

# Review on User Interaction for Robotic Arm in Digital Twin

Aiman Hakim Azahari<sup>1</sup>, Mohd Khalid Mokhtar<sup>1</sup>, Nazreen Abdullasim<sup>1</sup>, Mohd Hafiz Zakaria<sup>1</sup>,  
Asniyani Nur Haidar Abdullah<sup>1</sup>, Shafina Abd Karim Ishigaki<sup>1</sup>, Ikmal Faiq Albakri Mustafa Albakri<sup>1</sup>,  
Muhammad Najib Zamri<sup>2</sup>

<sup>1</sup> Pervasive Computing & Educational Technology, Department of Media Interactive Fakulti Teknologi Maklumat  
Komunikasi, Universiti Teknikal Malaysia Melaka, Malaysia

<sup>2</sup> IoT and Smart Technologies Research Group, University of Southampton Malaysia, Iskandar Puteri, Johor, Malaysia

## ARTICLE INFORMATION

### Article History:

Received 15 January 2026

Revised 19 April 2026

Accepted 18 May 2026

### Keywords:

Virtual Robotic Arm Control;  
Human-Computer Interaction;  
Motion Controllers;  
Hand Gesture Recognition;  
Usability Evaluation;  
Digital Twin;  
Virtual Reality;  
Haptic Feedback

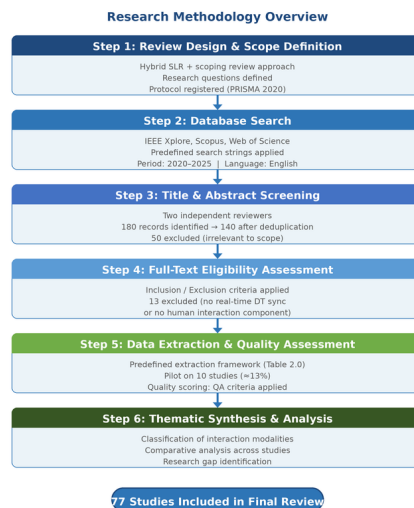
### Corresponding Author:

Mohd Khalid Mokhtar,  
Pervasive Computing &  
Educational Technology,  
Department of Media Interactive  
Fakulti Teknologi Maklumat  
Komunikasi, Universiti Teknikal  
Malaysia Melaka, Malaysia.  
Email:  
[khalid.mokhtar@utem.edu.my](mailto:khalid.mokhtar@utem.edu.my)

This work is open access under a  
[Creative Commons Attribution-Share  
Alike 4.0](https://creativecommons.org/licenses/by-sa/4.0/)



## ABSTRACT



The integration of human interaction techniques in digital twin (DT) systems has become increasingly important in manufacturing, industrial automation, and remote operations, particularly for robotic arm control. However, existing approaches joystick control, gesture-based input, and virtual reality (VR) are often disconnected across modalities, limiting effectiveness in real-time environments. The research contribution is a systematic literature review (SLR) that critically analyzes and synthesizes interaction techniques to identify performance trends, evaluation gaps, and design challenges in virtual robotic arm control within digital twin frameworks. The review covers studies published between 2020 and 2025, selected to reflect the rapid emergence of immersive technologies in real-time digital twin systems. Following PRISMA 2020 guidelines, 180 records were identified from IEEE Xplore, Scopus, and Web of Science, from which 77 peer-reviewed studies were selected. Interaction techniques were evaluated using task completion time, positional accuracy, NASA Task Load Index (NASA-TLX), and System Usability Scale (SUS). The findings reveal that VR-based techniques dominate due to their intuitiveness and immersive experience in human-in-the-loop control. However, evaluation remains inconsistent across studies, with significant variation in metrics and experimental setups. Latency and synchronization were identified as critical challenges in real-time control, where delays degrade precision and responsiveness. Traditional methods such as joysticks offer stability but lack the natural interaction of immersive techniques. These findings underscore the need for standardized evaluation frameworks and improved synchronization strategies, offering practical guidance for designing robust, human-centered digital twin interaction systems for robotic arms.

### Document Citation:

A. H. Azahari, M. K. Mokhtar, N. Abdullasim, M. H. Zakaria, A. N. H. Abdullah, S. A. K. Ishigaki, I. K. A. M. Albakri, and M. N. Zamri, "Review on User Interaction for Robotic Arm in Digital Twin," *Buletin Ilmiah Sarjana Teknik Elektro*, vol. 8, no. 2, pp. 619-637, 2026, DOI: 10.12928/biste.v8i2.15851.

## 1. INTRODUCTION

Digital twin (DT) technology has emerged as a foundational component of modern cyber physical systems, enabling real-time, bidirectional synchronization between physical assets and their high-fidelity virtual counterparts. A digital twin is defined as a continuously updated virtual model that supports monitoring, simulation, prediction, and control throughout the system lifecycle [1][2]. Unlike conventional simulation models that operate independently of physical systems, digital twins maintain a live connection with real world data, enabling dynamic updates and real-time decision support [3]. Within the widely referenced five dimension DT model comprising physical entities, virtual models, connection layers, data management, and services the interaction layer serves as the critical interface through which human operators perceive, command, and collaborate with the system in real time [1]. This architecture has positioned DT technology as a key enabler of intelligent manufacturing, advanced robotics, and human–robot collaboration frameworks. In robotics, digital twins play a critical role in enabling safer system development, remote operation, and operator training, particularly for robotic arm systems [4]. Robotic arms are widely deployed in industrial tasks including pick-and-place operations, assembly, and precision manipulation applications that demand accurate motion control, continuous feedback, and reliable real-time responsiveness. Digital twin systems allow these tasks to be simulated, monitored, and optimized in real time, reducing operational risks and improving efficiency [2],[5][6]. The integration of digital twins with human–robot collaboration frameworks further enhances system flexibility [7] and supports interactive control in complex, high-stakes environments [8].

