-type formulations to support dynamic and strategic interaction within collaborative settings [5]. While each contribution provides valuable insights, limitations persist. These include computational inefficiencies, limited adaptability to real-time operational contexts, and reduced robustness when processing noisy or ambiguous data [18]. To encapsulate these findings, Table 1 presents a comparative summary of the reviewed studies, highlighting the strengths of the respective methodologies alongside identified areas that warrant further enhancement.

Table 1
Problem Formulation

Author(s)	Techniques Involved	Advantages	Disadvantages
Wang and Jiao [27]	Non-Cooperative Game Theory, Bi-Level Optimization	Efficient for Competitive Task Allocation, Robust in Dynamic Environments	High Computational Complexity, Limited Real-Time Adaptability
Zhang et al. [32]	Collaboration Effectiveness, Complex Operations Allocation	Optimizes Collaboration, Considers Human Limitations	Simplifies Real-World Dynamics, Scalability Issues
Zeng et al. [31]	Task Allocation, Scheduling, Efficiency and Fatigue Optimization	Balances Efficiency and Fatigue, Boosts Productivity	Difficulty in Real-Time Adaptation, Human Fatigue Modelling Challenges
Cai et al. [3]	Task Matching, Ergonomics, Hybrid Assembly Cell	Integrates Ergonomics, Enhances Safety and Productivity	Complex Ergonomic Design, Limited to Non-Specialized Tasks
Alessio et al. [1]	Multicriteria Task Classification, Fuzzy Inference	Flexible, Handles Uncertainty Well	Can be Imprecise, Computationally Intensive

Wang and Jiao [27] present a cognitively intelligent framework for task allocation within human-automation collaboration, employing non-cooperative game theory in conjunction with bi-level optimisation. Their model captures agent interactions in competitive settings and yields optimal task distribution strategies that demonstrate resilience in dynamic industrial contexts. Nonetheless, the computational intensity of their method hinders its real-time responsiveness, thereby limiting its viability in high-speed assembly operations. Zhang et al. [32] propose a collaboration-oriented task allocation strategy aimed at aligning robotic capabilities with human cognitive and physical thresholds. This approach facilitates coordinated task execution by incorporating human-centred considerations. However, despite enhancing coordination and task performance, the strategy oversimplifies real-world dynamics and exhibits scalability limitations when extended to more complex or larger systems.

Zeng et al. [31] advance a task scheduling and allocation model that accounts for human productivity and fatigue, seeking to optimise collaboration through balanced workload distribution.

Their system enhances productivity while mitigating operator fatigue. Yet, its practical effectiveness is constrained by modelling inconsistencies in fatigue representation and a limited capacity for real-time adaptation, particularly under unanticipated operational conditions. Cai et al. [3] examine task alignment and ergonomic integration within human-robot hybrid assembly environments. Their approach aims to promote operator comfort, reduce work-related injuries, and boost group productivity through ergonomic task design. While effective in structured scenarios, this integration demonstrates limited applicability to complex ergonomic settings and tasks requiring extensive customisation or non-specialised group handling. Alessio et al. [1] introduce a fuzzy inference-based system for classifying tasks in collaborative assembly settings involving human-robot interaction. Their framework addresses uncertainty and ambiguity in task definitions by providing a flexible decision-making mechanism. However, dependence on fuzzy logic introduces potential drawbacks: imprecise outcomes and high computational demands in large-scale decision environments can limit its practicality, particularly for real-time applications.

Despite the notable advancements reported in these studies, several persistent limitations remain—chief among them high computational demands, insufficient real-time adaptability, simplistic modelling assumptions, and limited success in handling uncertainty. Most existing frameworks struggle to support dynamic task allocation and real-time decision-making while maintaining ergonomic integrity and operational efficiency. Furthermore, these approaches rarely integrate multimodal perception with game-theoretic optimisation in a cohesive manner, resulting in reduced adaptability in complex, evolving collaborative contexts. To address these shortcomings, the proposed framework integrates robust multimodal sensory data with adaptive game-theoretic models. This combined approach is designed to enhance system flexibility, responsiveness, and context-awareness. It enables dynamic task redistribution, strategic optimisation of collaboration, and ergonomically sound interaction, all with minimal computational overhead. The overarching aim is to construct a scalable, adaptive, and robust solution that enhances the synergistic interaction between humans and robots in industrial assembly operations.

3. Proposed System Model

The collaborative robot interaction framework was developed through the integration of distinct modules, each aligned with the system's key functional layers: perception, decision-making, and execution. Central to this architecture is the Multimodal Perception Layer, which processes real-time inputs derived from a variety of sensor sources, including visual (camera), auditory (microphone), and tactile modalities. These heterogeneous data streams are processed by a Sensor Fusion Engine responsible for temporal alignment and interpretive synthesis, enabling accurate comprehension of both human operator's behaviour and the surrounding environment. The consolidated sensory data are subsequently transmitted to the Human Intention Recognition Unit. This unit applies machine learning algorithms to infer the operator's intentions and to extract relevant contextual information critically to effective task coordination. The overall structure of the proposed system architecture is illustrated in Figure 1.

