Journal of Manufacturing Systems 80 (2025) 948–962

Contents lists available at ScienceDirect

Journal of Manufacturing Systems

journal homepage: www.elsevier.com/locate/jmansys

Technical paper

A novel digital twins-driven mutual trust framework for human–robot collaborations

Junfei Li [a], Enshen Zhu [a], Wenjun Lin [b], Simon X. Yang [a], Sheng Yang [a,*]

[a] University of Guelph, School of Engineering, Guelph, N1G 2W1, Ontario, Canada
[b] Algoma University, School of Computer Science & Technology, Brampton, L6V 1A3, Ontario, Canada

ARTICLE INFO

Keywords:
Trust
Digital twins
Human–robot collaborations
Human digital twins

ABSTRACT

Trust plays an important role and significantly influences human–robot collaborations (HRC). However, most previous research on trust only emphasizes the human attitude toward robots. There needs more understanding of human uncertainties that may also cause disruptions of trust in collaborations. This paper presents a novel mutual trust framework to provide a relatable vision for future development in HRC from an integrated perspective via the integration of human and robotic digital twins. More specifically, a comprehensive review of current trust research in HRC is first provided, including trust factors and state-of-the-art trust models. Second, a novel human–robot mutual trust framework based on 5-layer digital twins models is introduced. The mutual trust framework highlights the interactions amongst modules of artificial intelligence, simulation, and operation, which can provide wide services in HRC (e.g., task allocation and motion planning). A case study of solving a path planning problem is exemplified to evaluate the performance of the proposed mutual trust framework. Compared with singular trust models, the proposed framework enables robotic systems with real-time response and adaptation to human behavior. Some limitations and future work of the mutual trust framework are elaborated in the end.

1. Introduction

As robots become increasingly intelligent and capable, a new trend has emerged that views robots as teammates who can assist humans in accomplishing many complex tasks [1,2]. This trend needs to build and maintain the trust between humans and robots, which allows robotic teammates to dynamically adapt and respond to trust changes in real-time [3]. The complexity and uncertainty of human behaviors in collaborative tasks and environments make trust hard to build and maintain in human–robot relationships [4]. Especially for human operators with limited robotics experience, it is important to use human-friendly ways of building and maintaining trust to improve the efficiency of task completion [5–8]. Trust is commonly understood as a complex psychological state comprising beliefs and expectations about the trustworthiness of another based on previous experiences and interactions under conditions of uncertainty and risk [9]. In addition, trust is also recognized to have cognitive and emotional components. Cognitive processes are often seen as crucial in evaluating trustworthiness, specifically regarding how well the robotic system achieves its intended purpose [10]. Trust is viewed as an attitude that an entity (robot or human) will aid in achieving personal objectives in contexts marked by uncertainty and vulnerability [11].

In the context of HRC, trust is a critical determinant of successful teamwork and sustained human engagement. It provides the foundation that enables humans to reliably delegate tasks to robots, accept their decision-making processes, and adapt to shared work environments. However, most trust research in HRC only emphasizes the human attitude towards robots, which is based on the capability of the robots to operate in an uncertain environment. The uncertainty of human behaviors and health conditions may also cause disruptions in HRC, which means that human trust will fluctuate over time. If human behavior and status become more predictable, robots can have more trust in human operators and generate data-informed decisions to assist human operators. In addition, the modeling of human status and behaviors is difficult in HRC. Firstly, compared to robotic systems, human modeling requires a larger dataset, encompassing more complex geometric configurations, the acquisition of biological signals, and significant variations among different individuals. Secondly, human behavior modeling demands higher real-time performance due to the dynamic and unpredictable nature of human actions. This involves processing a continuous stream of data from multiple sensors to accurately track and respond to human movements and states. Lastly, human performance and behavior are also subject to subjective influences

* Corresponding author.
E-mail address: syang19@uoguelph.ca (S. Yang).

https://doi.org/10.1016/j.jmsy.2025.04.009
Received 3 September 2024; Received in revised form 1 April 2025; Accepted 15 April 2025
Available online 3 May 2025

such as psychological stress and fatigue levels, which can significantly impact both physical and cognitive functions.

To better capture human-related data, human digital twins (HDTs), a technology that creates a virtual representation of a physical human and enables simulation, prediction, and optimization of its physical counterpart, has attracted increasing attention [12–14]. HDTs are advanced simulations that replicate human physiological, cognitive, and behavioral states. These digital replicas are created using comprehensive datasets and are continuously updated with real-time data from various sensors to mirror the current state of an individual. HDTs can provide deeper insights into human conditions, allowing for better prediction and management of behaviors and health statuses. Therefore, a well-developed HDT can be used to improve robot trust in humans, but the modeling and simulation of HDT for human–robot collaboration is still very challenging.

This paper aims to provide a novel mutual trust framework for HRC based on integrated digital twins (DTs) of robotic and human operators. We introduce a comprehensive mutual trust framework that simulates trust between humans and robots based on multiple trust metrics. In addition, we present a more unified model to represent human-related data using the concept of HDTs. HDTs collect real-time multimodal data on human health, motion, cognition, and emotions to create a virtual human twin that supports trust-building and maintenance. Robotic digital twins (RDTs) collect real-time data on robot operations and status to form a virtual robotic twin that improves predictability and reliability. The DTs platform uses this real-time data to offer simulation, operation, and AI-inference functions, thereby fostering mutual trust and optimizing collaborative performance. The proposed framework can build and maintain trust at a reasonable level by considering the dynamics of both humans and robots. The contribution of this paper can be summarized as follows:

(1) A comprehensive mutual trust framework is proposed to simulate trust between humans and robots based on multiple trust metrics.
(2) A unified model is proposed to represent human-related data using the concept of HDTs.
(3) A DTs-based platform is proposed to offer simulation, operation, and AI-inference functions, thereby improving mutual trust and optimizing collaborative performance in a path planning case study.

This paper is structured in the following manner. In Section 2, the state-of-the-art analysis are provided for the concept of trust in HRC. Section 3 introduces the recent research of modeling trust in HRC. Section 4 introduces the proposed mutual trust framework and DT-based platform. Section 5 provided a case study for trust-based path planning based on the proposed mutual trust framework. In Section 6, the proposed framework and future research are summarized.

2. Concept of trust in HRC

In previous studies, there are two types of trust in HRC based on the perspective of humans or robots: human trust in robots (HTIR) and robot trust in humans (RTIH). HTIR refers to the degree of confidence that humans have in the capabilities, reliability, and intentions of robotic systems. This concept is critical in HRC as it influences how effectively and safely humans and robots can collaborate. RTIH pertains to the trust level by which robots evaluate human behaviors and decisions. This is essential for ensuring smooth and efficient collaboration in HRC. In HTIR research, researchers have found that many robot-related factors are related to increasing the trust from humans in robots. These factors include shape [15], voice [16], color [17], communication [18], speed [19], and system transparency [20].

This section presents a state-of-the-art analysis of the factors that constitute the recent understanding of trust in HRC. Some examples of trust factors are shown in Fig. 1. Earlier studies have categorized trust determinants into three main categories: robot-related factors (such as shape, color, and speed), human-related factors (including prior experiences, self-confidence, and situational awareness), and environment-related factors (like risk level and task complexity) [21]. However, evidence from previous studies suggested that robot-related factors are the most significant contributors to trust [22]. Accordingly, this section focuses on the robot-related factors, which are organized into four key areas: performance and reliability, personalization and adaptability, safety and predictability, and transparency and explainability.

2.1. Performance and reliability

The performance of the robot, including battery energy [27], task completion [28], robotic speed [19], and error rates [23,24], is crucial for trust building. Humans are more likely to trust robots that demonstrate high efficiency and reliability in HRC. Desai et al. [29] tested the trust between humans and robots in low-reliability scenarios. The results showed that a decrease in reliability influences trust levels, and these reliability fluctuations impact self-evaluations of their performance. Wright et al. [20] tested how the reliability of a robotic member affected human trust. Results suggested that reliability may have a stronger influence on human trust and robot perceptions than other factors, such as transparency. Chavaillaz et al. [30] studied the influence of reducing reliability on the automatic system. Results indicated that a decrease in automation reliability negatively impacted operator trust, and the operator still relied on the automation system unaffected by variations in reliability.

In the context of trust in HRC, reliability is defined as the extent to which a robot consistently meets expected performance standards, thus reinforcing human confidence in its capabilities. This process depends on the human capacity to perceive, interpret, and evaluate the robot's actions through cognitive functions, ultimately shaping the trust relationship between humans and robots. Xu et al. [31] investigated that the first impression of the robot can significantly influence human trust if the robot provides a correct answer from the start. Ahmad et al. [32] demonstrated an inverse correlation between trust and cognitive load, indicating that increased cognitive load among participants was associated with a decline in trust ratings. Gompei et al. [33] analyzed the attitude of familiarity as being associated with both cognitive and affective trust, whereas the errors made by the robot influenced cognitive trust.

2.2. Personalization and adaptability

Robots capable of personalizing their responses according to human preferences, behaviors, and feedback have significantly enhanced trust in their interactions. The integration of design uniqueness in terms of appearance and color significantly contributes to the overall acceptance and emotional connection users have with robots [17,34]. By incorporating aesthetically pleasing designs and customizable color schemes, robots can appeal to the personal tastes and preferences of a wider audience. Understanding and optimizing the visual interactions between humans and robots is fundamental to enhancing trust in HRC. Desai et al. [35] studied the influence of choosing semantic non-semantic indicators to guide humans and robots to finish tasks. The quality, tone, and clarity of robotic auditory outputs, such as speech and operational sounds, play a substantial role in building and maintaining trust between humans and robots [36]. Gender stereotypes are another factor that can influence trust. Kraus et al. [16] conducted research investigating the impact of gender voice on trust in both explicit and implicit gender attributions to robots. Participants exhibited a greater trust in the male robot, which they also rated as more reliable and competent than the robot with a female personality; conversely, the robot with a female personality was perceived as likable.

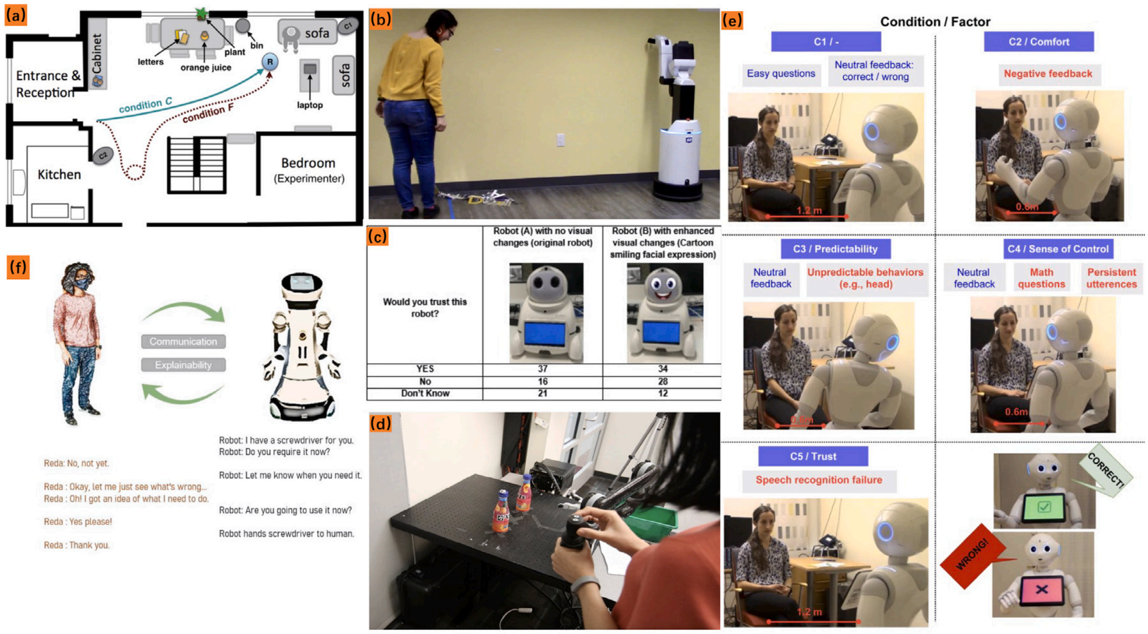


Fig. 1. Examples of trust tests and measurements via (a) correct and faulty navigation paths [23], (b) robots making some errors [24], (c) visual changes [17], (d) mutual adaptation in the collaborative environment [25], (e) safety factors testing [26], and (f) communication [18]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The trust between robots and humans involves managing uncertainty and vulnerability throughout their collaborative process. High adaptability enhances the confidence that collaborators will perform as anticipated in achieving the designated objectives despite diverse uncertainties and vulnerabilities. Therefore, adaptability is important in affecting trust between humans and robots. Fischer et al. [37] provided qualitative analyses for the influence of robot adaptability on human trust in medical interactions. Results showed that high adaptability could improve human trust when considering specific task requirements. Nikolaidis et al. [25] introduced a formalism for mutual adaptation in human–robot interaction to maintain their trust, even in instances that may contravene personal preferences.