User interaction is a central component of digital twin–based robotic arm systems, directly influencing usability, task performance, and operator workload. Traditional methods such as joysticks and teach pendants remain widely used for their reliability, but they often lack intuitiveness when controlling robotic arms with multiple degrees of freedom, increasing cognitive load during complex manipulation tasks. To address these limitations, recent research has explored more natural and immersive approaches, including virtual reality (VR)-based teleoperation, gesture-based interfaces, and haptic feedback mechanisms [9][10]. These human-centered approaches aim to improve spatial awareness, enhance user experience, and enable more intuitive mapping between human motion and robotic control within human-in-the-loop environments [11]. Advancing such interaction designs carries direct practical significance: more intuitive interfaces reduce operator error and cognitive workload, improve task precision and safety in high-risk environments [12], and support more effective human–robot collaboration across industrial and remote operation contexts. Despite the growing body of research, existing studies on user interaction for robotic arm digital twins remain disconnected across interaction modalities, evaluation methods, and system design frameworks. Many works focus primarily on technical aspects such as system architecture or communication protocols, while human-centered evaluation and usability analysis are often limited or inconsistently reported. Latency and synchronization between the physical robotic arm and its digital twin also present significant ongoing challenges, as even minor delays can negatively affect control stability, responsiveness, and user experience [11]. These issues are compounded by trade-offs between synchronization strategies including local processing, cloud-based architectures, and predictive techniques none of which has emerged as a clear standard [13].

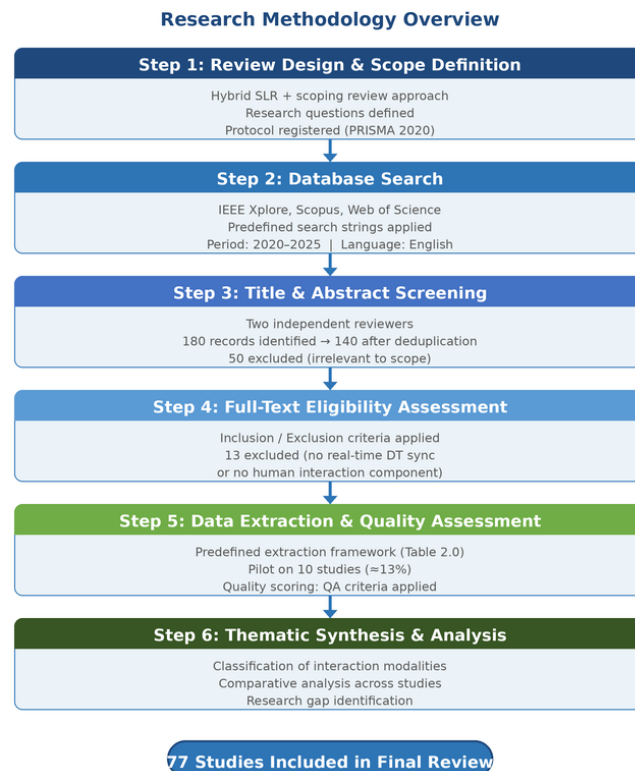
This review focuses specifically on robotic arms due to their widespread industrial deployment and their higher interaction complexity relative to other robotic systems such as mobile robots or unmanned aerial vehicles. Unlike mobile or aerial systems, robotic arm manipulation demands precise multi-axis coordination, continuous real-time feedback, and intuitive operator control making the quality of human interaction a decisive factor in system performance and safety. The review scope covers studies published between 2020 and 2025, a period marked by rapid advancement in immersive technologies and real-time digital twin infrastructure, ensuring the analysis captures the most current developments in the field. To address these gaps, this paper conducts a systematic literature review (SLR) following PRISMA 2020 guidelines, searching IEEE Xplore, Scopus, and Web of Science for peer-reviewed studies published between 2020 and 2025. Unlike prior reviews that focus narrowly on a single modality or system architecture, this review provides a cross-modal, human-centered synthesis that integrates interaction classification, evaluation critique, and research gap identification within a unified structured framework. The research contribution of this paper is threefold. First, it provides a structured classification of user interaction techniques for robotic arms in digital twin environments. Second, it critically analyzes existing evaluation methods and highlights inconsistencies in human-centered assessment across studies. Third, it identifies key research gaps and proposes directions for developing more effective, responsive, and human-centered digital twin interaction systems for robotic arms [14].

## 2. METHODS

### 2.1. Review Type and Strategy

This study adopts a hybrid systematic and scoping review approach to analyze user interaction techniques for robotic arms in digital twin (DT) environments. A systematic literature review (SLR) methodology is employed to ensure rigor, transparency, and reproducibility through predefined protocols, structured search processes, and explicit inclusion and exclusion criteria [15][16]. Systematic reviews are widely recognized for minimizing selection bias and enabling consistent synthesis of evidence across studies, particularly in multidisciplinary domains. At the same time, elements of a scoping review are incorporated to capture the breadth and diversity of research in this rapidly evolving field. Scoping reviews are particularly suitable for mapping key concepts, identifying research trends, and exploring heterogeneous study designs without restricting the analysis to narrowly defined experimental frameworks [17]. This is important in the context of digital twin based robotic systems, where studies vary significantly in terms of interaction modalities, system architectures, and evaluation approaches.

Established guidelines for rigorous review conduct further reinforce the structured screening and synthesis protocols applied in this study, ensuring that methodological decisions are grounded in validated evidence synthesis practices [18]. To operationalize this hybrid strategy, the systematic component is applied during the study selection process, including database searching, screening, and eligibility assessment based on predefined criteria. In contrast, the scoping component is applied during data synthesis, where studies are categorized and analyzed based on interaction type, platform, and evaluation methods. This approach ensures methodological rigor while allowing flexibility in capturing emerging trends and diverse research contributions. By combining systematic and scoping methodologies, this study provides both a structured and comprehensive understanding of user interaction techniques for robotic arms in digital twin environments. To maintain consistency and reduce potential bias, screening and data extraction were conducted independently by two reviewers. Any disagreements arising from differences between the systematic and scoping components such as borderline eligibility decisions or categorization conflicts were resolved through discussion and consensus, with a third reviewer consulted where agreement could not be reached. Figure 1 provides an overview of the overall research methodology, illustrating the sequential steps from review design and database searching through study screening, eligibility assessment, data extraction, and final synthesis.



**Figure 1.** Overview of Research Methodology

Figure 1 presents the overall research methodology adopted in this study, following a structured hybrid Systematic Literature Review (SLR) approach guided by PRISMA 2020 standards. The process begins with

the review design and scope definition, where research questions are formulated and a review protocol is established to ensure methodological rigor. This is followed by the database search phase using IEEE Xplore, Scopus, and Web of Science with predefined keyword combinations and filters (2020–2025, English, peer-reviewed) to ensure the relevance and quality of retrieved studies. Subsequently, title and abstract screening is conducted by two independent reviewers after removing duplicate records to identify studies aligned with the research scope. The selected studies then undergo full-text eligibility assessment based on defined inclusion and exclusion criteria, excluding those lacking real-time digital twin synchronization or human interaction components. Next, data extraction and quality assessment are performed using a predefined framework, including pilot testing and quality scoring to ensure consistency and reliability. Finally, thematic synthesis and analysis are carried out by classifying interaction modalities, comparing findings across studies, and identifying research gaps. This systematic process results in a final selection of 27 studies, forming a comprehensive foundation for further analysis and discussion.