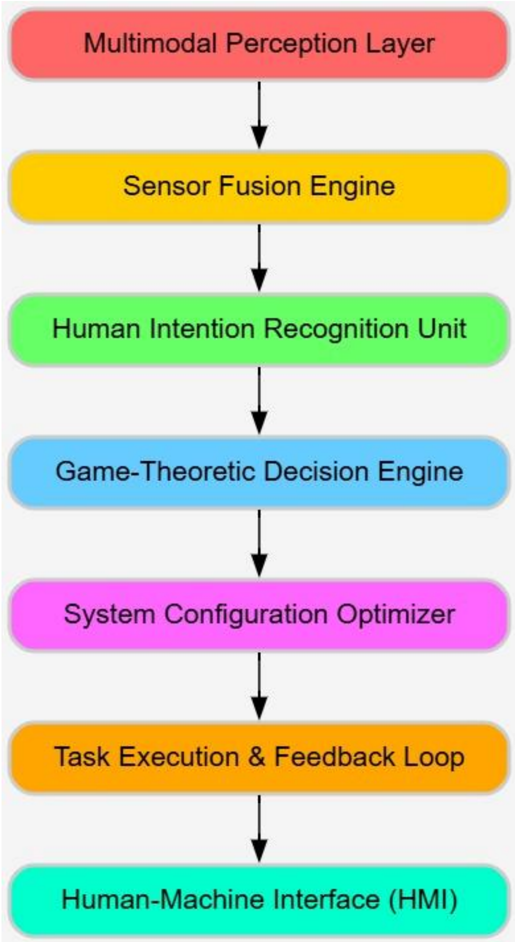

Fig.1: System Architecture

Building upon the established framework, an additional module—the Game-Theoretic Decision Engine—has been integrated to enhance adaptability and strategic collaboration. Within this system, both human and robotic agents are modelled as rational participants in a game-theoretic construction. To effectively manage potential conflicts, the engine evaluates a range of possible actions by estimating associated utility values, subsequently determining optimal strategies using equilibrium-based methodologies such as Nash or Stackelberg equilibria. The resulting decisions are then relayed to the System Configuration Optimiser, which dynamically adjusts robot-specific parameters including task prioritisation, motion planning, and interaction timing, in accordance with real-time system constraints and perceptual inputs. Supported by continuous interaction with the perception layer, the Task Execution Module implements the selected strategies, learning and refining its performance during operation. To foster transparency, trust, and cooperative engagement, visual and auditory feedback is provided to the operator via the HMI. The proposed architecture thereby facilitates adaptive human–robot collaboration, ensuring elevated levels of operational safety, flexibility, and efficiency in evolving industrial assembly settings.

3.1 Step 1: Multimodal Perception Layer

The Multimodal Perception Layer is responsible for acquiring environmental and interactional data through an array of heterogeneous sensors, including visual, auditory, and tactile modalities. These sensors enable the robot to interpret various elements of its operational environment. Visual sensors (such as cameras) provide still images or continuous video streams, auditory sensors capture acoustic signals and human speech patterns, and tactile sensors detect physical contact as well as spatial proximity. By processing these diverse sensory inputs, the robot is able to construct a cohesive and enriched understanding of its surroundings. The fusion of data from multiple sources significantly increases system robustness, as the individual strengths of each sensor type serve to mitigate the limitations of the others. This data integration is achieved via a mathematical approach involving a weighted summation model, as described in equation (1), which ensures the generation of a unified perceptual representation.

$$D_{fused} = W1D_{vision} + W2D_{audio} + W3D_{tactile} \quad (1)$$

Here, $W1, W2, W3$ is defined as the weights assigned to every sensor, $D_{tactile}$, D_{audio} and D_{vision} are defined as the tactile, audio and vision sensors depending on their relevance to the task at hand. The integrated sensory data provides a foundational input for the robot's interpretation of human behaviour, enabling it to modify its actions accordingly (Salehzadeh, Gong, & Jalili, 2022).

3.2 Step 2: Human Intention Recognition

The recognition of human intentions by robots is enabled through the analysis of data derived from sensor fusion processes. This capability is essential in collaborative robotics, as it allows robots to anticipate operator actions and align their responses accordingly. Machine learning, particularly supervised learning techniques and neural networks, underpins this functionality by identifying recurring patterns linked to human motions [6]. Following the fusion of sensory data, the robot applies trained algorithms to interpret operator gestures and movements, such as directional shifts or gesture towards objects, which may signify intentions to grasp, pause, or initiate alternate tasks. These models are trained on annotated datasets of human behaviours to improve their interpretative accuracy over time. During operation, the robot evaluates the likelihood of specific human intentions based on real-time input, allowing it to adjust its actions strategically. This predictive capability enables robots to operate in synchrony with human collaborators, thereby facilitating seamless and responsive joint task execution. The prescient interpretation of human intent forms a crucial foundation for effective and adaptive human–robot teamwork [11].

3.3 Step 3: Game-Theoretic Decision Engine

Decision-making in collaborative robotic systems is governed by the Game-Theoretic Decision Engine, which models both the robot and the human as rational agents seeking to optimise their respective utilities. Within this interactive framework, each agent's decisions directly influence the outcomes experienced by the other, thereby establishing a dynamic interdependence. Game theory provides an effective means of capturing such interactions, as it accounts for the mutual awareness each agent possesses regarding the other's objectives and behaviours. For example, upon recognising a human operator's intent to manipulate an object, the robot may elect to reposition itself to allow the human to complete the task safely. Conversely, humans may adapt their trajectories to avoid the robot's operational path. Through continuous feedback and evaluation, the decision engine dynamically reconfigures task strategies to enhance the utility of both participants. This iterative process not only mitigates potential conflicts but also supports the mutual adjustment of actions, fostering seamless coordination. The result is a cooperative working environment in which both entities operate with minimal disruption and enhanced efficiency [34].

3.4 Step 4: System Configuration Optimization

Upon determining the optimal strategies, the robot advances to system configuration optimisation through the game-theoretic engine. This process involves analysing internal parameters such as velocity, trajectory, and energy consumption to enhance execution performance. Adjusting these parameters provides an effective means of balancing task demands, including performance, safety, and operational efficiency [8]. For example, the robot may reduce its speed in confined spaces to minimise collision risks, even if this results in slightly longer task completion times. The optimisation framework also considers the physical limitations of the robot's components, such as motor capacity, battery life, and maximum allowable operating speed. To satisfy essential performance criteria—such as task duration and energy efficiency—the robot must maintain operation costs within acceptable limits [33]. Through this phase, the system achieves optimal operational performance and safety during task execution, thereby supporting the handling of complex operations while preserving the robot's longevity and safety standards [24]. This adaptability enables the robot to adjust its parameters online in response to varying environmental conditions, ensuring optimal functionality across diverse operational settings.