2.3. Safety and predictability

Several studies have addressed the significance of safety in HRC, including the application of various technologies, such as estimation and evaluation, software devices designed, impact detection systems, and collision-free strategies [38]. Particularly in the context of industrial robotic collaboration, ensuring safety is of critical importance [39–41]. Akalin et al. [26] studied the principal determinants of perceived safety, including comfort, familiarity, predictability, sense of control, transparency, and trust.

Robotic systems characterized by high predictability are more likely to gain trust as they enhance efficiency and facilitate human operators' comprehension. On the one hand, accurately predicting future positioning can minimize potential harm to humans. On the other hand, foreseeing and acknowledging errors or faults in robotic systems can help maintain human trust, especially when anticipating such faults. Washburn et al. [24] tested how expectations for robot functionality affect perceptions of robot trust. Results showed that participants with low initial expectations exhibited enhancements in trust after interacting with the robot, whereas participants with high initial expectations showed no changes in trust. Reinhardt et al. [42] tested the influence of trust and predictability with different movement strategies in proximal cooperation. Schadenberg et al. [43] tested the effect of unpredictability on the performance evaluation of the robot by human operators, and they found that unpredictability can cause low-performance evaluation.

2.4. Transparency and explainability

Transparency is a significant subject of trust in humans. Transparency typically refers to the clear communication by robotic systems to human users about their operating principles, decision-making processes, and functional limitations. Wright et al. [20] tested how transparency influenced the trust and performance of the robot when it finished different tasks. Matthews et al. [44] analyzed the implication of transparency to the robot trust, and they found that the transparency information should be personalized for the mental model to support accurate trust calibration. Lewis et al. [45] introduced both computational and human-based findings to enhance the transparency of deep reinforcement learning networks through object saliency visualizations of internal states.

Explainability constitutes a crucial aspect of HRC, enabling robots to execute actions at varying levels and communicate with humans in a manner that is trustworthy and conducive to human-friendly interaction. One challenge is the design of new strategies and algorithms for generating explanations based on previous experiences and particular human operators [46]. Ambsdorf et al. designed a competitive board game to demonstrate the necessity and potential for the explanation in HRC and the comprehensive evaluation of its impacts. Shin studied the impact of explainability in artificial intelligence (AI) on user trust and attitudes towards AI [47]. Dehkordi et al. [18] proposed a mental model to help robots understand human behavior and achieve five desirable traits: fluent behavior, adaptability, trust-building, effective communication, and explainability.

3. Modeling of trust in HRC

Modeling trust in HRC involves developing mathematical models to quantify and manage the trust dynamics between humans and robots. This process includes defining trust metrics, understanding the dynamic nature of trust, and maintaining the trust at a reasonable level. Most human–robot trust models are based on the robot's performance, and the majority focus on human-to-robot trust. This focus emphasizes how well a robot performs tasks and how reliably it interacts with humans, thereby influencing human confidence and reliance on the robot. Researchers have developed various metrics to assess human trust, such as task completion rates, accuracy, and the robot's ability to adapt to new situations.

Uni-directional trust is typically based on performance history with personalized trust metrics. Guo et al. [3] developed a personalized trust prediction model, employing Bayesian inference for parameter estimation. Several studies have evaluated trust by examining a robot ability to perform specific tasks and querying individual trust [31, 48]. These methods for trust modeling represent significant progress in understanding and articulating human trust in robots. However, performance-based trust models rely on historical data regarding the robot performance on specific tasks, which limits their applicability for transferring trust across different tasks. Research on human–robot trust often employs probabilistic models. Fooladi et al. [49] developed a foundational model for predicting trust during interactions between humans and multiple robots. Lee et al. [50] demonstrated that robots capable of assessing and adjusting to human intentions and abilities in decision-making processes can foster greater trust from humans. Yu et al. [51] proposed a trust-aware decision-making framework for HRC using a partially observable Markov decision process and syntactically co-safe linear distribution temporal logic to model and optimize trust-sensitive task execution through a probabilistic labeling function and modified point-based value iteration. Xu et al. [52] proposed a machine learning-based trust management model for connected and automated vehicles to enhance trust evaluation in dynamic traffic environments. Current trust frameworks often rely on performance metrics. Thus, performance-oriented trust models effectively encapsulate human trust in robots.

There is only a limited number of trust models that have considered the modeling of RTIH. Wang et al. [53] proposed a computational model of a robot's trust in its human co-worker based on various performance factors and validated through experiments. Dorbala et al. [54] proposed a deep reinforcement learning-based trust-driven robot navigation model that quantifies affective features from human language-based instructions to compute a human trust metric to enable trust-aware decision-making for efficient and reliable language-guided navigation. Li et al. [55] proposed a trust model based on linear Gaussian and sparse Gaussian processes and used the Monte Carlo tree search method to optimize the decision-making process of the robot with the goal of improving human–robot trust in emergency scenarios.

Mutual or bi-directional trust is a challenging issue in HRC. Azevedo-Sa et al. [56] introduced a novel capabilities-based bi-directional trust model for predicting trust from human or robotic trustors, which outperforms existing models and is useful for control authority allocation in human–robot teams. Rahman et al. [57] developed a real-time trust measurement to improve assembly performance by considering mutual trust between human and robot. Li et al. [58] introduced a bidirectional trust-based variable autonomy control approach for semi-autonomous mobile robots to reduce operator workload and improve system usability.

However, the existing research on trust modeling has primarily emphasized one-directional trust, particularly in building human trust in robots to enhance cooperation efficiency. Some studies have explored bi-directional trust [56,59], but these often focus on a single criterion (e.g., human intention) of trust between humans and robots. In addition, previous research has paid little attention to the impact of human factors, particularly the uncertainties in human behavior that can significantly influence trust building. Human behavior is complex and influenced by a multitude of variables, including psychological states, physical conditions, and cognitive contexts. The proposed mutual trust framework has distinctive features compared to other trust-based research, as shown in Table 1. The proposed approach employs a mutual trust framework and incorporates multiple trust factors, enabling it to perform various tasks.

Table 1
The proposed mutual trust framework in comparison with other trust-based research in HRC.

Method	Trust direction	Trust factor	Task type
Guo and Yang [3]	HTIR	Single	Trust estimation
Fooladi Mahani et al. [49]	HTIR	Single	Predicting trust
Lee et al. [50]	HTIR	Multiple	Adjusting interaction
Yu et al. [51]	HTIR	Multiple	Automatic driving
Xu et al. [52]	HTIR	Multiple	Trust evaluation
Dorbala et al. [54]	RTIH	Single	navigation
Wang et al. [53]	RTIH	Multiple	Trust estimation
Li et al. [55]	RTIH	Multiple	Emergency response
Azevedo-Sa et al. [56]	Mutual	Single	Task allocation
Rahman and Wang [57]	Mutual	Single	Industrial assembly
Li et al. [58]	Mutual	Multiple	Improve usability
Proposed	Mutual	Multiple	Multiple tasks

4. Mutual trust framework based on DTs

The overview of the mutual trust framework is shown in Fig. 2. The proposed framework emphasizes the advantages of DT in improving collaborations between robots and human operators. The red box indicates the collected real-time HDT data and framework that can provide a virtual human twin to support the trust-building and maintenance process. The blue box indicates the real-time RDT data and framework, where the virtual robotic twin can provide different responses to improve trust and collaborative performance. The human–robot digital twin's platform can use the real-time DT data from RDT and HDT to provide different functions, such as simulation functions (e.g., visualization), operation functions (e.g., status diagnosis, emotional analysis, and health monitoring) and AI-inference functions (e.g., status prediction, learning, and decision suggestion). Based on real-time DT data and platform integration, many services can be achieved in HRC that can outperform existing technologies. This section introduces the structures of RDT and HDT within the proposed framework. In addition, it offers an introduction to the proposed mutual trust framework and trust model. Finally, the services that the framework can provide are discussed.

4.1. Robotic digital twins

Robotic DT has risen as a significant area of research that presents an effective approach for mirroring specific attributes of a physical robot to create a virtual twin [60]. However, the structure of robotic DT based on trust with humans has yet to be discussed, with only isolated research on modeling a safe and flexible collaborative framework or environment. Human trust in robots is subjective cognition, such as individual preferences for color, shape, and sound, or an objective evaluation of the robotic systems, such as requirements for safety and stability [15]. Therefore, after analyzing various factors that affect the trust in human–robot collaboration based on the previous studies, a five-layer architecture of the RDT is proposed (as shown in Fig. 2) that integrates the main trust factors in the common trust models and checks against the collaborative environment.

4.1.1. Physical layer

The physical layer mainly collects real-time physical state and performance data. The current physical information can be accomplished through several methods depending on the complexity of the task requirement. Let the physical state of the robot be represented by a vector $x(t) \in R^n$ at time t, where n is the number of state variables. The dynamics of the physical robot can be described as

$$\dot{x}(t) = f(x(t), u(t)) \tag{1}$$

where $u(t)$ is the control input vector and f is a nonlinear function representing the system dynamics. Most robots can equip sensors to collect

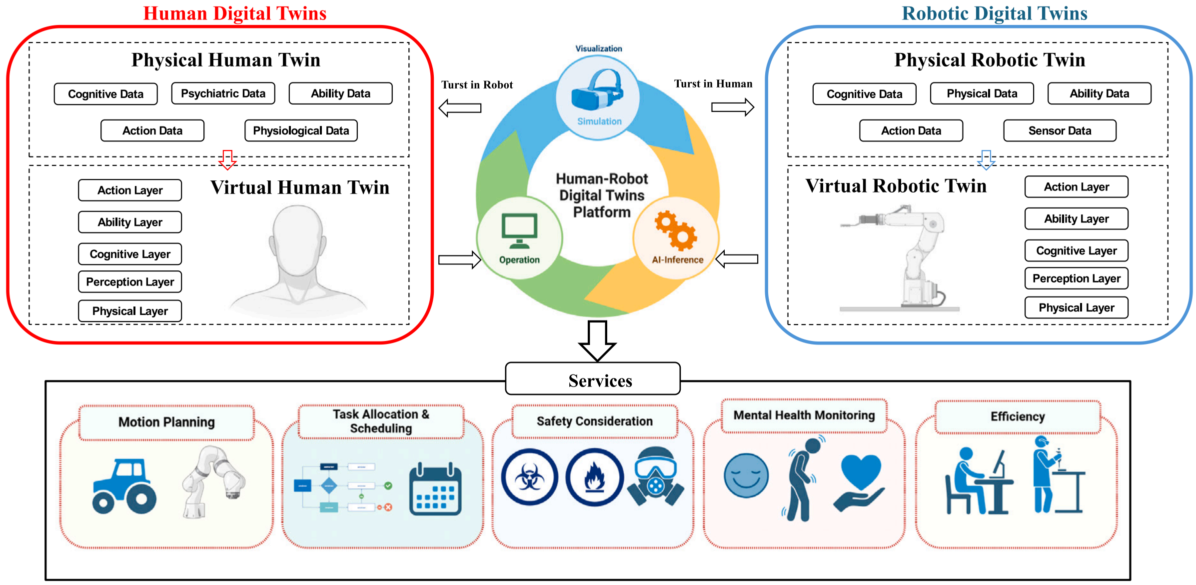


Fig. 2. The overview of the proposed mutual trust framework via integrating DTs of robot and human operators. Red box: the human digital twin. Blue box: the robotic digital. Black box: types of services the proposed framework can provide in HRC. The human–robot DTs platform achieves three primary functions by integrating real-time DT data. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

data about their state. The physical layer continuously collects real-time sensory data (e.g., position [61], force [62], and velocity [63]) to ensure precise task execution. This layer collects raw sensory data and processes and interprets this information to provide real-time feedback on the robot's physical interactions with its environment.