## 2.2. Search Databases and Keywords

A comprehensive and systematic search strategy was employed to identify relevant studies on user interaction techniques for robotic arms in digital twin (DT) environments. The search process was designed in accordance with the PRISMA 2020 guidelines to ensure transparency, reproducibility, and completeness in reporting the literature search procedure [19]. In addition, principles from PRISMA-S were considered to enhance the documentation and structure of search strategies across multiple databases [20]. Three major academic databases were selected for this review: IEEE Xplore, Scopus, and Web of Science. These databases were chosen due to their extensive coverage of peer-reviewed literature in robotics, human–computer interaction, and cyber–physical systems. Prior studies have emphasized that using multiple databases improves retrieval quality and reduces the risk of missing relevant publications in systematic reviews [21]. The combination of these databases ensures a broad and multidisciplinary coverage of the research domain. Table 1 summarizes the databases used, search terms, and filtering criteria applied in this study. The selection of databases and search strategy is supported by established systematic review guidelines and retrieval studies [19][20]. For Scopus, the adapted search string used was: TITLE-ABS-KEY (“digital twin” AND (“robotic arm” OR “robot manipulator”) AND (“user interaction” OR “teleoperation” OR “virtual reality” OR “gesture control” OR “haptic feedback”)) AND PUBYEAR > 2019 AND PUBYEAR < 2026 AND LANGUAGE (english). An equivalent field-tagged string was applied in Web of Science using the TS= (Topic) field operator. The complete search strings are available from the corresponding author upon request. To validate the search strategy, a pilot search was conducted prior to full execution, and the resulting string was reviewed by a subject expert to confirm coverage and minimize retrieval gaps.

Table 1 summarizes the search strategy applied across the three selected databases. IEEE Xplore was selected as the primary source due to its extensive coverage of robotics, control systems, and engineering publications [19][20]. Scopus and Web of Science were included to broaden interdisciplinary coverage across human–computer interaction, cyber–physical systems, and digital twin research [21]. All three databases were searched using consistent keyword combinations centred on digital twin, robotic arm, user interaction, teleoperation, virtual reality, gesture control, and haptic feedback, with filters restricting results to peer-reviewed English-language publications from 2020 to 2025. This multi-database approach ensures comprehensive retrieval and reduces the risk of missing relevant publications across disciplinary boundaries [19][20].

**Table 1.** Summary of Search Strategy and Data Sources

Database	Search Keywords / Query	Filters Applied	Supporting Reference	Purpose
IEEE Xplore	(“digital twin” AND “robotic arm” OR “robot manipulator”) AND (“user interaction” OR “teleoperation” OR “virtual reality” OR “gesture control” OR “haptic feedback”)	2020–2025, English, Peer-reviewed	[19][20]	Core database for robotics and engineering studies
Scopus	Adapted keyword combination based on database syntax	2020–2025, English, Peer-reviewed	[21]	Broad interdisciplinary coverage
Web of Science	Adapted keyword combination based on database syntax	2020–2025, English, Peer-reviewed	[21]	High-quality indexed journals and conferences

## 2.3. Inclusion and Exclusion Criteria

To ensure the relevance, quality, and consistency of the reviewed literature, explicit inclusion and exclusion criteria were defined prior to the screening process. Establishing clear selection criteria is a

fundamental step in systematic literature reviews, as it reduces selection bias and enhances transparency and reproducibility [22][23]. In addition, well-defined criteria help ensure that only studies aligned with the research objectives are included, particularly in multidisciplinary domains such as digital twin-based robotic systems [21]. Studies were included if they met the following conditions: (1) published in peer-reviewed journals or conference proceedings; (2) published between 2020 and 2025; (3) involved robotic arms or manipulators; (4) implemented a digital twin, defined in this study as a synchronized virtual representation of a physical system with real-time bidirectional data exchange; and (5) incorporated human interaction, such as teleoperation, control interfaces, or usability evaluation. The operational definition of a digital twin was applied to distinguish it from conventional simulation or visualization systems, ensuring that only studies with real-time synchronization and interaction capabilities were included.

Studies were excluded if they: (1) focused solely on simulation models without real-time synchronization; (2) involved robotic systems without human interaction, such as fully autonomous or mobile robots; (3) were non-peer-reviewed sources, including theses, technical reports, patents, or editorials; or (4) were not written in English. In addition to database filtering, backward snowballing was applied by reviewing the reference lists of selected studies to identify additional relevant publications that may have been missed during the initial search process [24]. For the purposes of this review, studies published between 2020 and 2025 are defined as "recent," reflecting a period of rapid technological advancement in immersive systems and digital twin infrastructure within the robotics domain. These inclusion and exclusion criteria ensure that the selected studies are recent, methodologically sound, and directly relevant to the research scope, thereby supporting a focused and reliable synthesis of user interaction techniques for robotic arms in digital twin environments. Table 2 provides a structured summary of the inclusion and exclusion criteria applied in this review.

Table 2 presents the inclusion and exclusion criteria applied during the study selection process. The inclusion criteria were designed to ensure that only studies directly relevant to human interaction with robotic arms in digital twin environments were retained, focusing on peer-reviewed publications from 2020 to 2025 with real-time synchronization and human interaction components [22][23]. The exclusion criteria systematically removed studies that lacked these core attributes, including autonomous systems without operator involvement, non-peer-reviewed sources, and studies reporting only technical performance without any human-centered usability assessment [24]. Together, these criteria ensured a rigorous and focused evidence base for the subsequent synthesis.