3.5 Step 5: Task Execution and Feedback Loop

Within the Task Execution and Feedback Loop, the robot carries out assigned tasks while dynamically adapting its behaviour to changes in the environment, human operator movements, and evolving operational requirements. It utilises optimised parameters and strategies during task performance, guided by real-time sensory feedback from visual, auditory, and tactile inputs to regulate its actions in response to environmental stimuli. Via the Game-Theoretic Decision Engine, the robot effectively engages with human operators, enabling it to adjust its activities according to their movements and intentions, thereby enhancing work efficiency and preventing conflicts. Continuous learning throughout task execution allows the robot to improve its functional capabilities by autonomously adapting to unforeseen circumstances and refining its strategies [14]. Through the Human–Machine Interface (HMI), operators receive system status updates and retain control access, permitting them to intervene during delays or system malfunctions. This ongoing feedback mechanism fosters operational excellence and heightens safety by ensuring effective collaboration between humans and robots throughout task completion [25].

3.6 Step 6: Human-Machine Interface (HMI)

In the final stage of the process, a user-friendly HMI is implemented to assist the operator in monitoring the robot's activities and making operational decisions [13]. The system delivers real-time updates on the robot's functionality, current task status, and any issues arising during task execution through the HMI. This platform features a graphical user interface that presents the robot's ongoing task progress, such as advancement towards objectives, completion of assigned operations, and requests to modify the task status [17]. Important alerts are conveyed to human operators via audio signals alongside visual cues when delays or sensor malfunctions occur. The HMI allows the operator to modify the robot's operation modes and parameters to ensure smooth task execution. Human–robot interaction is characterised by seamless connectivity, enabling the operator to intervene and control robot actions during critical moments. This component empowers the human operator to maintain control throughout the collaboration and task performance, establishing a direct communication channel that enhances the efficiency and effectiveness of joint human–robot operations [20].

4. Theoretical Framework

The game-theoretical models employed in this research are essential for managing the intricate strategic interactions between human operators and collaborative robots within smart manufacturing settings. These models offer a rigorous mathematical basis to represent dynamic decision-making, allowing cobots to optimise task distribution and adapt responsively to human behaviours. When integrated with multimodal perception, which encompasses visual, auditory, and tactile sensory inputs, the game-theoretic framework enhances the accurate recognition and interpretation of human communication signals, thereby improving real-time collaboration. This methodology supports maximising cooperative gains and resolving conflicts efficiently through the continuous adjustment of strategies informed by utility functions. As a result, embedding game theory within the multimodal interaction architecture fosters improved responsiveness, operational effectiveness, and safety in industrial assembly processes, substantiating its role as the primary analytical tool in this study.

5. Performance Evaluation

This section validates the integration of multimodal perception with game-theoretic optimisation by evaluating their performance in both simulated environments and real industrial assembly scenarios. The proposed approach was assessed against conventional methods with respect to safety performance and adaptability in task execution. Simulation results demonstrated that the proposed method achieved faster task completion times while reducing safety incidents compared to baseline systems. Enhanced human–robot interaction was facilitated by the fusion of visual, auditory, and tactile sensors, enabling more accurate interpretation of human intentions. The game-theoretic optimisation model further supported real-time decision-making, optimising the environment's adaptability in managing conflict resolution and task allocation under dynamic conditions. Testing in actual industrial settings confirmed that the proposed method surpassed traditional approaches by delivering improved collaboration effectiveness and operational flexibility. Overall, the findings indicate that the system substantially enhances human–robot interaction, outperforming existing techniques. Table 2 presents the principal simulation parameters employed to evaluate the method's performance.

Table 2
Simulation Parameters used for evaluating the Proposed Hybrid Cloud Resource Allocation Framework

Parameter	Value / Description
Simulation Environment	MATLAB R2023b / Python 3.10
Total Simulation Time	60 Minutes Per Run
Time Step Interval	1 Minute
Number of Agents	10 (5 Robots, 5 Humans)
Task Levels	10 (Level 1 to Level 10)
Iterations	10 Per Task Level
Performance Metric	Efficiency, Task Completion Rate, Processing Time Share
Resource Allocation Type	Dynamic (Based on Task and Capabilities)
Robot Processing Speed	2x Human Average
Human Flexibility Index	0.8
Collaboration Mode	Parallel & Sequential Strategies
Evaluation Criteria	Completion Time, Accuracy, Utilization Efficiency

Figure 2 illustrates the efficiency of human and robotic assembly work within industrial production, specifically analysing task allocation methods employed in human–robot collaboration. The robotic system requires 2 units of time to complete Task 1, whereas human workers need 3 units to achieve a comparable level of performance. Conversely, for Task 2, the situation is reversed: robots incur a cost of 4 units while human labour requires only 2 units to complete the task.

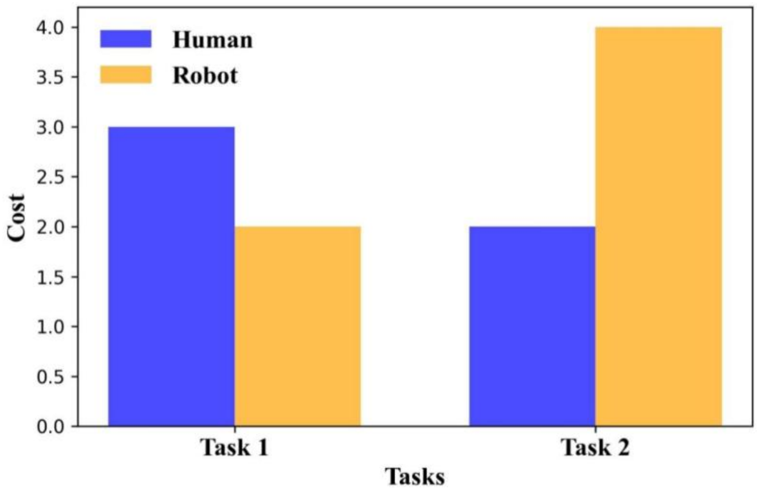

Fig.2: Cost Analysis

These variations in task performance efficiency across different task types underscore the importance of tailored task allocation strategies. The research aims to enhance collaborative robot (cobot) adaptability through the application of game-theoretic optimisation supported by multimodal perception. Real-time task allocation is facilitated by the proposed system, which utilises dynamic game models processing data from visual, auditory, and tactile sensors. Such intelligent frameworks are critical, as static or fixed decision-making approaches frequently lead to suboptimal resource utilisation and diminished efficiency, as evidenced by the data presented. This study contributes to the advancement of next-generation manufacturing by developing intelligent, perception-driven cobot systems. The corresponding heatmap is displayed in Figure 3.