4.1.2. Perception layer

The perception layer represents the sensing ability of other objects, such as the environment, humans, and other robots. Environmental sensors can collect information about the surroundings, including temperature, humidity, light levels, obstacles, or other objects. Let the perception data be denoted by a set of observations $z(t) = z_1(t), z_2(t), \dots, z_k(t)$, where k is the type of sensors. The perception layer can be described as

$$z(t) = o(h(t), e(t)) \tag{2}$$

where $h(t)$ and $e(t)$ are human and environmental information, respectively. Function $o(.)$ is the observation function. Robotic systems using ROS or similar middleware can obtain environmental data by subscribing to relevant sensor topics [64]. Common data collection methods include SLAM-based techniques [65–67], learning-based approaches like YOLO and SSD [68,69], and multimodal fusion of vision, audio, haptics, and physiological sensing [70]. These integrated techniques enhance robotic perception, enabling comprehensive environmental awareness.

4.1.3. Capability layer

The capability layer defines the robot's functional potential based on data from the physical and perception layers. Let the capabilities of the robot be represented by a set of functions $C = c_1, c_2, \dots, c_p$, where each function c_p represents a specific capability, such as mobility, manipulation, recognition, and communication [61,71–73], as well as adaptability to dynamic environments like ocean currents [74]. This layer ensures operational flexibility and robustness, allowing robots to perform specialized tasks in various conditions. For example, it specifies lifting capacity in industrial settings or navigation abilities for aerial and aquatic robots, optimizing efficiency and safety.

4.1.4. Cognitive layer

The cognitive layer represents the robotic understanding of human behavior based on the knowledge update and transfer within humans. Let $k(t)$ be the knowledge vector of the robot at time t, representing its understanding of human behavior and the collaboration task. The cognitive layer can be described as a learning process.

$$\dot{k}(t) = l(k(t), z(t)) \tag{3}$$

where l is a function that updates the robot knowledge based on its current knowledge and perception. Robots acquire knowledge through visual, auditory, textual, and sensor data [75]. Various learning methods, such as transfer, federated, and deep learning, enhance their ability to understand human behavior [76–78]. Recently, cognitive digital twins further advance this capability by integrating diverse data sources to refine predictions and improve decision-making [79]. Based on the cognitive digital twins, robots increasingly mimic human cognition and enable adaptive learning and real-time responses to dynamic environments. By analyzing past interactions, the cognitive layer continuously updates the robot's knowledge and enhance its ability to anticipate human actions and optimize collaboration.

4.1.5. Action layer

The action layer represents the physical behavior consequential to the collaboration task. The planning and control of robot behavior are the most essential factors in this layer, such as collision-free path planning and effective tracking and control of robots. Let the actions of the robot be denoted by $a_R(t) \in A$, where A is the action space. The action selection can be formulated as an optimization problem.

$$a^*(t) = \arg\max_{a_H \in A} E(x(t), h(t), a) \tag{4}$$

where $E(.)$ is a value function that measures the expected reward for taking action given the current state $x(t)$ and human behavior $h(t)$. In this layer, many soft computing techniques can enhance the action layer by enabling adaptive planning and real-time decision-making, such as neural networks, fuzzy systems, genetic algorithms, and swarm intelligence [80–83]. These methods allow robots to dynamically adjust behaviors based on inputs from the perception and cognitive layers, optimizing collaboration and improving task performance.

4.2. Human digital twins

The physical human provides different data to support the DT building of the human, as shown in the red box in Fig. 2. The physical human twin collects extensive real-time DT data, encompassing both biometric data derived from biological information collection and cognitive data that reflect human feeling, reasoning, and decision-making processes. Complementarily, the virtual human twin is delineated by a five-layered structure, each stratum contributing to the comprehensive digital embodiment. This structure coalesces physiological and mental metrics to provide an enriched virtual representation.

4.2.1. Physical layer

The physical layer collects the real-time states of a human, including brainwave activity, cardiac frequency, dermal conductance, and muscular dynamics [70]. Let the physical state of the human be represented by a vector $x_{H(t)} \in R^n$ at time t, where n is the number of state variables. Psychological data consists of electric and physical signals. Electric signals include EEG, EMG and ECG [84–86]. In addition, electrodermal activity information can also be used to analyze the physiological and cognitive states of the operator through wearable devices [87–89]. Physical signals, such as voice, gestures, strain, and pressure, provide insight into emotional states, stress, and physical exertion [70]. These signals enable a real-time assessment of human behavior, fatigue, and participation to enhance safety and collaboration in the HRC.

4.2.2. Perception layer

The perception layer represents sensory information in the human sense that helps human to become aware of objects and relationships [90]. Let the perception data be denoted by a set of sensory information $s(t) = s_1(t), s_2(t), \ldots, s_k(t)$, where k is the type of sense organ. The perception layer can be described as

$$s(t) = o_H(r(t), e(t)) \tag{5}$$

where $r(t)$ is the robotic information. Function $o_{H(.)}$ is the observation function. The perception layer enhances adaptability, particularly for users with sensory impairments. It compensates for diminished senses by augmenting other sensory channels. Human cognition and perception represent how humans receive external stimulation and react to the surrounding environment [91]. The process is influenced by an individual's personality, emotions, mental conditions and preserved knowledge. When simulating human psychological activities by modeling perception and cognition, biomedical and biomechanical sensors are commonly used to collect physiological data. For example, EEG captures brain electrical activity via scalp electrodes, EMG tracks muscle contractions and relaxations, galvanic skin response sensors measure variations in skin conductivity, and eye-tracking cameras monitor gaze direction. In addition, combining various sensor-based data facilitates the modeling of human psychological activities, such as emotions, intention, and mental workload, which contributes to building the high-resolution and high-fidelity digital human entity in the HDT framework.

4.2.3. Ability layer

The ability layer represents the ability of humans to perform tasks. Let the capabilities of the robot be represented by a set of functions $C = c_1, c_2, \ldots, c_H$, where each function c_H represents a specific capability. Human ability consists of physical, intellectual, and mental capacities. Physical ability includes strength and endurance, which influence task allocation and ergonomic design in human–robot collaboration. Robots can assist by handling tasks beyond human strength, improving efficiency. Intellectual ability encompasses problem-solving, communication, and specialized knowledge, optimizing task execution and decision-making. Mental ability involves psychological resilience and emotional stability, essential for managing stress and maintaining performance. Detecting mental status remains challenging, relying on behavioral cues, physiological signals, and multimodal sensory fusion [92,93].

4.2.4. Cognitive layer

The cognitive layer represents understanding, reasoning, and decision-making. The cognitive state largely affects human well-being and team performance as cognitive-based tasks become ubiquitous in collaboration tasks. The cognitive layer can be described by a memory and knowledge update process

$$\dot{c}(t) = l_H(c(t), s(t)) \tag{6}$$

where l_H is a function that updates human knowledge based on its current knowledge $c(t)$ and perception $s(t)$, which is collected from human feedback. The cognitive layer enables robots to learn from past interactions using AI-driven adaptive learning techniques, improving trust-based collaboration. Techniques for assessing cognitive state include subjective and objective measures, such as heartbeat, breathing, blood pressure, and eye movement [94]. Since human–robot interaction primarily relies on vision and language, the cognitive layer must handle tasks like Visual Question Answering (VQA), Vision-and-Language Navigation (VLN), and spatial reasoning [95–98]. In addition, various cognitive factors influence trust in human–robot interaction, including voice, facial expressions, and gender similarity [99,100]. As cognitive information is shaped by human perception and feedback, individual differences in interpretation affect decision-making and collaboration [7].

4.2.5. Action layer

The action layer represents the physical interaction with others, environments, and task executions based on perceptual and cognitive abilities. Let the actions of the robot be denoted by $a_H(t) \in A$, where A is the action space. Recent techniques for collecting action data include YOLO, SSD, SVM, and Bayesian networks for human position detection, CNN-based models for activity recognition, and PoseNet or 3D SSD for human pose estimation [68,69,101–106]. The action layer leverages these technologies to interpret human posture, gestures, and movements, enabling real-time adjustments for safer and more effective collaboration. By analyzing subtle body language cues and motion changes, robots can dynamically adapt their responses, improving interaction fluidity.

4.3. Trust building and maintenance via DT platform

The human–robot DT platform is designed to improve collaborations between humans and robots through many intelligent functions. It incorporates advanced technologies divided into three main components: simulation, operation, and AI-based inference. This section introduces the functionalities of these components and how they contribute to building trust between humans and robots.

4.3.1. Simulation

The simulation component visualizes human and robotic operations in real time, aiding understanding and decision-making. Advanced visualization technologies enhance: (1) Clarity: Providing a clear view of the robot's environment and actions for better predictability; (2) Prediction & Planning: Helping operators anticipate outcomes, reduce risks, and optimize performance; (3) Trust: Increasing transparency to build confidence in robotic behavior.

4.3.2. Operation

The operation component integrates several critical functions, including status diagnosis, emotional analysis, and health monitoring, to ensure the smooth execution of HRC tasks. (1) Status diagnosis: monitors and diagnoses the robot's operational status in real-time, quickly addressing any issues to maintain reliability and safety. (2) Emotional analysis: adapts the robot's responses based on human emotional states, enhancing interaction comfort and personalization. (3) Health monitoring: this approach focuses on the physical well-being of humans, adjusting robot tasks to prevent fatigue and ensure safety, further building trust.

4.3.3. AI inference

The AI-inference component enhances robotic operations through: (1) Status Prediction: Identifying potential issues in human and robot behavior to prevent failures and improve performance, fostering trust; (2) Learning: Continuously adapting to human behaviors and preferences for more personalized, context-aware interactions; (3) Decision Support: Using real-time DT data to assist human decision-making, reinforcing trust in robotic collaboration.

This paper proposes a novel trust model for HRC, which integrates the impact of various trust factors identified in previous research. Based on a comprehensive literature review on trust factors in Section 2, the trust model can be defined as

$$T = \sum_{i=1}^{n} \alpha_i F_i^{+} - \sum_{j=1}^{m} \beta_j F_j^{-} \tag{7}$$

where T represents the trust level. Variables F_i^{+} and F_j^{-} are related to positive and negative trust factors based on RDT and HDT collection, respectively. Parameters α_i and β_j are the weights assigned to each positive and negative factor, respectively. Variables n and m are the number of positive and negative trust factors, respectively. The weights reflect the relative influence of each factor on the overall trust, as derived from empirical data and expert assessments. It is important to note that the proposed model applies to the mutual perspective for both human and robotic operations. For the TRIH, the necessary trust factors can be directly obtained through the HDT. The HDT effectively captures and analyzes human behaviors, intentions, and reliability, translating these into trust factors that inform the robot's trust level. Similarly, the THIR can be assessed by gathering the corresponding trust factors from the RDT. The RDT comprehensively represents the robot's operational status, capabilities, and historical performance data. These factors are crucial for evaluating the robot's trustworthiness from a human perspective. The weights of trust factors may vary considerably across individuals. The common approach derives factor weights through questionnaire-based methods, as subjective perceptions of trust often rely on qualitative feedback from user populations. In the this case study, these weights are primarily informed by the findings presented in Section 2. In future experiments involving real human participants, the further data-driven methods can be used to offer additional rigor and tailor the weights more precisely to specific application contexts.

4.4. Services of mutual trust framework

This section discusses essential services that enhance human–robot collaboration using DT technology. It highlights robot motion planning, task allocation and scheduling, safety considerations, mental workload monitoring, and work efficiency. These services promote a secure, efficient, and trustful environment for HRC.

4.4.1. Robot motion planning

Motion planning is essential in robotics, enabling precise, efficient, and safe navigation in complex environments. The mutual trust framework enables robots to plan and execute motions in a way that might be efficient, intuitively understandable, and predictable to human collaborators. From a human-to-robot perspective, it builds a deeper level of trust from the human operators, who can anticipate and adapt to the actions of the robot, leading to a more harmonious and effective collaboration. From the robot-to-human perspective, the robot can understand human behaviors and intelligently generate a trust-based trajectory in collaborative environments.

4.4.2. Task allocation and scheduling

The application of DT is to coordinate the different skills and capabilities of humans and robots, building a trust-based collaborative environment. The suitability of a task for robotic execution is influenced by many factors, such as the physical and geometric attributes of components, safety considerations, and the methods of feeding and joining. DT plays a critical role in assessing these factors by simulating various task scenarios, therefore aiding in the optimal allocation and scheduling of tasks. The proposed framework is focused on ensuring efficiency and building and maintaining trust between robots and human operators. Through accurate simulations and predictive analytics, DT provides a reliable framework for decision-making, enhancing mutual understanding and cooperation in human–robot interactions.