**Table 2.** Inclusion and Exclusion Criteria for Study Selection

#	Inclusion Criteria	Exclusion Criteria	Ref.
1	Published in peer-reviewed journals or conference proceedings	Simulation models without real-time synchronization	[22]
2	Published between 2020 and 2025 (recent period)	Robotic systems without human interaction (fully autonomous or mobile robots)	[21]
3	Study involves robotic arms or manipulators	Non-peer-reviewed sources (theses, technical reports, patents, editorials)	[23]
4	Implements a digital twin with real-time bidirectional data exchange	Studies not written in English	[22][23]
5	Incorporates human interaction (teleoperation, control interfaces, or usability evaluation)	Studies focused solely on autonomous robotic control without any human interaction component	[24]

#### 2.4. Study Selection Process

The study selection process was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines to ensure transparency, reproducibility, and methodological rigor in identifying and selecting relevant studies [19],[25]. The PRISMA framework is widely adopted in systematic literature reviews, including engineering and technology domains, as it provides a structured approach to minimize selection bias and improve reporting clarity [26]. The selection process consisted of four main stages: identification, screening, eligibility assessment, and final inclusion. During the identification stage, a total of 180 records were retrieved from IEEE Xplore, Scopus, and Web of Science using predefined search strategies. After removing duplicate records ( $n = 40$ ), 140 unique records remained for further analysis. In the screening stage, titles and abstracts of the retrieved records were reviewed to assess their relevance to robotic arms, digital twin systems, and user interaction. During this stage, 50 records were excluded due to irrelevance to the research scope, including studies not related to robotic manipulators or lacking a digital twin component. The remaining 90 studies were retained for full-text evaluation. During the eligibility stage, full-text articles were assessed based on the inclusion and exclusion criteria defined in Section 3.3. A total of 13 studies were excluded at this stage due to not meeting the operational definition of a digital twin (e.g., absence of real-time synchronization) or lacking a human interaction component. This step ensured that only studies aligned with the research objectives were included.

Following the eligibility assessment, a total of 77 studies were selected for inclusion in this review. The complete study selection process is illustrated in Figure 2, which presents the PRISMA 2020 flow diagram for this review. The diagram provides a transparent overview of the systematic filtering process across all four stages. In the identification stage, 180 records were retrieved from three databases, from which 40 duplicates were removed, yielding 140 unique records. During screening, 50 records were excluded based on title and abstract review due to irrelevance to robotic arms, digital twin systems, or human interaction. In the eligibility stage, 13 full-text articles were excluded for failing to meet the operational definition of a digital twin or for lacking a human interaction component. The remaining 77 studies were retained for final inclusion and form the evidence base for this review. Each reduction stage is documented in the diagram to ensure full reproducibility and transparency in accordance with PRISMA 2020 reporting standards [19],[25].

Figure 2 illustrates the PRISMA flow diagram of the study selection process used in this research. The process begins with the identification of 180 records from three databases, namely IEEE Xplore, Scopus, and Web of Science. After removing 40 duplicate records, 140 records remain for screening. During the title and abstract screening stage, 40 records are excluded due to irrelevance to digital twin robotic arms or not being peer-reviewed. The remaining 90 articles are then assessed for full-text eligibility, where an additional 50 records are excluded for not meeting the study criteria, including lack of relevance to digital twin robotic arms or non-peer-reviewed status. Subsequently, 77 studies are included for further consideration; however, 13 full-text articles are excluded due to limitations such as simulations without real-time digital twin implementation or absence of human interaction components. This systematic filtering process results in a final selection of 77 studies, ensuring that only relevant and high-quality sources are included for analysis.

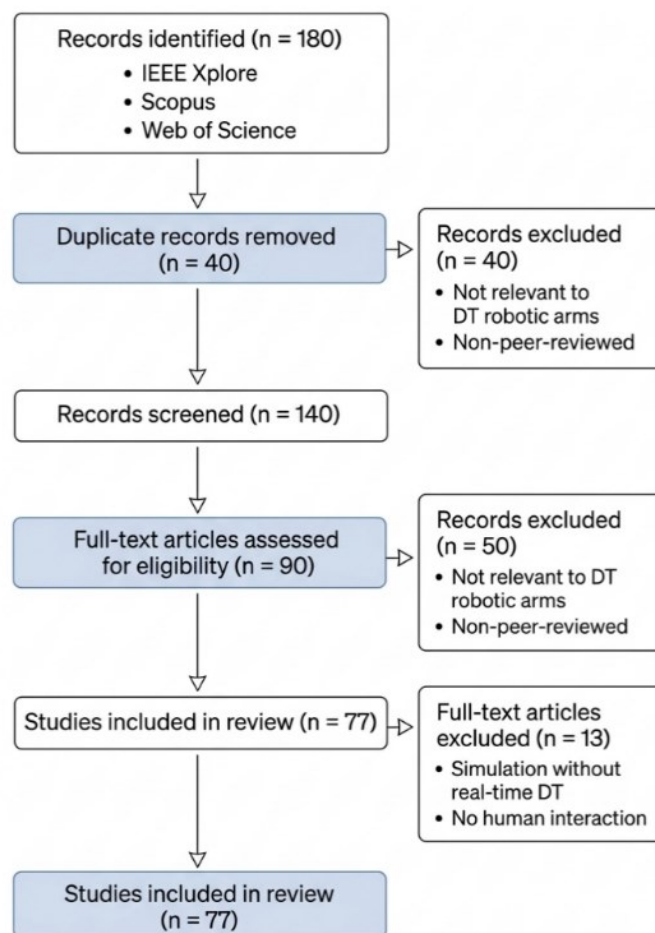


Figure 2. PRISMA Flow Diagram of the Study Selection Process

## 2.5. Data Extraction and Quality Assessment

Following the study selection process, a structured data extraction and quality assessment procedure was conducted to systematically analyze the selected studies. Data extraction is a critical step in systematic literature reviews, as it enables consistent comparison and synthesis of findings across multiple studies [22][23]. In

addition, systematic mapping principles were considered to support the categorization and organization of research evidence within this domain [27][28]. Table 3 summarizes the data extraction categories and quality assessment criteria applied in this study. A predefined data extraction framework was developed to capture key information from each study. The extracted data were categorized into three main components: interaction type, platform, and evaluation metrics. Interaction type refers to the user interface modality used to control or interact with the robotic arm, including joystick-based control, virtual reality (VR) interfaces, gesture-based interaction, and haptic feedback systems. Platform information includes both software environments, such as simulation engines and middleware frameworks, and hardware components, including robotic manipulators, sensors, and input devices. Evaluation metrics consist of both objective measures, such as task completion time, accuracy, and latency, as well as subjective measures, including the System Usability Scale (SUS) and NASA Task Load Index (NASA-TLX), which are widely used to quantify user experience and workload in interactive systems [29].