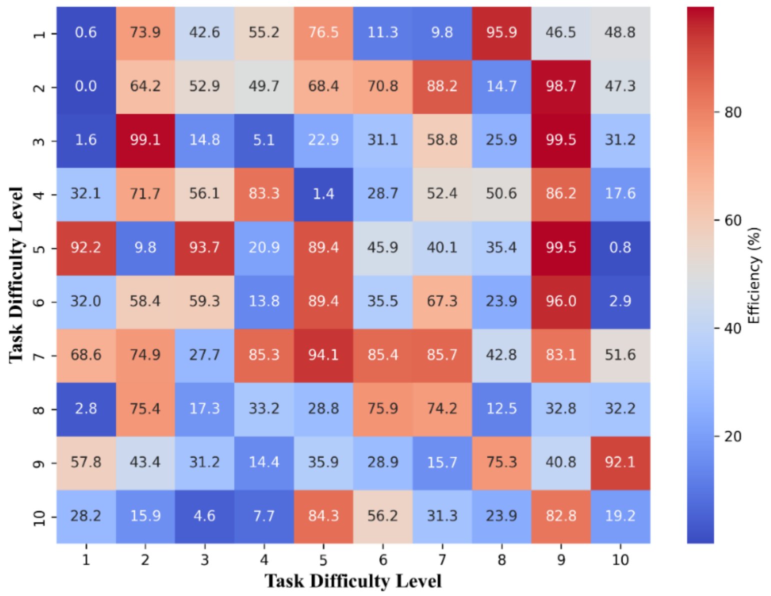

Fig.3: Heatmap Analysis

Figure 4 depicts the progressive enhancement of process efficiency across successive iterations of the system, primarily attributable to learning, optimisation, and training within the robotic or collaborative framework. The chart's axes represent iteration numbers, ranging from 1 to 10, alongside efficiency values between 0.85 and 0.98. The blue line, marked with circular points, illustrates an upward trend in efficiency over time, identified in the legend as "Efficiency Improvement." During the initial five iterations, the system demonstrates a rapid increase in efficiency, indicating significant adaptation and learning during this pivotal development phase. From iteration six onwards, the system exhibits a typical optimisation curve characterised by more gradual, incremental improvements. This pattern suggests that the foundational algorithm progressively refines system performance with repeated executions until it attains an optimal operational level.

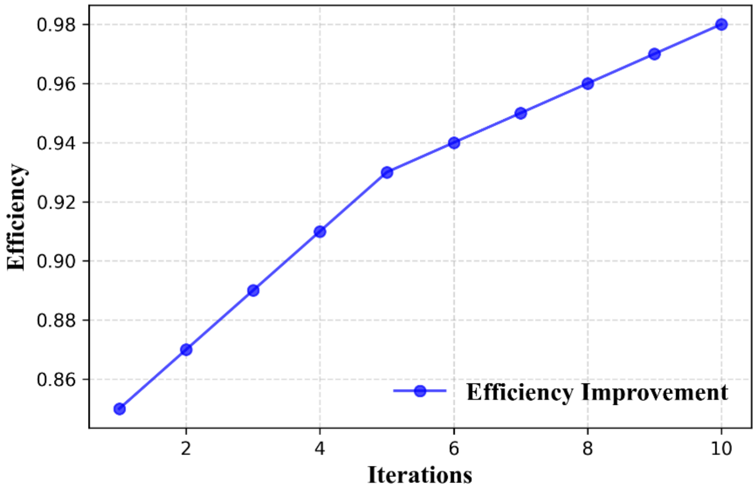

Fig.4: Efficiency Validation

Figure 5, accompanied by a trendline, reveals the relationship between task complexity and efficiency. As task complexity increases from 1 to 10, efficiency declines, demonstrated by the consistently downward-sloping red dashed trendline. Efficiency measurements at various complexity levels are represented by green data points, which align closely with the linear trend. This data highlights a strong negative correlation between the complexity of tasks and efficiency outcomes, indicating that greater task complexity corresponds to lower efficiency. This depiction underscores the challenge of maintaining optimal performance amid increasing task demands and emphasises the necessity for strategic enhancements and robotic support to effectively manage such complexity.

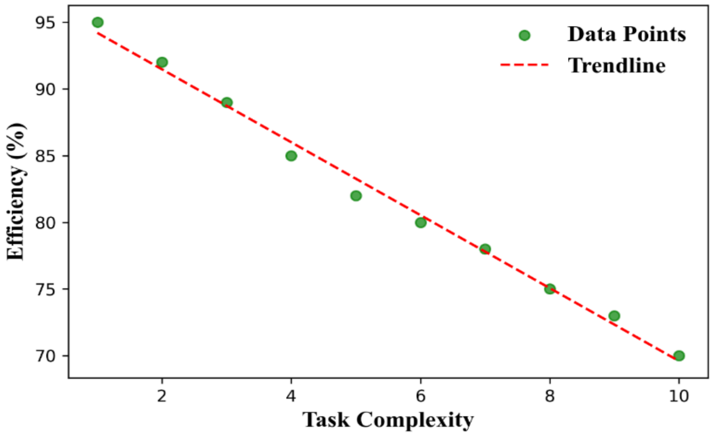

Fig.5: Task Complexity

Figure 6 presents robot performance percentages across task complexity levels from 1 to 10. Each box plot illustrates the distribution of performance outcomes at each complexity level, with the red line representing the median, the blue box indicating the interquartile range (IQR), and whiskers extending to encompass the full data range, excluding outliers marked as individual circles. At lower complexity levels (1–3), median robot performance ranges between 70% and 73%, accompanied by a relatively widespread and several notable outliers, particularly at levels 1 and 3, where performance dips to as low as 53% and 54%, respectively. This variability suggests fluctuations in robot behaviour on simpler tasks, potentially due to over-adaptation or limited engagement. Between levels 4 and 6,

Decision Making: Applications in Management and Engineering
Volume 7, Issue 1 (2024) 735-751

the median remains relatively stable (73–75%), while the performance range narrows, indicating increased consistency. Although outliers persist, especially at level 6, the narrowing IQR points to improved reliability. At higher complexity levels (7–10), robot performance demonstrates a clear upward trajectory, with median values rising steadily from 77% to 81%. The upper whiskers exceed 90% at level 10, and the overall box positions rise accordingly. This trend implies that robots adapt more effectively to complex tasks, potentially due to greater task structure or optimised learning strategies. In summary, robot performance exhibits greater variability and inconsistency at lower complexity levels but becomes more reliable and reaches peak efficiency as task complexity increases. Such adaptive learning and optimisation capabilities suggest that future robotic systems may surpass human performance and excel within demanding industrial environments.