4.4.3. Safety considerations

DT plays a pivotal role in the realm of safety considerations, simultaneously enhancing safety and reinforcing trust in robotic systems. The initial step involves crafting programming directives focused on safety compliance. After task allocation and identification, DT simulates robotic tasks in a virtual environment to develop safe robot motion patterns. Continuous comparison of simulated and actual signals is crucial for detecting failure modes, with any observed discrepancies prompting immediate data feedback for simulation optimization. This process allows for the issuance of temporary, safety-oriented instructions that temporarily override standard programming. Trust is bolstered as DT ensures the safety of operations and the reliability of the system, which reverts to its original tasks once normalcy is restored.

4.4.4. Mental workload monitoring

Integrating DT in workload monitoring significantly advances, especially in building and maintaining trust. DT offers a real-time, accurate digital representation of physical systems, enhancing transparency and reliability in operations. DT technology enables continuous monitoring and predictive analysis, fostering trust through consistent performance and anticipatory problem-solving. By providing clear insights and simulating various scenarios, DT aids decision-making and strategic planning, which are crucial for establishing trust in complex, dynamic environments. Consequently, DT optimizes workload management and plays a vital role in strengthening the trust between human operators and automated systems.

4.4.5. Work efficiency

DT significantly enhances work efficiency and trust in automated systems through accurate, real-time replication of physical operations. By providing transparent insights into operational processes, DT establishes a foundation of trust, assuring stakeholders of system accuracy and dependability. The capacity of DT to simulate various scenarios aids in proactive decision-making and trust reinforcement, ensuring operational resilience.

5. A case study of motion planning using the mutual trust framework

As shown in Fig. 3, to demonstrate the effectiveness of the mutual-trust framework, a case study of path planning in a complex environment is exemplified. This case study illustrates how mutual trust between a human operator and a robotic partner influences the collaborative path-planning process, highlighting the benefits of enhanced transparency and predictability in human–robot collaborations.

The case study is set in the human–robot team navigate to a target location. This setting allows for examining how RDT and HDT contribute to enhancing mutual trust and how this trust influences collaborative path planning. The environment consists of various obstacles and dynamic elements that require the human–robot team to adapt their path planning based on real-time information and mutual trust. The virtual space of DTs is crucial for digitalization, visualization, and visualization of the HRC system, yet rarely explored compared to the data acquisition research on either humans or robots. This paper will particularly focus on the trust modeling and simulation aspects while adopting existing data acquisition approaches in prior studies (e.g., robot status [22] and human fatigue [107]). The expected outcomes of this case study can be summarized as follows

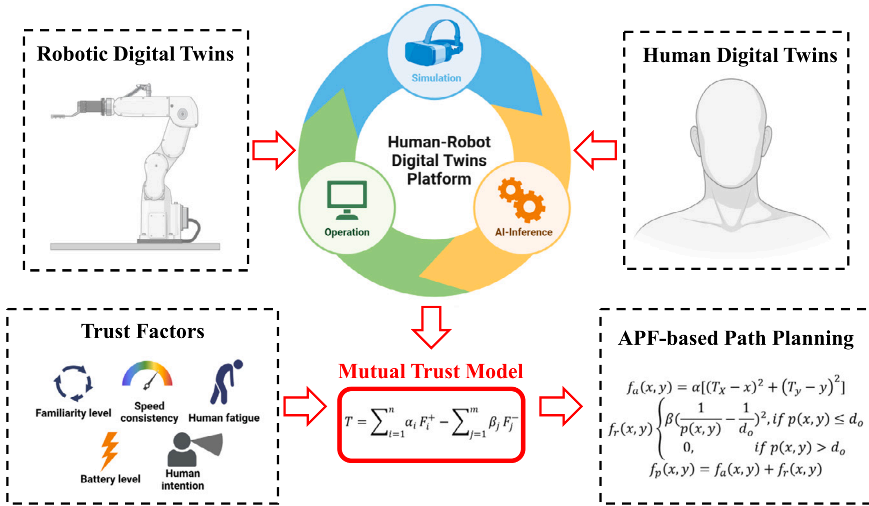


Fig. 3. The trust-aware path planning based on the proposed mutual trust framework.

(1) This is a demonstration of how the RDT and HDT collect data on trust factors and support collaborative path planning.
(2) Compared to the unidirectional trust model, there is improved efficiency, safety, and task completion time in HRC.
(3) Insights into the dynamics of mutual trust and its influence on the collaborative decision-making process.

The proposed mutual trust framework can be tested for physical HRC through experimental validation in real-world scenarios where human operators and robots work together on dynamic tasks. The experimental scenario will consider a controlled collaborative manufacturing assembly line, where robots equipped with DT information for both human operators and dynamically adjust their actions based on physiological, cognitive, and behavioral data from humans in real time. The testing should involve various trust-sensitive tasks, including adaptive motion planning, real-time task allocation, and safety-critical interactions, to evaluate how the framework enhances trust between humans and robots. In addition, human operator trust levels towards robots can be assessed using survey questionnaires, physiological indicators (e.g., breathing rate, heart rate variability, galvanic skin response), and behavioral analysis (e.g., reaction time, error rate). Meanwhile, robots' trust levels towards humans can be evaluated by monitoring predictive accuracy in human intention recognition, reduction of errors in task execution, and general adaptability of the system.

5.1. Trust factors via integration of RDT and HDT

Based on the comprehensive reviews of existing literature and empirical evidence in Section 2, the case study chooses five pivotal and non-exclusive trust factors, which are essential for collaborative decision-making and path planning. These factors are discussed in detail in the following sections.

5.1.1. Battery level

Low battery levels may decrease human trust in the robot's ability to complete the task. In this case study, an exponential decay model is used to mimic the reduction in battery power.

$$B(t) = B_0 e^{-kt} \tag{8}$$

where $B(t)$ is the battery energy at time t; B_0 is the initial energy and k is the discount constant. A low battery level can create perceptions of unreliability and reduce confidence in the robot's performance. This is especially critical in tasks requiring sustained power, such as medical assistance or emergency response. Trust in robots is based on reliability, and visible power depletion may lead users to question their capability. DTs technology have emerged as a key tool for battery modeling and estimation, leveraging IoT and cloud computing. DTs can monitor battery conditions in real time, visualize internal states, and apply advanced diagnostics for dynamic battery management [108,109].

5.1.2. Speed consistency

Inconsistent or erratic speed changes may reduce human trust in the feeling of the robot performance [15]. Both the robot and the human maintain a constant speed, engaging in motion that brings them closer to each other and then further apart. When the distance between the robot and a human is less comfortable, the decrease in trust will be based on the difference in speed. The greater the speed difference, the faster the trust decreases. Human trust in robotic systems is largely dependent on the predictability of robot actions. When robots exhibit erratic behaviors, such as abrupt speed changes, it can trigger a sense of unpredictability, which is often associated with a lack of control or safety. This perception is particularly pronounced in environments where humans and robots work closely together, such as in collaborative manufacturing settings or service industries where spatial navigation around humans is frequent.

5.1.3. Familiarity level

The complexity of the human–robot interaction can obscure the decision-making process, making it harder for humans to anticipate robot actions [110,111]. As complexity increases, so does the mental effort required of the human operator. Overly demanding tasks can lead to errors and frustrations, further eroding trust in the robotic system. When complexity generates stress or anxiety, humans can become more risk-averse and less likely to trust automated systems. In this case study, familiarity is considered as one critical dimension that interacts with complexity. The familiarity of the operator with robotic systems and their associated tasks reduces perceived complexity. As familiarity grows, operators can develop mental models of robot behavior, improving predictability and reducing the cognitive load that often accompany complex interactions. In turn, this supports higher levels of trust. The familiarity level increases linearly as the time spent on collaboration increases. This increase in trust is mainly attributed to the accumulation of positive interactions and the demonstrated ability to navigate and operate effectively within the environment. Over time, as the robot consistently shows competence in familiar settings, the predictability of its actions increases human confidence in the robot's capabilities.

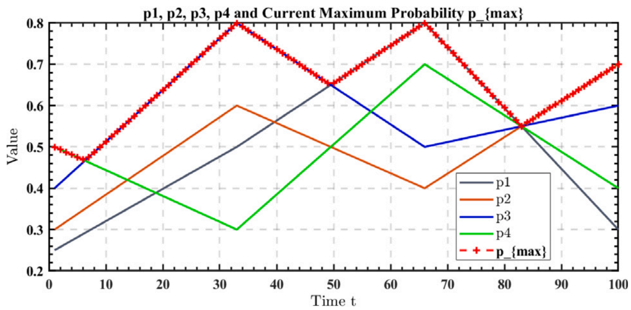


Fig. 4. Simulation results of prediction capability on human intention.

5.1.4. Human fatigue

Higher fatigue may decrease the robot's trust in humans. This section uses a nonlinear function to simulate the increase in fatigue over time. The growth of fatigue with time can be described by a quadratic function

$$F(t) = 1 + f_a t^2 \tag{9}$$

where f_a is the fatigue growth rate. As fatigue in humans grows, it can significantly affect their performance, consistency, and reliability in collaborative tasks, which may influence the robot trust in human partners. In scenarios where robots rely on human input or collaboration, the increasing fatigue can lead to more frequent errors, slower response times, and generally less predictable behavior from the human operators. Robots might interpret these inconsistencies as indicators of decreased reliability if the robot is designed to adapt to and learn from human actions. The DT can analyze the human body joints to detect biomechanical fatigue as a factor of change in back, elbow, and knee joint angles [107]. Fatigue may manifest as reduced physical strength or diminished cognitive performance, which can be evaluated through decreased efficiency and increased error rates in task execution [112].

5.1.5. Human intention

In psychology, intention refers to a person's plan or idea for action. Research categorizes intentions into three types: binary (whether to act), categorical (choosing from set actions), and trajectory-based (deciding movement paths) [113]. Insufficiently estimated human intention can result in a lower degree of trust from the robot to the human. This case study examines how individuals intend to move in a specific direction, considering their head pose [114] and gaze direction [115]. Based on these factors, it creates a probability distribution for the person's next moving direction and calculates the corresponding level of intention prediction capability using the following equation

$$\begin{aligned} L(p_1, p_2, \ldots, p_N) &= \frac{\max(p_1, p_2, \ldots, p_N) - \alpha}{1 - \alpha} \\ \alpha &= \frac{1}{N} \end{aligned} \tag{10}$$

where L is the level of prediction capability on human intention, varying between 0 and 1, and determined by the probability distribution $(p_1, p_2, \ldots, p_N)$ over the N possible movement directions and α is the smallest possible value of $\max(p_1, p_2, \ldots, p_N)$. If the human can move forward, backward, left, or right ($N = 4$), then at least one of these directions must have a probability of 25% or higher. As shown in Fig. 4, if each of the four possible movement directions carries a probability of 25%, the intention level L becomes 0, indicating that the robot cannot determine the human's intended moving direction.