To ensure consistency and reliability, a pilot data extraction was conducted on ten studies (approximately 13% of the total sample), selected to represent the range of interaction modalities and study designs in the corpus. The extraction framework was refined based on this pilot process to improve clarity, resolve ambiguous category boundaries, and reduce inter-rater inconsistency. The finalized framework was then applied independently by two reviewers to all 77 included studies, ensuring uniform data collection and enabling systematic comparison across studies. Disagreements between reviewers were resolved through discussion; where consensus could not be reached, a third reviewer provided a final judgment. In addition to data extraction, a quality assessment was performed to evaluate the methodological rigor and relevance of the included studies. Each study was assessed based on predefined criteria adapted from established evaluation frameworks [22], including: (1) publication type (journal or conference), (2) indexing status (e.g., Scopus or Web of Science), (3) clarity of digital twin implementation, (4) presence of experimental validation or user study, and (5) reporting of evaluation metrics. Studies that demonstrated clear digital twin implementation and included user-centered evaluation were considered to have higher methodological quality. In terms of publication weighting, journal articles were assigned higher quality weight than conference papers, reflecting the more rigorous peer-review process typically applied in journal publication [22].

The extracted data were analysed using a qualitative comparative approach, where studies were grouped according to interaction modalities and platform characteristics. This analysis enabled the identification of research trends, strengths, and limitations across different interaction techniques. As summarized in Table 2, the framework captures the full spectrum of relevant study attributes, from interaction modality and platform to evaluation metrics and quality indicators. The distribution of studies across these categories reveals that VR-based interaction and objective performance metrics dominate the current literature, while haptic and multimodal approaches remain underrepresented and subjective usability evaluation is inconsistently applied. This pattern directly motivates the comparative analysis and gap identification presented in the subsequent sections.

Table 3 summarizes the structured data extraction framework applied to all 77 included studies. The framework captures interaction modality, platform characteristics, and evaluation metrics to enable systematic comparison across studies [27][28]. Quality assessment criteria were applied to each study based on publication type, indexing status, clarity of digital twin implementation, and presence of experimental validation or user study [22]. Studies demonstrating clear real-time synchronization and human-centered evaluation were assigned higher methodological quality scores, informing the weighting of evidence during data synthesis [22][23]. This structured approach ensured consistency and reproducibility throughout the review process.

**Table 3.** Data Extraction and Quality Assessment Framework

Category	Element	Description	Reference
Interaction Type	Control Interface	Joystick, VR, gesture-based, haptic	[27][28]
Platform	Software	Unity, ROS, simulation engines	[27]
Platform	Hardware	Robotic arm, sensors, controllers	[27]
Evaluation Metrics	Objective	Task completion time, accuracy, latency	[29]
Evaluation Metrics	Subjective	SUS, NASA-TLX	[29]
Quality Assessment	Publication Type	Journal or conference (journals weighted higher)	[22]
Quality Assessment	Indexing	Scopus / Web of Science	[21]
Quality Assessment	DT Implementation	Real-time synchronization (Yes/No)	[22][23]
Quality Assessment	Validation	User study or experiment present	[22]

### 3. LITERATURE REVIEW

#### 3.1. Digital Twin for Robotic Arms

Digital twin (DT) technology has been widely recognized as a key component in cyber-physical production systems, enabling the integration of physical assets with their virtual counterparts for real-time

monitoring, simulation, and control. A digital twin is typically defined as a dynamic and synchronized virtual representation of a physical system, supported by continuous data exchange between the physical and digital environments [30]-[32]. Unlike conventional simulation models that operate independently of physical systems, digital twins maintain a live bidirectional connection with real-world data, enabling dynamic updates and real-time decision support. This capability allows DT systems to support predictive analysis, optimization, and decision-making processes across the system lifecycle. In the context of robotics, digital twin technology has been increasingly applied to robotic arm systems, particularly in intelligent manufacturing and industrial automation. DT-driven robotic systems enable real-time visualization, remote control, and performance optimization, thereby improving operational efficiency and reducing system risks [33]. Furthermore, the integration of human-in-the-loop mechanisms within DT environments allows operators to interact with robotic systems more intuitively, enhancing control flexibility and enabling safer human-robot collaboration [34].

Despite these advancements, the implementation of digital twins in robotic arm applications presents several challenges. One key issue lies in the balance between simulation fidelity and real-time synchronization, where higher model accuracy often increases computational complexity and latency [35]. This trade-off directly affects system responsiveness and user interaction quality, particularly in applications involving teleoperation or immersive control interfaces. Overall, existing studies demonstrate that digital twin technology provides a strong foundation for enhancing robotic arm systems through real-time interaction and intelligent control [36]. However, variations in system architectures, interaction mechanisms, and synchronization strategies indicate that the field remains heterogeneous and lacks standardized frameworks. These limitations highlight the need for a more structured analysis [37] of interaction techniques and evaluation methods, which is further explored in the subsequent sections.

### 3.2. Virtual Reality (VR)

Virtual reality (VR) has emerged as a prominent interaction technique for robotic arm control in digital twin environments, particularly in applications involving teleoperation and human-robot collaboration. VR-based systems enable immersive interaction by providing users with real-time visualization and spatial mapping [38] between human motion and robotic actions. This approach enhances situational awareness and allows operators to perform complex manipulation tasks more intuitively compared to conventional control interfaces [39]-[41]. In addition to improving usability, VR technologies support more natural human-robot interaction by integrating motion tracking and intuitive control mechanisms. These capabilities are aligned with the broader development of human-centered robotic systems, where intuitive interaction plays a key role in improving efficiency and safety [42]. Furthermore, advances in wearable and sensor-based systems, such as myosignal-based control interfaces, have contributed to more seamless interaction between humans and robotic systems, supporting more adaptive and responsive control strategies [43].

Despite these advantages, VR-based interaction also presents several challenges. System latency, hardware dependency, and potential user fatigue during prolonged usage can affect the overall effectiveness of VR systems [44]. These limitations highlight the need for optimized synchronization and efficient system design to ensure stable and responsive interaction in real-time digital twin environments. Overall, VR-based interaction provides a strong foundation for intuitive and immersive robotic control. However, variations in implementation approaches and performance constraints indicate that further investigation is required to evaluate its effectiveness across different application contexts. A more detailed comparative analysis of VR-based interaction techniques in relation to other modalities is presented in the subsequent section.

### 3.3. Gesture-Based Interfaces

Gesture-based interfaces have emerged as a natural and intuitive approach for human robot interaction in digital twin based robotic systems. By allowing users to control robotic arms through hand movements and body gestures, these interfaces reduce reliance on traditional input devices [45] and enable more direct and human-centered interaction. In digital twin environments, gesture-based control is often integrated with real-time visualization and feedback mechanisms [46], supporting human-in-the-loop operation and enhancing system flexibility [34]. In addition to improving interaction intuitiveness, gesture-based systems are closely aligned with the development of cyber physical production systems, where seamless communication between human operators and robotic systems is essential [47]. These systems often rely on vision-based or sensor-based technologies to interpret user gestures and translate them into robotic commands [48][49]. Furthermore, real-time processing techniques, such as motion interpolation and smoothing, are commonly applied to improve the stability and continuity of gesture-based control, particularly in dynamic manipulation tasks [50].