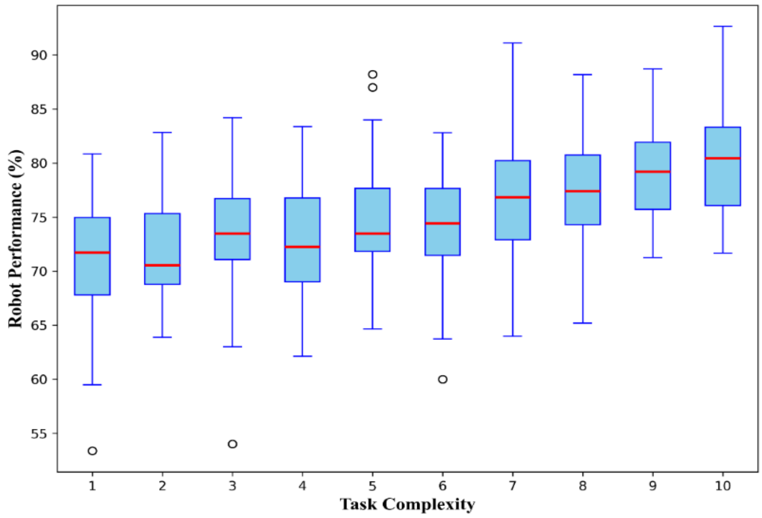

Fig.6: Robot Performance

Figure 7 compares the efficiency of human workers, robots, and human-robot collaborative systems across varying levels of task complexity, from level 1 to level 10. Efficiency, expressed as a percentage, is plotted on the y-axis, while task complexity increases along the x-axis. At the lowest complexity level (1), all three systems—human, robot, and collaborative—operate at near-maximum efficiency, close to 100%, with collaborative performance marginally lower than that of humans and robots. As task complexity rises, efficiency declines steadily for both humans and robots. By complexity level 5, human efficiency is approximately 85%, robot efficiency nearly 82%, whereas collaborative systems maintain a superior efficiency of around 89%. This advantage becomes more pronounced with increasing complexity. At level 8, human efficiency decreases to roughly 70%, robot efficiency falls just below 68%, while the collaborative system continues to outperform both, with an efficiency near 75%. At the highest complexity level (10), human and robot efficiencies drop to approximately 60% and 62%, respectively, whereas collaborative performance remains relatively robust at about 65%. These findings clearly demonstrate that while individual human and robotic performance diminishes as tasks become more complex, collaborative human-robot systems consistently exhibit higher efficiency. This evidence underscores the value and adaptability of collaborative strategies in managing complexity within industrial settings, serving as a crucial factor in sustaining operational productivity in human-robot teamwork.

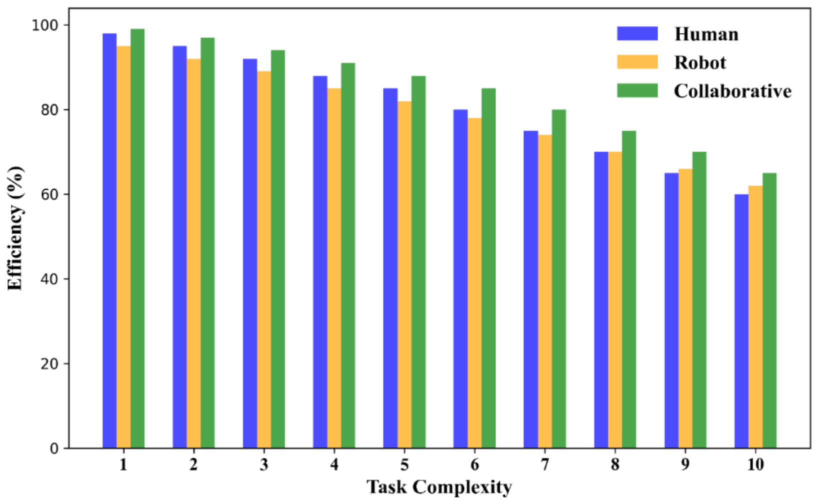

Fig.7: Task Complexity

Figure 8 presents a comparison of task completion rates between humans and robots over a 60-minute period within an industrial assembly setting. Both humans and robots begin with a 0% completion rate at time zero. After 10 minutes, humans reach approximately 10% of task completion, whereas robots achieve only 5%. At the 20-minute mark, human performance increases to 30%, outperforming the robots, which stand at 20%. This trend continues at 30 minutes, with humans having completed 50% of the task, while robots have reached 40%. By 40 minutes, human completion rises to 70%, whereas robots lag at around 60%. At 50 minutes, humans approach 85% completion, closely followed by robots at 80%. Ultimately, both humans and robots attain 100% task completion by the end of the 60-minute interval. The data indicates that although humans initially progress faster, the completion rates of robots converge over time. This diminishing gap suggests that robots improve in efficiency and adaptability during prolonged tasks, demonstrating that robot performance can approximate human levels in extended operations. These findings advocate for leveraging the combination of robotic consistency and human intuition to optimise performance in collaborative industrial environments.