5.1.6. RDT and HDT with trust factors

After identifying the trust factors, the study initially describes the RDT structure, which is utilized to collect and reflect these factors. The overall architecture of the RDT can be represented as

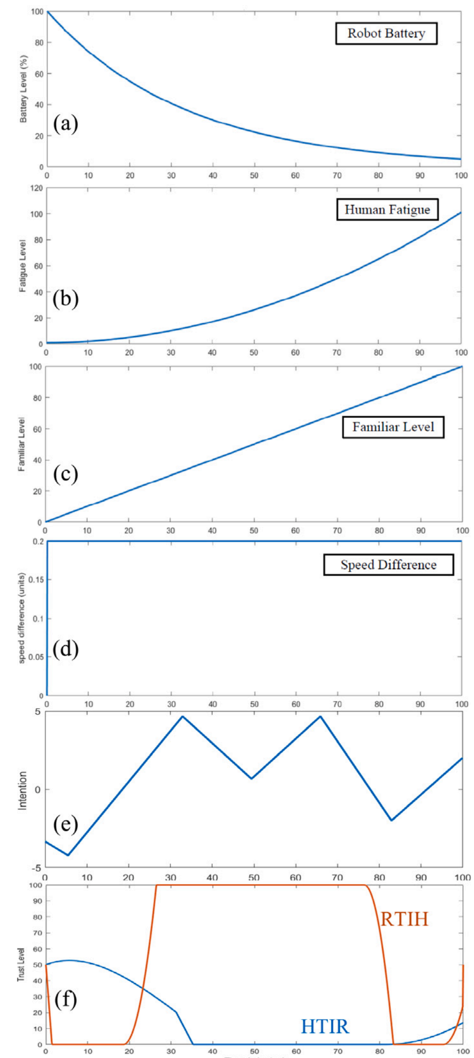
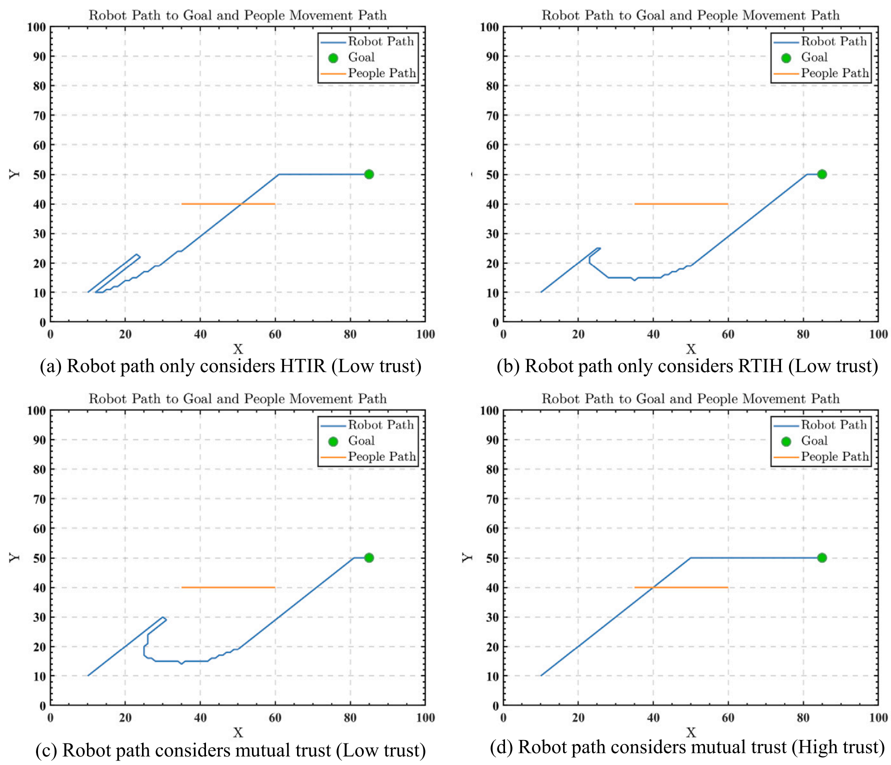


Fig. 5. Simulation results of trust level without strategies. (a) Battery level; (b) Fatigue level; (c) Familiar level; (d) Speed difference; (e) Human intension; (f) Trust level.

$$\begin{aligned} \dot{x}(t) &= f(x(t), u(t)) \\ z(t) &= o(h(t), e(t)) \\ \dot{k}(t) &= l(k(t), z(t)) \\ a^*(t) &= \arg\max_{a_R \in A} E(x(t), h(t), a) \\ u(t) &= c\left(a^*(t)\right) \end{aligned} \tag{11}$$

In this case study, the physical layer $\dot{x}(t)$ includes the robot's battery, speed, direction, and position. The perception layer $z(t)$ enables the robot to detect the environment and recognize human fatigue and intentions. The capability layer C consists of a camera and laser scanner. For cognition, a reinforcement learning-based method updates the robot's understanding of human behavior, with alternatives like graph learning [116], cloud deep learning [117], and intelligent agent models [118]. The action layer focuses on reaching the target while maintaining human trust.



(a) Robot path only considers HTIR (Low trust)
(b) Robot path only considers RTIH (Low trust)
(c) Robot path considers mutual trust (Low trust)
(d) Robot path considers mutual trust (High trust)

Fig. 6. Simulation results of path planning based on the APF method and proposed mutual trust framework. (a) Robot path only considers HTIR; (b) Robot path only considers RTIH; (c) Mutual trust-based path to the target without adaptive strategies; (d) Mutual trust-based path to the target with adaptive strategies.

Similar to RDT, the overall architecture of the HDT can be represented as

$$\begin{aligned} & x_H(t) \\ s(t) &= o_H(r(t), e(t)) \\ C &= \{c_1, c_2, \dots, c_H\} \\ \dot{c}(t) &= l_H(c(t), s(t)) \\ & a_H(t) \in A \end{aligned} \tag{12}$$

The physical layer monitors heart rate, muscle activity, and electrodermal signals, which can indicate fatigue [119,120]. The perception layer analyzes human reactions to the environment, using sensory responses to infer intentions [121]. The ability layer tracks changes in performance due to stress or prolonged effort. The cognitive layer assesses interactions with robots to gauge user willingness and acceptance [122]. The action layer adapts in real time by detecting behavioral shifts like slower movements or increased reaction times, helping mitigate fatigue effects.

5.2. Trust evaluation based on human–robot DTs platform

HTIR integrates the impact of battery depletion, the speed difference between humans and robots, and the level of familiarity with the robot on the trust towards the robot. The trust evaluation of human perspective can be defined as

$$\begin{aligned} E_b &= 0.1\left(100 - \frac{B(t)}{B_{\text{int}}} \cdot 100\right) \\ E_s &= c_1 + S_{\text{diff}} \\ E_f &= c_2 + \frac{F(t)}{100} \\ T_R &= \alpha_1 E_b + \alpha_2 E_s + \alpha_3 E_f \end{aligned} \tag{13}$$

where E_b, E_s and E_f are trust evaluation of the battery energy, speed consistency and familiarity level, respectively. The robot trust in humans is based on two factors: the fatigue level and the intention angle. The trust evaluation of the robot perspective can be defined as

$$\begin{aligned} E_{\text{int}} &= \begin{cases} -0.1(\theta_{\text{curr}} - \theta_{\text{th}}) & \text{if } \theta_{\text{curr}} > \theta_{\text{th}} \\ 1 & \text{if } \theta_{\text{curr}} \le \theta_{\text{th}} \end{cases} \\ E_{f_a} &= 0.1(F(t) - 1) \\ T_h &= \alpha_4 E_{\text{int}} + \alpha_5 E_{f_a} \end{aligned} \tag{14}$$

where E_{int} and E_{fa} are trust evaluations of the intention and fatigue, respectively. In this simulation, the parameters are set as $c_1 = 0.25, c_2 = 0.01, \theta_{th} = 30$. Fig. 5 illustrates the dynamic evolution of the HTIR. Initially, trust levels in the robot increase as it demonstrates familiarity with its surroundings. However, a decline in trust is observed as the battery levels decrease. After that, a notable decrease can be observed when the robot is close to the human operator below a predefined comfort distance. This decrease is exacerbated by discrepancies in the velocities of the human and robot. As the robot fails to maintain a consistent and comfortable distance, trust precipitously declines to 0. Subsequently, as the distance between the human and the robot increases, thereby alleviating immediate discomfort and perceived risk, trust begins to repair gradually. Fig. 5 shows the dynamics of the trust in the human operator. Initially, as the intention remains within a predefined acceptable angle and fatigue levels are low, the trust in the human correspondingly increases. This early stage of interaction suggests a positive assessment of the human's reliability and capability to manage tasks effectively. However, a subsequent decline in trust is observed when the human's intentions exceed predefined acceptable angles. The robot interprets these deviations as indicators of a lack of focused attention, which leads to a reduction in trust. The trend of decreasing trust continues with further increasing levels of human fatigue.

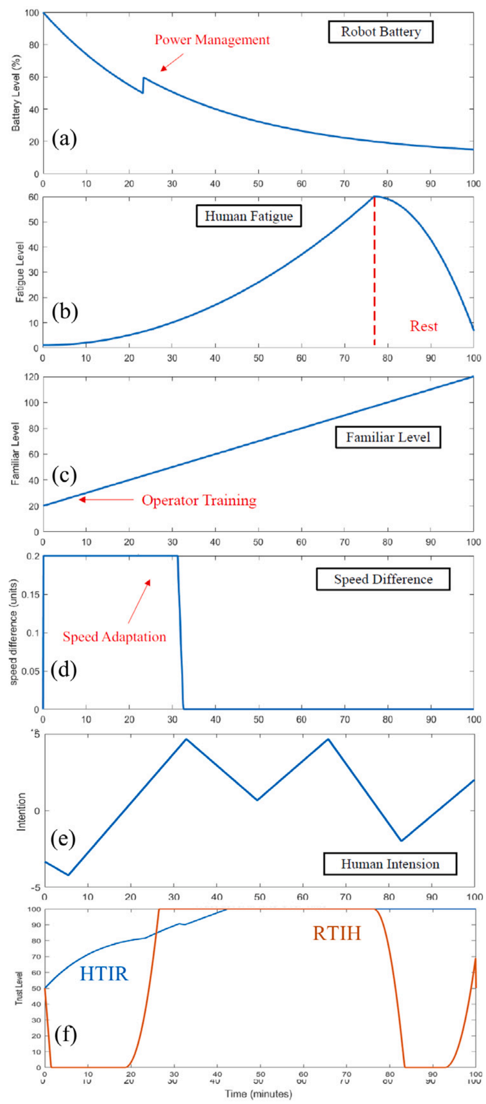


Fig. 7. Simulation results of trust level with strategies. (a) Battery level; (b) Fatigue level; (c) Familiar level; (d) Speed difference; (e) Human intension; (f) Trust level.

5.3. Mutual trust framework for Artificial Potential Field-based motion planning

The Artificial Potential Field (APF) method is commonly used in robot path planning. It guides the robot from the starting point to the target and avoids obstacles. In this method, the target generates an attractive force to guide the robot towards the goal, while obstacles generate repulsive forces to prevent the robot from colliding. The attractive and repulsive forces can be written as [123]

$$f_a(x,y) = \alpha\left[(T_x - x)^2 + (T_y - y)^2\right]$$
$$f_r(x,y) = \begin{cases} \beta\left(\frac{1}{p(x,y)} - \frac{1}{d_o}\right)^2, & \text{if } p(x,y) \le d_o \\ 0, & \text{if } p(x,y) > d_o \end{cases} \tag{15}$$
$$f_p(x,y) = f_a(x,y) + f_r(x,y)$$

where repulsive range d_o is the key factor in determining the path generation.

For trust-based motion planning, the repulsive range d_o is changed based on the mutual trust of humans and robots. If both the RTIH and the HTIR at the current time are below 50, the repulsion range is increased by 5. If both mutual trusts at the current time are above 50, the repulsion range will decrease by 1. Otherwise, the repulsion range is increased by 1. Fig. 6(a) and (b) show the generated path only considering HTIR and RTIH. Fig. 6(c) shows the mutual trust-based path planning without strategies, which illustrates an expansive repulsion zone around the human becasue of the human operator's low trust towards the robot.

5.4. DT-based strategies for motion planning

The mutual trust model can support the simulation of different operation strategies to better improve trust. Three examples are given, including power management strategy, speed adaptation, and human operator training.

5.4.1. Power management

Effective power management within the human–robot digital twin platform ensures that robots have sufficient battery life to complete tasks without interruptions, thereby reducing operational uncertainty and enhancing reliability. The simulation component of the platform can visualize energy consumption patterns and predict future energy requirements based on the task's complexity and duration. This predictive analysis allows for strategic battery management, ensuring that robots are adequately charged before task execution and can autonomously seek charging stations when nearing low power levels during tasks. As shown in Fig. 6, robots can increase the trust by demonstrating consistent operational capabilities without power failures.

5.4.2. Speed adaptation

The proximity-based speed adaptation involves adjusting the movement speed of robots when close to humans to ensure safety and comfort. This adaptation is crucial in environments characterized by varying risk and task complexity. The operational component of the platform uses real-time data to dynamically adjust the robot's speed, ensuring that movements are neither too abrupt nor too slow, which could either startle or frustrate human operators. This careful modulation of speed, dictated by the proximity sensors integrated into the digital twin framework, enhances safety perceptions and trust in robotic systems. As shown in Fig. 7, by integrating these strategies into the human–robot DTs platform, the overall system adapts to human counterparts' physical presence and operational rhythms. Simulation tools within the platform can model scenarios where these strategies are applied, providing a dynamic representation of interactions and allowing for fine-tuning parameters to improve cohesiveness in human–robot teams.

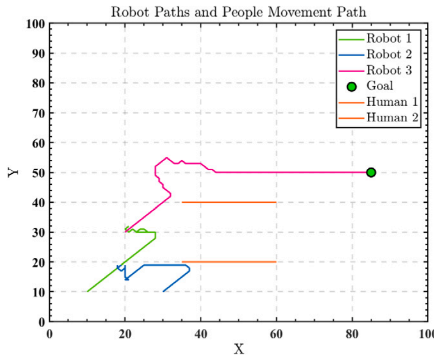
5.4.3. Human operator training

When the DT platform detects that the operator is unfamiliar with the system, it will initiate a training service. This training is designed to enhance the operator's proficiency and ensure they are well-equipped to interact with and manage the system effectively. As shown in Fig. 7, the DT platform leverages real-time data and analytics to identify gaps in the operator's knowledge, providing targeted training sessions that are tailored to address specific areas of improvement. By offering this training, the platform aims to optimize system performance and maintain high operational standards.