However, gesture-based interaction faces several technical challenges that can affect system performance and reliability. One key limitation is sensitivity to environmental conditions, such as lighting variations and occlusions, which can reduce recognition accuracy [51]. In addition, latency and synchronization issues

between the physical robot and its digital twin can lead to delayed or inconsistent responses, negatively impacting user experience [52]. Recent approaches, including edge computing enabled architectures, have been proposed to mitigate these issues by reducing communication delays [53] and improving real-time responsiveness [54]. Overall, gesture-based interfaces provide a promising alternative to conventional control methods by enabling more natural and flexible interaction. When combined with haptic feedback [55], gesture-based systems can further enhance user perception by providing tactile confirmation of virtual contact events, creating a richer and more realistic human-in-the-loop interaction experience. Nevertheless, variations in implementation techniques and system constraints indicate that further evaluation is required to assess their effectiveness across different application scenarios. A detailed comparison of gesture-based interaction with other modalities is presented in the subsequent section.

### 3.4. Evaluation Metrics in Existing Studies

Evaluation metrics play a crucial role in assessing the effectiveness of user interaction techniques in digital twin-based robotic systems. In the context of robotic arm control, evaluation is essential not only for measuring system performance but also for understanding user experience and interaction quality [56]. Existing studies commonly adopt a combination of objective and subjective metrics to provide a comprehensive assessment of system capabilities and usability [57],[58]. Objective evaluation metrics are widely used to quantify system performance and operational efficiency. These metrics typically include task completion time, positional accuracy, trajectory smoothness, and system latency. Latency and synchronization accuracy are critical factors in digital twin environments, as delays between the physical system and its virtual representation can significantly affect control responsiveness and stability [59]. Such metrics are especially important in teleoperation scenarios, where real-time feedback and precise control are required to ensure safe and effective manipulation.

In addition to objective measurements, subjective evaluation metrics are commonly employed to assess user experience and cognitive workload. Standardized instruments such as the System Usability Scale (SUS) and NASA Task Load Index (NASA-TLX) are widely used to evaluate usability, mental effort, and overall interaction satisfaction [29],[60]. These metrics provide valuable insights into how users perceive different interaction techniques, particularly in immersive environments such as virtual reality-based teleoperation systems. Furthermore, usability studies often incorporate user feedback and experimental observations to complement quantitative performance data [61],[62]. Despite the availability of these evaluation methods, inconsistencies exist in how metrics are applied and reported across studies. Many works emphasize technical performance while providing limited user-centered evaluation, leading to an imbalance between system efficiency and usability considerations. In addition, differences in experimental design, participant selection, and task scenarios make it difficult to directly compare results across studies. These limitations highlight the lack of standardized evaluation frameworks for digital twin-based robotic interaction. The root causes of this fragmentation include disciplinary differences between the robotics and human-computer interaction communities, the absence of shared benchmark tasks, and the varying latency constraints across deployment contexts from sub-20 ms requirements for haptic feedback loops to 80–300 ms typical of WiFi-based gesture systems and the improvements achievable with 5G-enabled architectures [52],[54],[59].

Overall, existing evaluation approaches provide valuable insights into system performance and user experience; however, the absence of consistent evaluation methodologies limits the comparability and generalizability of findings. These challenges underscore the need for more structured and standardized evaluation frameworks, which are further examined in the subsequent section. Table 7 provides a structured comparison of evaluation metrics applied across different interaction modalities in the reviewed studies.

## 4. RESULT AND DISCUSSION

### 4.1. Classification of Interaction Methods

The selected studies were analyzed and categorized based on the type of interaction techniques used for controlling robotic arms in digital twin environments. The analysis reveals four primary categories of interaction modalities: virtual reality (VR)-based interaction, gesture-based interfaces, haptic feedback systems, and hybrid or multimodal approaches. This classification provides a structured overview of how user interaction has evolved toward more intuitive and human-centered control mechanisms in recent research [22],[3]. Among these categories, VR-based interaction emerges as the most dominant approach. VR systems provide immersive visualization and intuitive spatial control, enabling users to interact with robotic manipulators in a more natural and efficient manner. Several studies highlight the effectiveness of VR in teleoperation tasks, particularly in improving user awareness and precision during complex manipulation [63],[40]. This dominance can be attributed to the rapid advancement of immersive technologies and their strong alignment with human-robot collaboration requirements [8]. However, it is important to note that this

dominance may partly reflect the maturity and widespread availability of VR development tools and research infrastructure, rather than representing a definitive conclusion about the inherent superiority of VR over other interaction modalities.

Gesture-based interaction represents another significant category, focusing on enabling contactless and natural control through hand movements. These systems are particularly useful in collaborative environments where traditional input devices may limit flexibility. However, despite their intuitive nature, gesture-based approaches are less widely adopted due to challenges in recognition accuracy and environmental robustness [10]. Haptic feedback systems, on the other hand, aim to enhance interaction by providing tactile and force-based feedback to users [64]. These systems improve control accuracy and user perception, especially in precision tasks such as teleoperation. However, their adoption remains limited due to higher implementation complexity, hardware requirements, and cost considerations [65][66]. In addition to individual interaction techniques, several studies explore hybrid or multimodal approaches that combine multiple interaction modalities, such as VR with haptic feedback or gesture-based control with visual interfaces. These approaches aim to leverage the strengths of different techniques to improve overall system performance and user experience [67]. The classification of interaction techniques is summarized in Table 4, which highlights the characteristics, strengths, and limitations of each interaction modality. A detailed comparative analysis of these modalities in terms of performance, usability, and implementation complexity is presented in the following section.

Table 4 reveals that VR-based interaction accounts for the largest proportion of studies, reflecting the broader trend toward immersive human-in-the-loop control in digital twin environments. Gesture-based interfaces represent the second most studied modality, driven by the appeal of contactless and natural interaction, though their real-world adoption lags behind VR due to robustness limitations. Haptic feedback and hybrid approaches, while offering complementary strengths in precision and multimodal integration, remain underrepresented in the reviewed literature, indicating that these areas present the most significant opportunities for future investigation. The distribution of studies across these four categories underscores the field's current focus on immersive and intuitive interfaces while highlighting critical gaps in evaluation standardization and human-centered design.