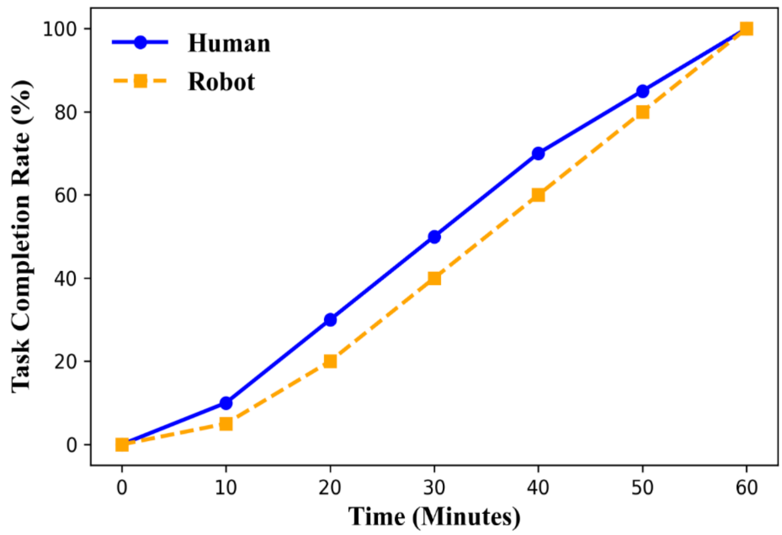

Fig.8: Task Completion Rate

Figure 9 depicts the time contributions of humans and robots, measured in minutes, across four industrial assembly tasks. In Task 1, humans dedicated 10 minutes while robots contributed 5 minutes, resulting in a total duration of 15 minutes. For Task 2, the workload increased, with humans working for 15 minutes and robots for 10 minutes, summing to 25 minutes overall. The task with the greatest human involvement required 25 minutes, alongside 20 minutes of robotic participation, amounting to 45 minutes in total. Similarly, Task 4 lasted 45 minutes; however, the distribution differed, with robots accounting for 25 minutes (more than half of the total time) and humans contributing 20 minutes. This progressive variation demonstrates how task complexity and the nature of collaboration lead to a division of labour, whereby robots assume a greater share of responsibility for executing more complex tasks within advanced human-robot cooperative frameworks.

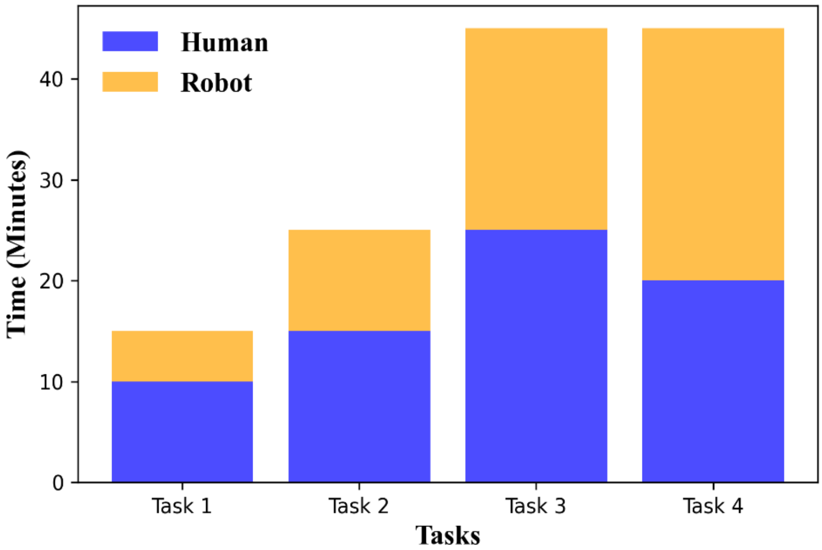

Fig.9: Processing Time

6. Discussion

The experimental findings offer a thorough assessment of the proposed framework’s strengths and limitations in comparison to conventional systems. As illustrated in Figure 4, the system achieved a 15% increase in efficiency from iteration 1 to 10, rising from 0.85 to 0.98, which clearly demonstrates effective adaptation and optimisation over time. Moreover, Figure 7 reveals that collaborative systems consistently maintained an efficiency advantage of 5 to 8% above both standalone human and robotic systems across all levels of task complexity. At the highest complexity level (level 10), collaborative efficiency remained at 65%, compared to 60% for humans and 62% for robots, underscoring the tangible benefits of synergistic collaboration. Insights from Figure 6 show notable variance in robot performance; at lower complexity levels (1–3), median performance ranged between 70% and 73% but exhibited considerable fluctuation and low outliers (down to 53%), suggesting potential over-adaptation or insufficient task challenge. In contrast, performance at higher complexity levels (7–10) improved steadily, with median values reaching 81% alongside reduced variability, indicating enhanced robot engagement and learning when confronted with more demanding tasks.

Figure 8 further confirms that although robots initially complete tasks more slowly than humans, they attain parity with a 100% completion rate within 60 minutes, emphasising their sustained efficiency and adaptability over longer durations. Correspondingly, workload distribution analysis (Figure 9) indicates a progressive shift, whereby Task 5 robots undertake over 55% of the total task time, reflecting their growing proficiency in complex operations. Collectively, these results substantiate the system’s capacity for dynamic task allocation, conflict resolution, and adaptation to operational fluctuations via real-time, utility-based decision-making.

These outcomes suggest that integrating multimodal perception with game-theoretic decision frameworks in collaborative robots significantly boosts productivity, particularly in complex and variable industrial contexts. This approach facilitates more strategic and flexible human–robot interaction, positioning it as a promising avenue for advancing smart manufacturing systems aimed at increasing efficiency without compromising safety or ergonomic considerations. Nevertheless, challenges persist regarding long-term adaptability, primarily due to factors such as sensor reliability, environmental variability, and the inherent unpredictability of human behaviour. The reliance on high-quality multimodal inputs and assumptions of rational human decision-making may constrain performance under dynamic real-world conditions. Additionally, scalability to larger and more intricate industrial settings necessitates optimisation to preserve real-time responsiveness. Addressing these constraints through enhanced robustness, adaptive learning mechanisms, and more efficient algorithms will be essential for effective deployment of this framework across diverse manufacturing environments.