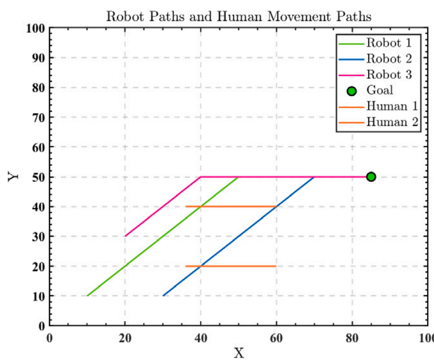
5.4.4. Human operator rest

When the DT platform detects operator fatigue, it will mandate a rest period for the operator. This intervention is designed to ensure the safety and well-being of the operator, as well as to maintain optimal system performance. By monitoring real-time data, the DT platform can accurately identify signs of fatigue and promptly notify the operator to take a break. This proactive approach helps prevent errors, reduce the risk of accidents, and enhance overall efficiency within the system.

As shown in Fig. 7, the strategies based on the DT platform lead to an increase in trust rather than a decrease. Fig. 6(d) shows that due to implementing a series of strategies, trust levels were maintained at a high level, allowing the robot to take the shortest path to reach the target point.



(a) Multi-robot paths considers mutual trust (Low trust)



(b) Multi-robot paths considers mutual trust (High trust)

Fig. 8. Simulation results in multi-agent environments.

5.5. Multi-agent environments

In many real-world industrial settings, it is indeed common for multiple robots to coordinate with multiple human operators simultaneously. The proposed mutual trust framework can still adapt to multi-agent environments, as shown in Fig. 8. Each robot maintains a local trust model that is periodically updated based on both individual and shared observations. By exchanging information regarding task performance, reliability, and user feedback, the robots collectively adjust their trust parameters to account for interactions involving multiple human collaborators.

6. Conclusion

In conclusion, this study proposed a mutual trust framework that integrates DTs of human operators and robots to enhance collaboration. The proposed framework addresses the critical role of trust in HRC by incorporating data-informed strategies to adapt to trust-level changes. Through the use of DTs, the proposed approach not only enhances the understanding of robot capabilities but also provides insights into human conditions, which is crucial for maintaining effective and secure collaborations. In addition, the framework facilitates improved robot design by incorporating DTs of humans and robots, enabling continuous refinement of robot behavior to meet human needs and limitations better. This novel perspective on mutual trust significantly advances the adaptability and responsiveness of robots in collaborative environments. The case study on path planning demonstrates a promising direction for future research in this area, particularly in terms of refining the DT models and expanding the scope of the framework to include more complex and varied HRC scenarios. The DT platform can provide a virtual simulation for different strategies to improve HRC performance. It should be noted that the path planning case is only one of the APF method, and the mutual trust framework can be useful for other applications (e.g., safety and task allocation).

However, this study presents some limitations. Firstly, the proposed framework was designed and validated in single-human HRC environments. Real-world applications often involve large teams, dynamic task assignments, and potentially conflicting human–robot interactions. As the complexity and scale increase, computational demands and the risk of overfitting pre-defined trust models also rise. The future work plan will employ distributed architectures where trust estimation occurs locally and is continuously aggregated to update the global state, ensuring robust performance in multi-robot or multi-human environments. In addition, the proposed framework did not fully address how trust estimation would function under uncertain conditions. Future work proposes adaptive learning-based trust models that leverage online updates to address this gap. As a robot interacts with multiple users in real-time, it can refine its trust estimations using dynamic feedback, reinforcement signals, or Bayesian updating to handle user behavior and intent uncertainties.

CRediT authorship contribution statement

Junfei Li: Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Enshen Zhu:** Methodology, Investigation, Formal analysis. **Wenjun Lin:** Writing – review & editing, Validation, Supervision, Methodology. **Simon X. Yang:** Writing – review & editing, Validation, Supervision, Resources, Methodology, Investigation, Funding acquisition. **Sheng Yang:** Writing – review & editing, Visualization, Validation, Supervision, Project administration, Methodology, Investigation, Funding acquisition.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Sheng Yang reports financial support was provided by Social Sciences and Humanities Research Council (SSHRC). If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research received financial support from the Social Sciences and Humanities Research Council (SSHRC), Canada through the New Frontier Research Fund - Exploaration Grant (NFRFE-2022-00081).

References

[1] Semeraro F, Griffiths A, Cangelosi A. Human–robot collaboration and machine learning: A systematic review of recent research. Robot Comput-Integr Manuf 2023;79:102432. http://dx.doi.org/10.1016/j.rcim.2022.102432.

[2] Khavas ZR, Ahmadzadeh SR, Robinette P. Modeling trust in human-robot interaction: A survey. In: International conference on social robotics. Springer; 2020, p. 529–41. http://dx.doi.org/10.1007/978-3-030-62056-1_44.

[3] Guo Y, Yang XJ. Modeling and predicting trust dynamics in human–robot teaming: A Bayesian inference approach. Int J Soc Robot 2021;13(8):1899–909. http://dx.doi.org/10.48550/arXiv.2007.13151.

[4] Glikson E, Woolley AW. Human trust in artificial intelligence: Review of empirical research. Acad Manag Ann 2020;14(2):627–60. http://dx.doi.org/10.5465/annals.2018.0057.

[5] Baker AL, Phillips EK, Ullman D, Keebler JR. Toward an understanding of trust repair in human-robot interaction: Current research and future directions. ACM Trans Interact Intell Syst (TiiS) 2018;8(4):1–30. http://dx.doi.org/10.1145/3181671.

[6] Kok BC, Soh H. Trust in robots: Challenges and opportunities. Curr Robot Rep 2020;1(4):297–309. http://dx.doi.org/10.1007/s43154-020-00029-y.

[7] Law T, Scheutz M. Trust: Recent concepts and evaluations in human-robot interaction. Trust Human- Robot Interact 2021;27–57. http://dx.doi.org/10.1016/B978-0-12-819472-0.00002-2.