**Table 4.** Frequency of User Interaction Methods for Robotic Arms in Digital Twin Studies

Interaction Type	Description	Strengths	Limitations	Key References
Virtual Reality (VR)	Immersive teleoperation with real-time visualization	High immersion, intuitive control, improved spatial awareness	Latency, hardware dependency, user fatigue	[63],[40],[8]
Gesture-Based	Hand/body motion control using vision or sensors	Natural interaction, contactless, flexible	Recognition errors, environmental sensitivity	[10]
Haptic Feedback	Force/tactile feedback for interaction	Improved precision, realistic feedback	High cost, complex integration	[65]
Hybrid / Multimodal	Combination of multiple techniques	Enhanced interaction, better usability	Increased system complexity	[67]

#### 4.2. Comparative Analysis

The comparative analysis of interaction techniques highlights significant differences in performance, usability, and implementation complexity across virtual reality (VR), gesture-based, haptic, and hybrid interaction approaches. While Section 4.1 provides a classification of these techniques, this section critically evaluates their relative strengths and limitations to identify underlying trends and challenges in digital twin-based robotic systems. VR-based interaction demonstrates superior performance in terms of user immersion and spatial awareness, making it highly effective for teleoperation and training applications. For example, Gao *et al.* [63] demonstrated significant improvements in task accuracy and operator situational awareness when using VR-based teleoperation within a digital twin framework compared to conventional control methods, while Liu *et al.* [33] demonstrated enhanced manipulation precision in digital twin-driven manufacturing tasks using immersive interfaces. However, the effectiveness of VR systems is often influenced by system latency, hardware requirements, and user fatigue, particularly in prolonged operations. Notably, some studies report that under high-latency conditions exceeding 200 ms, VR-based performance degrades to levels comparable to or worse than traditional joystick control, suggesting that the advantage of VR is conditional on adequate synchronization infrastructure [67],[57]. These trade-offs suggest that while VR offers strong usability advantages, its practical deployment depends on system optimization and infrastructure support.

Gesture-based interaction, in contrast, emphasizes natural and contactless control but exhibits lower reliability in real-world environments. The performance of gesture recognition systems is highly dependent on environmental conditions and sensor accuracy, which can lead to inconsistent interaction outcomes. As a result,

although gesture-based systems enhance intuitiveness, their adoption remains limited compared to VR-based approaches. This indicates a gap between conceptual usability and practical robustness in gesture-based interaction systems. Haptic feedback systems provide significant advantages in precision and user perception by delivering tactile feedback during manipulation tasks. These systems are particularly effective in applications requiring fine control, such as teleoperation and medical robotics. However, their implementation is constrained by hardware complexity and cost, which limits their scalability in broader industrial applications. Furthermore, integrating haptic feedback with digital twin systems requires precise synchronization to ensure realistic interaction. Another important observation is the lack of standardized evaluation frameworks across studies. Different research works employ varying metrics and experimental setups, making it difficult to perform direct comparisons between interaction techniques. For instance, some studies focus on objective performance metrics such as task completion time, while others emphasize subjective usability measures, leading to inconsistent evaluation outcomes [58],[57]. This inconsistency is further influenced by differences in study design, participant selection, and application scenarios. The root causes of this standardisation gap can be attributed to disciplinary fragmentation between the robotics and human computer interaction communities, the relatively recent emergence of digital twin interaction as a research domain, and the absence of agreed benchmark tasks or shared evaluation environments that would enable reproducible cross study comparisons.

These findings suggest that the differences between interaction techniques are not solely due to technological capabilities but also influenced by methodological variations and application contexts. While VR-based systems dominate current research due to their maturity and usability advantages, gesture-based and haptic approaches offer complementary benefits that can enhance interaction when properly integrated. This has led to the emergence of hybrid interaction systems, which aim to combine multiple modalities to overcome individual limitations. The specific findings and emerging trends identified across the 77 reviewed studies are synthesized in the following section. The comparative analysis is summarized in Table 5, which provides a structured comparison of interaction techniques based on key performance and usability factors.

As shown in Table 5 VR-based interaction achieves the highest ratings for both usability and accuracy, supported by its immersive visualization and spatial control capabilities. However, its moderate robustness and high implementation complexity indicate that deployment challenges remain significant, particularly in resource-constrained industrial settings. Gesture-based interaction scores high on intuitive usability but low on robustness, confirming that its practical value is undermined by environmental sensitivity and recognition inconsistency. Haptic feedback demonstrates the highest robustness and accuracy among individual modalities but carries the greatest implementation burden, making scalability a persistent concern. Hybrid and multimodal systems offer the most comprehensive interaction profile, combining high usability, accuracy, and robustness, yet their very high complexity reinforces the need for simplified integration frameworks. These observations collectively confirm that no single modality fully satisfies all performance and usability requirements, motivating the growing interest in adaptive and hybrid digital twin interaction systems. The key findings derived from this comparative analysis are discussed in detail in the following section.

**Table 5.** Comparative Analysis of Interaction Techniques in Digital Twin-Based Robotic Systems

Interaction Type	Usability	Accuracy	Robustness	Implementation Complexity	Key Limitation	References
Virtual Reality (VR)	High	High	Moderate	High	Latency, user fatigue	[67],[33]
Gesture-Based	High (intuitive)	Moderate	Low	Moderate	Recognition accuracy, environment sensitivity	[68],[10]
Haptic Feedback	Moderate	High	High	Very High	Cost, integration complexity	[69],[65]
Hybrid / Multimodal	Very High	High	Moderate–High	Very High	System complexity	[67],[57]

### 4.3. Key Findings

The analysis of the selected studies reveals several key findings and emerging trends in user interaction techniques for digital twin-based robotic arm systems. One of the most prominent observations is the dominance of virtual reality (VR)-based interaction, which has become the preferred approach for teleoperation and immersive control. VR systems provide enhanced spatial awareness, intuitive interaction, and improved task performance, making them highly suitable for complex robotic manipulation tasks [63],[34]. In addition, usability studies indicate that immersive interfaces can significantly improve user experience and reduce cognitive workload when compared to conventional control methods [61]. However, despite the widespread adoption of VR-based interaction, several studies highlight notable limitations and trade-offs. For instance, while VR improves usability and immersion, it may also introduce challenges such as system latency, hardware constraints, and user fatigue during extended operation. These findings suggest that the dominance of VR may

not solely reflect its superiority but also the maturity and availability of supporting technologies. This indicates a potential research bias toward VR-based systems rather than a definitive conclusion about its overall effectiveness.