7. Conclusion

Amid the increasing significance of adaptable and intelligent human–robot collaboration within contemporary industrial assembly settings, this study presents a notable advancement by proposing a novel framework that integrates multimodal perception with game-theoretic optimisation. The necessity for dynamic task allocation is underscored by the observed variability in performance efficiency between humans and robots across diverse tasks. The proposed system leverages data derived from visual, auditory, and tactile sensors to enable cobots to more precisely interpret complex human behaviours and environmental signals. Real-time, utility-based decisions are generated by conceptualising human–robot interaction as a dynamic strategic game, wherein decisions are optimised to allocate tasks, resolve conflicts, and adapt to fluctuating operational contexts. Validation in both simulated environments and real-world industrial applications substantiates the framework’s benefits in terms of efficiency, safety, and operational flexibility. This research thus advances smart manufacturing by advocating a perception-driven strategy for seamless and strategic cobot integration within industrial workflows, characterised by robustness and adaptability.

Notwithstanding these encouraging outcomes, the framework presents certain limitations. The system’s learning and adaptation processes are heavily contingent upon sensor fidelity and environmental factors; suboptimal lighting, excessive noise, or mechanical disturbances can adversely affect perceptual accuracy. Additionally, while the game-theoretic model demonstrates efficacy in short-term interactions, its performance may diminish in extended planning horizons or scenarios involving evolving objectives unless periodically updated. These constraints limit the framework’s applicability in uncontrolled or highly variable real-world contexts. Future investigations will prioritise enhancing long-term adaptability through the integration of reinforcement learning within the decision engine, thereby enabling cobots to refine their capabilities via prolonged interaction histories. Further developments aim to augment perceptual capacities through advanced tactile sensing and enhanced semantic vision. Incorporating user feedback mechanisms and modelling operator intent via physiological measures (such as electromyography or gaze tracking) could further bolster system responsiveness and operator trust. Ultimately, extending the framework to facilitate multi-agent coordination among multiple cobots and human operators represents a critical progression towards achieving scalable collaborative assembly in Industry 5.0 manufacturing environments.

References

[1] Alessio, A., Aliev, K., & Antonelli, D. (2024). Multicriteria task classification in human-robot collaborative assembly through fuzzy inference. *Journal of Intelligent Manufacturing*, *35*(5), 1909-1927. https://doi.org/10.1007/s10845-022-02062-4

[2] Andronas, D., Kampourakis, E., Papadopoulos, G., Bakopoulou, K., Kotsaris, P. S., Michalos, G., & Makris, S. (2023). Towards seamless collaboration of humans and high-payload robots: an automotive case study. *Robotics and Computer-Integrated Manufacturing*, *83*, 102544. https://doi.org/10.1016/j.rcim.2023.102544

[3] Cai, M., Liang, R., Luo, X., & Liu, C. (2023). Task allocation strategies considering task matching and ergonomics in the human-robot collaborative hybrid assembly cell. *International Journal of Production Research*, *61*(21), 7213-7232. https://doi.org/10.1080/00207543.2022.2147234

[4] Chen, L., Zhao, H., Shi, C., Wu, Y., Yu, X., Ren, W., Zhang, Z., & Shi, X. (2023). Enhancing Multi-Modal Perception and Interaction: An Augmented Reality Visualization System for Complex Decision Making. *Systems*, *12*(1), 7. https://doi.org/10.3390/systems12010007

[5] Conti, C. J., Varde, A. S., & Wang, W. (2020). Robot action planning by commonsense knowledge in human-robot collaborative tasks. 2020 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS), 1728196159. https://doi.org/10.1109/IEMTRONICS51293.2020.9216410

[6] Conti, C. J., Varde, A. S., & Wang, W. (2022). Human-robot collaboration with commonsense reasoning in smart manufacturing contexts. *IEEE Transactions on Automation Science and Engineering*, *19*(3), 1784-1797. https://doi.org/10.1109/TASE.2022.3159595

[7] Dei, C., Meregalli Falerni, M., Cilsal, T., Redaelli, D. F., Lavit Nicora, M., Chiappini, M., Storm, F. A., & Malosio, M. (2025). Design and testing of (A) MICO: a multimodal feedback system to facilitate the interaction between cobot and human operator. *Journal on Multimodal User Interfaces*, *19*(1), 21-36. https://doi.org/10.1007/s12193-024-00444-x

[8] Ding, X., Guo, J., Ren, Z., & Deng, P. (2021). State-of-the-art in perception technologies for collaborative robots. *IEEE Sensors Journal*, *22*(18), 17635-17645. https://doi.org/10.1109/JSEN.2021.3064588

[9] Duan, J., Zhuang, L., Zhang, Q., Zhou, Y., & Qin, J. (2024). Multimodal perception-fusion-control and human–robot collaboration in manufacturing: A review. *The International Journal of Advanced Manufacturing Technology*, *132*(3), 1071-1093. https://doi.org/10.1007/s00170-024-13385-2

[10] Franceschi, P., Cassinelli, D., Pedrocchi, N., Beschi, M., & Rocco, P. (2024). Design of an assistive controller for physical human–robot interaction based on cooperative game theory and human intention estimation. *IEEE Transactions on Automation Science and Engineering*. https://doi.org/10.1109/TASE.2024.3429643

[11] Hameed, A., Ordys, A., Możaryn, J., & Sibilska-Mroziewicz, A. (2023). Control system design and methods for collaborative robots. *Applied Sciences*, *13*(1), 675. https://doi.org/10.3390/app13010675

[12] Han, L., Zhang, J., & Wang, H. (2025). Shared Control in pHRI: Integrating Local Trajectory Replanning and Cooperative Game Theory. *IEEE Transactions on Robotics*. https://doi.org/10.1109/TRO.2025.3532510

[13] Li, S., Wang, R., Zheng, P., & Wang, L. (2021). Towards proactive human–robot collaboration: A foreseeable cognitive manufacturing paradigm. *Journal of manufacturing systems*, *60*, 547-552. https://doi.org/10.1016/j.jmsy.2021.07.017

[14] Li, S., Zheng, P., Liu, S., Wang, Z., Wang, X. V., Zheng, L., & Wang, L. (2023). Proactive human–robot collaboration: Mutual-cognitive, predictable, and self-organising perspectives. *Robotics and Computer-Integrated Manufacturing*, *81*, 102510.