[8] Ye Y, You H, Du J. Improved trust in human-robot collaboration with ChatGPT. IEEE Access 2023. http://dx.doi.org/10.48550/arXiv.2304.12529.
[9] Jones GR, George JM. The experience and evolution of trust: Implications for cooperation and teamwork. Acad Manag Rev 1998;23(3):531–46. http://dx.doi.org/10.2307/259293.
[10] Muir BM. Trust in automation: Part I. Theoretical issues in the study of trust and human intervention in automated systems. Ergonomics 1994;37(11):1905–22. http://dx.doi.org/10.1080/00140139408964957.
[11] Lee JD, See KA. Trust in automation: Designing for appropriate reliance. Hum Factors 2004;46(1):50–80. http://dx.doi.org/10.1518/hfes.46.1.50_30392.
[12] Malik AA, Brem A. Digital twins for collaborative robots: A case study in human-robot interaction. Robot Comput-Integr Manuf 2021;68:102092. http://dx.doi.org/10.1016/j.rcim.2020.102092.
[13] Zafar MH, Langås EF, Sanfilippo F. Exploring the synergies between collaborative robotics, digital twins, augmentation, and industry 5.0 for smart manufacturing: A state-of-the-art review. Robot Comput-Integr Manuf 2024;89:102769. http://dx.doi.org/10.1016/j.rcim.2024.102769.
[14] Li C, Zheng P, Li S, Pang Y, Lee CK. AR-assisted digital twin-enabled robot collaborative manufacturing system with human-in-the-loop. Robot Comput-Integr Manuf 2022;76:102321. http://dx.doi.org/10.1016/j.rcim.2022.102321.
[15] Robinette P, Howard AM, Wagner AR. Effect of robot performance on human–robot trust in time-critical situations. IEEE Trans Human- Mach Syst 2017;47(4):425–36. http://dx.doi.org/10.1109/THMS.2017.2648849.
[16] Kraus M, Kraus J, Baumann M, Minker W. Effects of gender stereotypes on trust and likability in soken human-robot interaction. In: Proceedings of the eleventh international conference on language resources and evaluation. LREC 2018, 2018.
[17] Pinney J, Carroll F, Newbury P. Human-robot interaction: The impact of robotic aesthetics on anticipated human trust. PeerJ Comput Sci 2022;8:e837. http://dx.doi.org/10.7717/peerj-cs.837.
[18] Dehkordi MB, Mansy R, Zaraki A, Singh A, Setchi R. Explainability in human-robot teaming. Procedia Comput Sci 2021;192:3487–96. http://dx.doi.org/10.1016/j.procs.2021.09.122.
[19] Sadrfaridpour B, Saeidi H, Burke J, Madathil K, Wang Y. Modeling and control of trust in human-robot collaborative manufacturing. Robust Intell Trust Auton Syst 2016;115–41. http://dx.doi.org/10.1007/978-1-4899-7668-0_7.
[20] Wright JL, Chen JY, Lakhmani SG. Agent transparency and reliability in human–robot interaction: The influence on user confidence and perceived reliability. IEEE Trans Human- Mach Syst 2019;50(3):254–63. http://dx.doi.org/10.1109/THMS.2019.2925717.
[21] Lewis M, Sycara K, Walker P. The role of trust in human-robot interaction. Found Trust Auton 2018;135–59. http://dx.doi.org/10.1007/978-3-319-64816-3_8.
[22] Hancock PA, Billings DR, Schaefer KE, Chen JY, De Visser EJ, Parasuraman R. A meta-analysis of factors affecting trust in human-robot interaction. Hum Factors 2011;53(5):517–27. http://dx.doi.org/10.1177/0018720811417254.
[23] Salem M, Lakatos G, Amirabdollahian F, Dautenhahn K. Would you trust a (faulty) robot? Effects of error, task type and personality on human-robot cooperation and trust. In: Proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction. 2015, p. 141–8. http://dx.doi.org/10.1145/2696454.2696497.
[24] Washburn A, Adeleye A, An T, Riek LD. Robot errors in proximate HRI: How functionality framing affects perceived reliability and trust. ACM Trans Human-Robot Interact (THRI) 2020;9(3):1–21. http://dx.doi.org/10.1145/3380783.
[25] Nikolaidis S, Zhu YX, Hsu D, Srinivasa S. Human-robot mutual adaptation in shared autonomy. In: Proceedings of the 2017 ACM/IEEE international conference on human-robot interaction. 2017, p. 294–302. http://dx.doi.org/10.48550/arXiv.1701.07851.
[26] Akalin N, Kristoffersson A, Loutfi A. Do you feel safe with your robot? Factors influencing perceived safety in human-robot interaction based on subjective and objective measures. Int J Hum-Comput Stud 2022;158:102744. http://dx.doi.org/10.1016/j.ijhcs.2021.102744.
[27] Yang Q, Parasuraman R. How can robots trust each other for better cooperation? a relative needs entropy based robot-robot trust assessment model. In: 2021 IEEE international conference on systems, man, and cybernetics. SMC, IEEE; 2021, p. 2656–63. http://dx.doi.org/10.48550/arXiv.2105.07443.
[28] Story M, Webb P, Fletcher SR, Tang G, Jaksic C, Carberry J. Do speed and proximity affect human-robot collaboration with an industrial robot arm? Int J Soc Robot 2022;14(4):1087–102. http://dx.doi.org/10.1007/s12369-021-00853-y.
[29] Desai M, Medvedev M, Vázquez M, McSheehy S, Gadea-Omelchenko S, Bruggeman C, Steinfeld A, Yanco H. Effects of changing reliability on trust of robot systems. In: Proceedings of the seventh annual ACM/IEEE international conference on human-robot interaction. 2012, p. 73–80. http://dx.doi.org/10.1145/2157689.2157702.
[30] Chavaillaz A, Wastell D, Sauer J. System reliability, performance and trust in adaptable automation. Appl Ergon 2016;52:333–42. http://dx.doi.org/10.1016/j.apergo.2015.07.012.
[31] Xu J, Howard A. The impact of first impressions on human-robot trust during problem-solving scenarios. In: 2018 27th IEEE international symposium on robot and human interactive communication (RO-MAN). IEEE; 2018, p. 435–41. http://dx.doi.org/10.1109/ROMAN.2018.8525669.
[32] Ahmad MI, Bernotat J, Lohan K, Eyssel F. Trust and cognitive load during human-robot interaction. 2019, http://dx.doi.org/10.48550/arXiv.1909.05160, arXiv preprint arXiv:1909.05160.
[33] Gompei T, Umemuro H. Factors and development of cognitive and affective trust on social robots. In: Social robotics: 10th international conference, ICSR 2018, qingdao, China, November 28-30, 2018, proceedings 10. Springer; 2018, p. 45–54. http://dx.doi.org/10.1007/978-3-030-05204-1_5.
[34] Biermann H, Brauner P, Ziefle M. How context and design shape human-robot trust and attributions. Paladyn, J Behav Robot 2020;12(1):74–86. http://dx.doi.org/10.1515/pjbr-2021-0008.
[35] Desai M, Kaniarasu P, Medvedev M, Steinfeld A, Yanco H. Impact of robot failures and feedback on real-time trust. In: 2013 8th ACM/IEEE international conference on human-robot interaction. HRI, IEEE; 2013, p. 251–8. http://dx.doi.org/10.1109/HRI.2013.6483596.
[36] Torre I, White L. Trust in vocal human–robot interaction: Implications for robot voice design. Voice Attractiveness: Stud Sexy, Likable, Charism Speak 2021;299–316. http://dx.doi.org/10.1007/978-981-15-6627-1_16.
[37] Fischer K, Weigelin HM, Bodenhagen L. Increasing trust in human–robot medical interactions: Effects of transparency and adaptability. Paladyn, J Behav Robot 2018;9(1):95–109. http://dx.doi.org/10.1515/pjbr-2018-0007.
[38] Robla-Gómez S, Becerra VM, Llata JR, Gonzalez-Sarabia E, Torre-Ferrero C, Perez-Oria J. Working together: A review on safe human-robot collaboration in industrial environments. Ieee Access 2017;5:26754–73. http://dx.doi.org/10.1109/ACCESS.2017.2773127.
[39] Proia S, Carli R, Cavone G, Dotoli M. Control techniques for safe, ergonomic, and efficient human-robot collaboration in the digital industry: A survey. IEEE Trans Autom Sci Eng 2021;19(3):1798–819. http://dx.doi.org/10.1109/TASE.2021.3131011.
[40] Villani V, Pini F, Leali F, Secchi C. Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications. Mechatronics 2018;55:248–66. http://dx.doi.org/10.1016/j.mechatronics.2018.02.009.
[41] Magrini E, Ferraguti F, Ronga AJ, Pini F, De Luca A, Leali F. Human-robot coexistence and interaction in open industrial cells. Robot Comput-Integr Manuf 2020;61:101846. http://dx.doi.org/10.1016/j.rcim.2019.101846.
[42] Reinhardt J, Pereira A, Beckert D, Bengler K. Dominance and movement cues of robot motion: A user study on trust and predictability. In: 2017 IEEE international conference on systems, man, and cybernetics. SMC, IEEE; 2017, p. 1493–8. http://dx.doi.org/10.1109/SMC.2017.8122825.
[43] Schadenberg BR, Reidsma D, Heylen DK, Evers V. "I see what you did there" understanding People's social perception of a robot and its predictability. ACM Trans Human- Robot Interact (THRI) 2021;10(3):1–28.
[44] Matthews G, Lin J, Panganiban AR, Long MD. Individual differences in trust in autonomous robots: Implications for transparency. IEEE Trans Human- Mach Syst 2019;50(3):234–44. http://dx.doi.org/10.1109/THMS.2019.2947592.
[45] Lewis M, Li H, Sycara K. Deep learning, transparency, and trust in human robot teamwork. In: Trust in human-robot interaction. Elsevier; 2021, p. 321–52. http://dx.doi.org/10.1016/B978-0-12-819472-0.00014-9.
[46] Setchi R, Dehkordi MB, Khan JS. Explainable robotics in human-robot interactions. Procedia Comput Sci 2020;176:3057–66. http://dx.doi.org/10.1016/j.procs.2020.09.198.
[47] Shin D. The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. Int J Hum-Comput Stud 2021;146:102551. http://dx.doi.org/10.1016/j.ijhcs.2020.102551.
[48] Geiskkovitch DY, Thiessen R, Young JE, Glenwright MR. What? That's not a chair!: How robot informational errors affect children's trust towards robots. In: 2019 14th ACM/IEEE international conference on human-robot interaction. HRI, IEEE; 2019, p. 48–56. http://dx.doi.org/10.1109/HRI.2019.8673024.
[49] Fooladi Mahani M, Jiang L, Wang Y. A Bayesian trust inference model for human-multi-robot teams. Int J Soc Robot 2021;13(8):1951–65. http://dx.doi.org/10.1007/s12369-020-00705-1.
[50] Lee J, Fong J, Kok BC, Soh H. Getting to know one another: Calibrating intent, capabilities and trust for human-robot collaboration. In: 2020 IEEE/RSJ international conference on intelligent robots and systems. IROS, IEEE; 2020, p. 6296–303. http://dx.doi.org/10.1109/IROS45743.2020.9340736.
[51] Yu P, Dong S, Sheng S, Feng L, Kwiatkowska M. Trust-aware motion planning for human-robot collaboration under distribution temporal logic specifications. In: 2024 IEEE international conference on robotics and automation. ICRA, IEEE; 2024, p. 12949–55. http://dx.doi.org/10.1109/ICRA57147.2024.10610874.
[52] Xu Q, Zhang L, Qin X, Zhou Y. A novel machine learning-based trust management against multiple misbehaviors for connected and automated vehicles. IEEE Trans Intell Transp Syst 2024. http://dx.doi.org/10.1109/TITS.2024.3422423.
[53] Wang Q, Liu D, Carmichael MG, Aldini S, Lin C-T. Computational model of robot trust in human co-worker for physical human-robot collaboration. IEEE Robot Autom Lett 2022;7(2):3146–53. http://dx.doi.org/10.1109/LRA.2022.3145957.
[54] Dorbala VS, Srinivasan A, Bera A. Can a robot trust you?: A DRL-based approach to trust-driven human-guided navigation. In: 2021 IEEE international conference on robotics and automation. ICRA, IEEE; 2021, p. 3538–45. http://dx.doi.org/10.1109/ICRA48506.2021.9561983.

[55] Li Y, Xu J, Guo D, Liu H. Trust-aware human-robot fusion decision-making for emergency indoor patrolling. IEEE Trans Autom Sci Eng 2024. http://dx.doi.org/10.1109/TASE.2024.3350639.
[56] Azevedo-Sa H, Yang XJ, Robert LP, Tilbury DM. A unified bi-directional model for natural and artificial trust in human–robot collaboration. IEEE Robot Autom Lett 2021;6(3):5913–20. http://dx.doi.org/10.1109/LRA.2021.3088082.
[57] Rahman SM, Wang Y. Mutual trust-based subtask allocation for human–robot collaboration in flexible lightweight assembly in manufacturing. Mechatronics 2018;54:94–109. http://dx.doi.org/10.1016/j.mechatronics.2018.07.007.
[58] Li Y, Cui R, Yan W, Zhang S, Yang C. Reconciling conflicting intents: Bidirectional trust-based variable autonomy for mobile robots. IEEE Robot Autom Lett 2024. http://dx.doi.org/10.1109/LRA.2024.3396100.
[59] Ali A, Azevedo-Sa H, Tilbury DM, Robert Jr LP. Heterogeneous human–robot task allocation based on artificial trust. Sci Rep 2022;12(1):15304. http://dx.doi.org/10.1038/s41598-022-19140-5.
[60] Kaigom EG, Roßmann J. Value-driven robotic digital twins in cyber–physical applications. IEEE Trans Ind Informatics 2020;17(5):3609–19. http://dx.doi.org/10.1109/TII.2020.3011062.
[61] Alatise MB, Hancke GP. A review on challenges of autonomous mobile robot and sensor fusion methods. IEEE Access 2020;8:39830–46. http://dx.doi.org/10.1109/ACCESS.2020.2975643.
[62] Cao MY, Laws S, y Baena FR. Six-axis force/torque sensors for robotics applications: A review. IEEE Sens J 2021;21(24):27238–51. http://dx.doi.org/10.1109/JSEN.2021.3123638.
[63] Nonomura Y. Sensor technologies for automobiles and robots. IEEJ Trans Electr Electron Eng 2020;15(7):984–94. http://dx.doi.org/10.1002/tee.23142.
[64] Macenski S, Foote T, Gerkey B, Lalancette C, Woodall W. Robot operating system 2: Design, architecture, and uses in the wild. Sci Robot 2022;7(66):eabm6074. http://dx.doi.org/10.1126/scirobotics.abm6074.
[65] Kazerouni IA, Fitzgerald L, Dooly G, Toal D. A survey of state-of-the-art on visual SLAM. Expert Syst Appl 2022;205:117734. http://dx.doi.org/10.1016/j.eswa.2022.117734.
[66] Khan MU, Zaidi SAA, Ishtiaq A, Bukhari SUR, Samer S, Farman A. A comparative survey of lidar-slam and lidar based sensor technologies. In: 2021 mohammad ali jinnah university international conference on computing. MAJICC, IEEE; 2021, p. 1–8. http://dx.doi.org/10.1109/MAJICC53071.2021.9526266.
[67] Chang L, Niu X, Liu T, Tang J, Qian C. GNSS/INS/LiDAR-SLAM integrated navigation system based on graph optimization. Remote Sens 2019;11(9):1009. http://dx.doi.org/10.3390/rs11091009.
[68] Jiang P, Ergu D, Liu F, Cai Y, Ma B. A review of yolo algorithm developments. Procedia Comput Sci 2022;199:1066–73. http://dx.doi.org/10.1016/j.procs.2022.01.135.
[69] Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C-Y, Berg AC. SSD: Single shot multibox detector. In: Computer vision–ECCV 2016: 14th European conference, amsterdam, the netherlands, October 11–14, 2016, proceedings, part i 14. Springer; 2016, p. 21–37. http://dx.doi.org/10.1007/978-3-319-46448-0_2.
[70] Wang T, Zheng P, Li S, Wang L. Multimodal human–robot interaction for human-centric smart manufacturing: A survey. Adv Intell Syst 2024;6(3):2300359. http://dx.doi.org/10.1002/aisy.202300359.
[71] Selvaggio M, Cognetti M, Nikolaidis S, Ivaldi S, Siciliano B. Autonomy in physical human-robot interaction: A brief survey. IEEE Robot Autom Lett 2021;6(4):7989–96. http://dx.doi.org/10.1109/LRA.2021.3100603.
[72] van Amsterdam B, Clarkson MJ, Stoyanov D. Gesture recognition in robotic surgery: A review. IEEE Trans Biomed Eng 2021;68(6). http://dx.doi.org/10.1109/TBME.2021.3054828.
[73] Wan S, Gu Z, Ni Q. Cognitive computing and wireless communications on the edge for healthcare service robots. Comput Commun 2020;149:99–106. http://dx.doi.org/10.1016/j.comcom.2019.10.012.
[74] Cong Y, Gu C, Zhang T, Gao Y. Underwater robot sensing technology: A survey. Fundam Res 2021;1(3):337–45. http://dx.doi.org/10.1016/j.fmre.2021.03.002.
[75] Mukherjee D, Gupta K, Chang LH, Najjaran H. A survey of robot learning strategies for human-robot collaboration in industrial settings. Robot Comput-Integr Manuf 2022;73:102231. http://dx.doi.org/10.1016/j.rcim.2021.102231.
[76] Li H, Baucom B, Georgiou P. Linking emotions to behaviors through deep transfer learning. PeerJ Comput Sci 2020;6:e246. http://dx.doi.org/10.7717/peerj-cs.246.
[77] Wang Y, Bennani IL, Liu X, Sun M, Zhou Y. Electricity consumer characteristics identification: A federated learning approach. IEEE Trans Smart Grid 2021;12(4):3637–47. http://dx.doi.org/10.1109/TSG.2021.3066577.
[78] Chatterjee A, Gupta U, Chinnakotla MK, Srikanth R, Galley M, Agrawal P. Understanding emotions in text using deep learning and big data. Comput Hum Behav 2019;93:309–17. http://dx.doi.org/10.1016/j.chb.2018.12.029.
[79] Zheng X, Lu J, Kiritsis D. The emergence of cognitive digital twin: Vision, challenges and opportunities. Int J Prod Res 2022;60(24):7610–32. http://dx.doi.org/10.1080/00207543.2021.2014591.
[80] Li J, Xu Z, Zhu D, Dong K, Yan T, Zeng Z, Yang SX. Bio-inspired intelligence with applications to robotics: A survey. 2022, http://dx.doi.org/10.48550/arXiv.2206.08544, arXiv preprint arXiv:2206.08544.
[81] Nguyen A-T, Taniguchi T, Eciolaza L, Campos V, Palhares R, Sugeno M. Fuzzy control systems: Past, present and future. IEEE Comput Intell Mag 2019;14(1):56–68. http://dx.doi.org/10.1109/MCI.2018.2881644.
[82] Liu J, Chen Z, Zhang Y, Li W. Path planning of mobile robots based on improved genetic algorithm. In: Proceedings of the 2020 2nd international conference on robotics, intelligent control and artificial intelligence. 2020, p. 49–53. http://dx.doi.org/10.1145/3438872.3439054.
[83] Schranz M, Di Caro GA, Schmickl T, Elmenreich W, Arvin F, Şekercioğlu A, Sende M. Swarm intelligence and cyber-physical systems: Concepts, challenges and future trends. Swarm Evol Comput 2021;60:100762. http://dx.doi.org/10.1016/j.swevo.2020.100762.
[84] Buerkle A, Eaton W, Lohse N, Bamber T, Ferreira P. EEG based arm movement intention recognition towards enhanced safety in symbiotic human-robot collaboration. Robot Comput-Integr Manuf 2021;70:102137. http://dx.doi.org/10.1016/j.rcim.2021.102137.
[85] Bi L, Guan C, et al. A review on EMG-based motor intention prediction of continuous human upper limb motion for human-robot collaboration. Biomed Signal Process Control 2019;51:113–27. http://dx.doi.org/10.1016/j.bspc.2019.02.011.
[86] Villani V, Righi M, Sabattini L, Secchi C. Wearable devices for the assessment of cognitive effort for human–robot interaction. IEEE Sensors J 2020;20(21):13047–56. http://dx.doi.org/10.1109/JSEN.2020.3001635.
[87] Gervasi R, Aliev K, Mastrogiacomo L, Franceschini F. User experience and physiological response in human-robot collaboration: A preliminary investigation. J Intell Robot Syst 2022;106(2):36. http://dx.doi.org/10.1007/s10846-022-01744-8.
[88] Zhao B, Wang Z, Yu Z, Guo B. EmotionSense: Emotion recognition based on wearable wristband. In: 2018 IEEE smartWorld, ubiquitous intelligence & computing, advanced & trusted computing, scalable computing & communications, cloud & big data computing, internet of people and smart city innovation (smartWorld/SCALCOM/UIC/aTC/cBDCom/IOP/sCI). IEEE; 2018, p. 346–55. http://dx.doi.org/10.1109/SmartWorld.2018.00091.
[89] Verna E, Puttero S, Genta G, Galetto M. A novel diagnostic tool for human-centric quality monitoring in human-robot collaboration manufacturing. J Manuf Sci Eng 2023;145(12). http://dx.doi.org/10.1115/1.4063284.
[90] Efron R. What is perception? In: Proceedings of the boston colloquium for the philosophy of science 1966/1968. Springer; 1969, p. 137–73. http://dx.doi.org/10.1007/978-94-010-3378-7_4.
[91] Gentsch A, Weber A, Synofzik M, Vosgerau G, Schütz-Bosbach S. Towards a common framework of grounded action cognition: Relating motor control, perception and cognition. Cognition 2016;146:81–9. http://dx.doi.org/10.1016/j.cognition.2015.09.010.
[92] Sharma N, Gedeon T. Objective measures, sensors and computational techniques for stress recognition and classification: A survey. Comput Methods Programs Biomed 2012;108(3):1287–301. http://dx.doi.org/10.1016/j.cmpb.2012.07.003.
[93] Carneiro D, Castillo JC, Novais P, Fernández-Caballero A, Neves J. Multimodal behavioral analysis for non-invasive stress detection. Expert Syst Appl 2012;39(18):13376–89. http://dx.doi.org/10.1016/j.eswa.2012.05.065.
[94] Critchley HD, Garfinkel SN. The influence of physiological signals on cognition. Curr Opin Behav Sci 2018;19:13–8. http://dx.doi.org/10.1016/j.cobeha.2017.08.014.
[95] Uppal S, Bhagat S, Hazarika D, Majumder N, Poria S, Zimmermann R, Zadeh A. Multimodal research in vision and language: A review of current and emerging trends. Inf Fusion 2022;77:149–71. http://dx.doi.org/10.1016/j.inffus.2021.07.009.
[96] Kenfack FK, Siddiky FA, Balint-Benczedi F, Beetz M. Robotvqa—a scene-graph-and deep-learning-based visual question answering system for robot manipulation. In: 2020 IEEE/RSJ international conference on intelligent robots and systems. IROS, IEEE; 2020, p. 9667–74. http://dx.doi.org/10.1109/IROS45743.2020.9341186.
[97] Thomason J, Murray M, Cakmak M, Zettlemoyer L. Vision-and-dialog navigation. In: Conference on robot learning. PMLR; 2020, p. 394–406. http://dx.doi.org/10.48550/arXiv.1907.04957.
[98] Venkatesh SG, Biswas A, Upadrashta R, Srinivasan V, Talukdar P, Amrutur B. Spatial reasoning from natural language instructions for robot manipulation. In: 2021 IEEE international conference on robotics and automation. ICRA, IEEE; 2021, p. 11196–202. http://dx.doi.org/10.1109/ICRA48506.2021.9560895.
[99] Ghazali AS, Ham J, Barakova EI, Markopoulos P. Effects of robot facial characteristics and gender in persuasive human-robot interaction. Front Robot AI 2018;5:73. http://dx.doi.org/10.3389/frobt.2018.00073.
[100] You S, Robert Jr LP. Human-robot similarity and willingness to work with a robotic co-worker. In: Proceedings of the 2018 ACM/IEEE international conference on human-robot interaction. 2018, p. 251–60. http://dx.doi.org/10.1145/3171221.3171281.
[101] Schuldt C, Laptev I, Caputo B. Recognizing human actions: A local SVM approach. In: Proceedings of the 17th international conference on pattern recognition, 2004. ICPR 2004.. vol. 3, IEEE; 2004, p. 32–6. http://dx.doi.org/10.1109/ICPR.2004.1334462.