https://doi.org/10.1016/j.rcim.2022.102510

[15] Liu, S., Wang, L., & Vincent Wang, X. (2022). Multimodal data-driven robot control for human–robot collaborative assembly. *Journal of Manufacturing Science and Engineering*, *144*(5), 051012. https://doi.org/10.1115/1.4053806

[16] Malobický, B., Hruboš, M., Kafková, J., Krško, J., Michálik, M., Pirník, R., & Kuchár, P. (2025). Towards Seamless Human–Robot Interaction: Integrating Computer Vision for Tool Handover and Gesture-Based Control. *Applied Sciences*, *15*(7), 3575. https://doi.org/10.3390/app15073575

[17] Martinetti, A., Chemweno, P. K., Nizamis, K., & Fosch-Villaronga, E. (2021). Redefining safety in light of human-robot interaction: A critical review of current standards and regulations. *Frontiers in chemical engineering*, *3*, 666237. https://doi.org/10.3389/fceng.2021.666237

[18] Mukherjee, D., Gupta, K., Chang, L. H., & Najjaran, H. (2022). A survey of robot learning strategies for human-robot collaboration in industrial settings. *Robotics and Computer-Integrated Manufacturing*, *73*, 102231. https://doi.org/10.1016/j.rcim.2021.102231

[19] Proia, S., Carli, R., Cavone, G., & Dotoli, M. (2021). Control techniques for safe, ergonomic, and efficient human-robot collaboration in the digital industry: A survey. *IEEE Transactions on Automation Science and Engineering*, *19*(3), 1798-1819. https://doi.org/10.1109/TASE.2021.3131011

[20] Rakkolainen, I., Farooq, A., Kangas, J., Hakulinen, J., Rantala, J., Turunen, M., & Raisamo, R. (2021). Technologies for multimodal interaction in extended reality—a scoping review. *Multimodal Technologies and Interaction*, *5*(12), 81. https://doi.org/10.3390/mti5120081

[21] Rautiainen, S., Pantano, M., Traganos, K., Ahmadi, S., Saenz, J., Mohammed, W. M., & Martinez Lastra, J. L. (2022). Multimodal interface for human–robot collaboration. *Machines*, *10*(10), 957. https://doi.org/10.3390/machines10100957

[22] Riar, M., Weber, M., Ebert, J., & Morschheuser, B. (2025). Can Gamification Foster Trust-Building in Human-Robot Collaboration? An Experiment in Virtual Reality. *Information Systems Frontiers*, 1-26. https://doi.org/10.1007/s10796-024-10573-z

[23] Ricci, A., Ronca, V., Capotorto, R., Giorgi, A., Vozzi, A., Germano, D., Borghini, G., Di Flumeri, G., Babiloni, F., & Aricò, P. (2025). Understanding the Unexplored: A Review on the Gap in Human Factors Characterization for Industry 5.0. *Applied Sciences*, *15*(4), 1822. https://doi.org/10.3390/app15041822

[24] Rodriguez-Guerra, D., Sorrosal, G., Cabanes, I., & Calleja, C. (2021). Human-robot interaction review: Challenges and solutions for modern industrial environments. *Ieee Access*, *9*, 108557-108578. https://doi.org/10.1109/ACCESS.2021.3099287

[25] Sun, Z., Zhu, M., Shan, X., & Lee, C. (2022). Augmented tactile-perception and haptic-feedback rings as human-machine interfaces aiming for immersive interactions. *Nature communications*, *13*(1), 5224. https://doi.org/10.1038/s41467-022-32745-8

[26] Tchane Djogdom, G. V., Otis, M. J.-D., & Meziane, R. (2025). A Theoretical Foundation for Erroneous Behavior in Human–Robot Interaction. *Journal of Intelligent & Robotic Systems*, *111*(1), 1-24. https://doi.org/10.1007/s10846-025-02221-8

[27] Wang, S., & Jiao, R. J. (2025). Cognitive intelligent task allocation for human-automation symbiosis in Industry 5.0 manufacturing systems via non-cooperative game theory: a bi-level optimization approach. *The International Journal of Advanced Manufacturing Technology*, *136*(3), 1717-1739. https://doi.org/10.1007/s00170-024-14890-0

[28] Wang, T., Zheng, P., Li, S., & Wang, L. (2024). Multimodal human–robot interaction for human-centric smart manufacturing: a survey. *Advanced Intelligent Systems*, *6*(3), 2300359. https://doi.org/10.1002/aisy.202300359

[29] Xu, J., Sun, Q., Han, Q.-L., & Tang, Y. (2025). When embodied AI meets industry 5.0: human-centered smart manufacturing. *IEEE/CAA Journal of Automatica Sinica*, *12*(3), 485-501. https://doi.org/10.1109/JAS.2025.125327

[30] Xu, W., Yang, H., Ji, Z., & Ba, M. (2024). Cognitive digital twin-enabled multi-robot collaborative manufacturing: Framework and approaches. *Computers & Industrial Engineering*, *194*, 110418. https://doi.org/10.1016/j.cie.2024.110418

[31] Zeng, F., Fan, C., Shirafuji, S., Wang, Y., Nishio, M., & Ota, J. (2025). Task allocation and scheduling to enhance human–robot collaboration in production line by synergizing efficiency and fatigue. *Journal of manufacturing systems*, *80*, 309-323. https://doi.org/10.1016/j.jmsy.2025.03.006

[32] Zhang, F., Zhang, Y., & Xu, S. (2022). Collaboration effectiveness-based complex operations allocation strategy towards to human–robot interaction. *Autonomous Intelligent Systems*, *2*(1), 20. https://doi.org/10.1007/s43684-022-00039-x

[33] Zhang, X., Sun, Y., & Zhang, Y. (2021). Evolutionary game and collaboration mechanism of human-computer interaction for future intelligent aircraft cockpit based on system dynamics. *IEEE Transactions on Human-Machine Systems*, *52*(1), 87-98. https://doi.org/10.1109/THMS.2021.3116115

[34] Zhou, Y., Peng, Y., Li, W., & Pham, D. T. (2022). Stackelberg model-based human-robot collaboration in removing screws for product remanufacturing. *Robotics and Computer-Integrated Manufacturing*, *77*, 102370. https://doi.org/10.1016/j.rcim.2022.102370