[102] Park S, Aggarwal JK. A hierarchical Bayesian network for event recognition of human actions and interactions. Multimedia Syst 2004;10:164–79. http://dx.doi.org/10.1007/s00530-004-0148-1.

[103] Tu Z, Xie W, Qin Q, Poppe R, Veltkamp RC, Li B, Yuan J. Multi-stream CNN: Learning representations based on human-related regions for action recognition. Pattern Recognit 2018;79:32–43. http://dx.doi.org/10.1016/j.patcog.2018.01.020.

[104] Mutegeki R, Han DS. A CNN-LSTM approach to human activity recognition. In: 2020 international conference on artificial intelligence in information and communication. ICAIIC, IEEE; 2020, p. 362–6. http://dx.doi.org/10.1109/ICAIIC48513.2020.9065078.

[105] Jlidi N, Snoun A, Bouchrika T, Jemai O, Zaied M. PTLHAR: PoseNet and transfer learning for human activities recognition based on body articulations. In: Twelfth international conference on machine vision. ICMV 2019, vol. 11433, SPIE; 2020, p. 187–94. http://dx.doi.org/10.1117/12.2559567.

[106] Gao Q, Liu J, Ju Z. Robust real-time hand detection and localization for space human–robot interaction based on deep learning. Neurocomputing 2020;390:198–206. http://dx.doi.org/10.1016/j.neucom.2019.02.066.

[107] Sharotry A, Jimenez JA, Mediavilla FAM, Wierschem D, Koldenhoven RM, Valles D. Manufacturing operator ergonomics: A conceptual digital twin approach to detect biomechanical fatigue. Ieee Access 2022;10:12774–91. http://dx.doi.org/10.1109/ACCESS.2022.3145984.

[108] Li W, Rentemeister M, Badeda J, Jöst D, Schulte D, Sauer DU. Digital twin for battery systems: Cloud battery management system with online state-of-charge and state-of-health estimation. J Energy Storage 2020;30:101557. http://dx.doi.org/10.1016/j.est.2020.101557.

[109] Wang W, Wang J, Tian J, Lu J, Xiong R. Application of digital twin in smart battery management systems. Chin J Mech Eng 2021;34(1):57. http://dx.doi.org/10.1186/s10033-021-00577-0.

[110] Verna E, Puttero S, Genta G, Galetto M. Exploring the effects of perceived complexity criteria on performance measures of human-robot collaborative assembly. J Manuf Sci Eng 2023;145(10):101014. http://dx.doi.org/10.1115/1.4063232.

[111] Capponi M, Gervasi R, Mastrogiacomo L, Franceschini F. Assessing perceived assembly complexity in human-robot collaboration processes: A proposal based on Thurstone's law of comparative judgement. Int J Prod Res 2024;62(14):5315–35. http://dx.doi.org/10.1007/s10846-022-01744-8.

[112] Habay J, Van Cutsem J, Verschueren J, De Bock S, Proost M, De Wachter J, Tassignon B, Meeusen R, Roelands B. Mental fatigue and sport-specific psychomotor performance: A systematic review. Sports Med 2021;51:1527–48. http://dx.doi.org/10.1007/s40279-021-01429-6.

[113] Losey DP, McDonald CG, Battaglia E, O'Malley MK. A review of intent detection, arbitration, and communication aspects of shared control for physical human–robot interaction. Appl Mech Rev 2018;70(1):010804. http://dx.doi.org/10.1115/1.4039145.

[114] Liu H, Liu T, Zhang Z, Sangaiah AK, Yang B, Li Y. ARHPE: Asymmetric relation-aware representation learning for head pose estimation in industrial human–computer interaction. IEEE Trans Ind Inform 2022;18(10):7107–17. http://dx.doi.org/10.1109/TII.2022.3143605.

[115] Yang B, Huang J, Chen X, Li X, Hasegawa Y. Natural grasp intention recognition based on gaze in human–robot interaction. IEEE J Biomed Heal Inform 2023;27(4):2059–70. http://dx.doi.org/10.1109/JBHI.2023.3238406.

[116] Mortlock T, Muthirayan D, Yu S-Y, Khargonekar PP, Al Faruque MA. Graph learning for cognitive digital twins in manufacturing systems. IEEE Trans Emerg Top Comput 2021;10(1):34–45. http://dx.doi.org/10.1109/TETC.2021.3132251.

[117] Shi Y, Shen W, Wang L, Longo F, Nicoletti L, Padovano A. A cognitive digital twins framework for human-robot collaboration. Procedia Comput Sci 2022;200:1867–74. http://dx.doi.org/10.1016/j.procs.2022.01.387.

[118] Lv J, Li X, Sun Y, Zheng Y, Bao J. A bio-inspired LIDA cognitive-based digital twin architecture for unmanned maintenance of machine tools. Robot Comput-Integr Manuf 2023;80:102489. http://dx.doi.org/10.1016/j.rcim.2022.102489.

[119] Segerstrom SC, Nes LS. Heart rate variability reflects self-regulatory strength, effort, and fatigue. Psychol Sci 2007;18(3):275–81. http://dx.doi.org/10.1111/j.1467-9280.2007.01888.x.

[120] Guidi A, Greco A, Felici F, Leo A, Ricciardi E, Bianchi M, Bicchi A, Valenza G, Scilingo EP. Heart rate variability analysis during muscle fatigue due to prolonged isometric contraction. In: 2017 39th annual international conference of the IEEE engineering in medicine and biology society. EMBC, IEEE; 2017, p. 1324–7. http://dx.doi.org/10.1109/EMBC.2017.8037076.

[121] Bogue R. Sensors for robotic perception. Part one: Human interaction and intentions. Ind Robot: An Int J 2015;42(5):386–91. http://dx.doi.org/10.1108/IR-05-2015-0098.

[122] Honig S, Oron-Gilad T. Understanding and resolving failures in human-robot interaction: Literature review and model development. Front Psychol 2018;9:351644. http://dx.doi.org/10.3389/fpsyg.2018.00861.

[123] Park MG, Jeon JH, Lee MC. Obstacle avoidance for mobile robots using artificial potential field approach with simulated annealing. In: ISIE 2001. 2001 IEEE international symposium on industrial electronics proceedings (cat. no. 01TH8570). vol. 3, IEEE; 2001, p. 1530–5. http://dx.doi.org/10.1109/ISIE.2001.931933.