Another important trend is the increasing use of standardized evaluation metrics, particularly subjective measures such as the NASA Task Load Index (NASA-TLX), to assess user workload and interaction quality [60]. These metrics provide valuable insights into user experience; however, their application remains inconsistent across studies. Some works prioritize objective performance metrics, while others emphasize subjective usability measures, resulting in fragmented evaluation approaches. This inconsistency limits the ability to establish direct comparisons between different interaction techniques. Latency and synchronization have been identified as critical challenges in digital twin-based robotic systems, particularly in teleoperation scenarios. Delays between the physical system and its virtual representation can significantly affect control accuracy, responsiveness, and overall user experience. Studies report latency ranges of 50 to 200 milliseconds for local processing architectures and 100 to 400 milliseconds for cloud-based systems, with even small variations leading to noticeable degradation in performance, especially in precision tasks [59]. Network protocol selection also plays a critical role: WiFi-based systems typically introduce gesture recognition delays of 80 to 300 milliseconds, while 5G-enabled architectures reduce this to 20 to 60 milliseconds, representing a substantial improvement for real-time digital twin interaction [52],[54]. Haptic feedback systems are the most latency-sensitive, requiring sub-20 millisecond response times to maintain stable force feedback loops, which severely constrains their viability in cloud-dependent architectures [59]. This issue becomes more pronounced in distributed systems, highlighting the urgent need for latency-aware and predictive control strategies, particularly as digital twin deployments scale across geographically distributed industrial environments.

In contrast to VR-based approaches, gesture-based interaction systems demonstrate strong potential for intuitive and natural control but are limited by reliability and robustness issues. Factors such as environmental variability and sensor accuracy can negatively impact gesture recognition performance, leading to inconsistent interaction outcomes [68]. Similarly, haptic feedback systems enhance user perception and control precision by providing tactile information; however, their adoption remains constrained by hardware complexity and integration challenges [70]. Notably, some contradictory findings exist in the literature: certain studies report that gesture-based interfaces achieve higher user satisfaction scores than VR in short-duration or collaborative tasks where physical immersion is less critical, while others find that traditional joystick control outperforms VR in high-precision tasks under constrained latency conditions [58],[57]. These contradictions reinforce the importance of task context and system conditions in evaluating interaction modality effectiveness and highlight the risk of drawing universal conclusions from studies conducted under varying experimental parameters. Overall, the findings indicate a clear shift toward human-centered and immersive interaction techniques, with increasing emphasis on usability, real-time responsiveness, and system adaptability. However, the coexistence of multiple interaction approaches, each with distinct strengths and limitations, suggests that no single technique fully satisfies all requirements. This has led to growing interest in hybrid and multimodal interaction systems, which aim to combine the advantages of different approaches to achieve more effective and robust human-robot interaction. The critical research gaps and ongoing challenges underlying these findings are examined in detail in the following section.

#### 4.4. Research Gaps and Challenges

The analysis of existing studies reveals several critical research gaps in digital twin-based robotic arm interaction systems. Although significant progress has been made in developing intuitive and immersive interaction techniques, limitations in system design, evaluation methodologies, and real-time performance continue to hinder practical implementation. These gaps persist for interconnected reasons: the multidisciplinary nature of digital twin interaction research slows the convergence of standards across robotics and HCI communities; the hardware requirements for advanced modalities such as haptic and multimodal systems create resource barriers for many research groups; and the absence of shared benchmark datasets or standardized test environments prevents reproducible comparison across studies. Table 6 shows a summary of the key research gaps identified in the reviewed studies, along with their impacts and potential future directions. One of the most prominent gaps is the lack of standardized evaluation frameworks for assessing interaction techniques. As discussed in previous sections, studies employ a wide range of objective and subjective metrics, making it difficult to establish consistent benchmarks across different systems. This inconsistency limits the comparability of results and reduces the reliability of conclusions drawn from existing research [57]. Furthermore, human-centered evaluation remains underexplored, with many studies focusing primarily on technical performance rather than user experience and cognitive factors [71]. Another significant gap lies in latency and synchronization challenges within digital twin environments. Although real-time interaction is a fundamental requirement, many systems struggle to maintain accurate synchronization between physical and

virtual components. Latency issues can negatively impact control accuracy, user perception, and overall system responsiveness, particularly in teleoperation scenarios [59]. Despite the recognition of this problem, limited research has focused on predictive or latency-aware approaches to mitigate these effects.

**Table 6.** Research Gaps in Digital Twin-Based Interaction Systems

Research Area	Identified Gap	Impact	Future Direction	References
Evaluation Framework	Lack of standardized metrics	Difficult comparison across studies	Develop unified evaluation framework [75]	[57],[71]
Latency & Synchronization	Delays in DT interaction	Reduced accuracy and responsiveness	Latency-aware and predictive models	[59]
Human-Centered Design	Limited user-focused evaluation	Poor usability understanding	Integrate usability and cognitive metrics	[71]
Haptic & Multimodal Systems	High complexity and cost	Limited real-world adoption	Simplified and scalable architectures	[70]
Explainability (AI)	Lack of system transparency	Reduced trust and interpretability	Apply explainable AI techniques	[73]

In addition, the integration of advanced interaction techniques, such as haptic feedback and multimodal systems, remains constrained by hardware complexity and system interoperability challenges. While haptic feedback enhances user perception and precision, its implementation often requires specialized devices and complex system integration, limiting its scalability in real-world applications [70]. Similarly, multimodal interaction systems introduce additional complexity in coordination and synchronization, which has not been fully addressed in current research. Another emerging gap is the lack of explainability and transparency in intelligent robotic systems. With the increasing use of AI-driven components in digital twin environments, understanding system behavior and decision-making processes becomes critical for user trust and system reliability [72]. Explainable AI approaches have been proposed to address this issue; however, their application in digital twin-based robotic interaction remains limited [73][74]. Overall, these gaps collectively reveal a field that is technically advancing but methodologically fragmented. The coexistence of diverse interaction approaches without shared evaluation standards limits the ability to establish which techniques are most effective for specific task types, user groups, or deployment contexts. Addressing these gaps requires coordinated effort across the robotics, HCI, and digital twin communities to develop standardized evaluation methodologies, improve real-time synchronization performance, and integrate human-centered design principles throughout the system development lifecycle. The findings from this review directly inform the research directions and practical recommendations outlined in the conclusion.

Table 6 summarises the key research gaps identified across the 77 reviewed studies. The most critical gap is the absence of standardized evaluation frameworks, which prevents meaningful cross-study comparison and limits the generalizability of findings [57],[71]. Latency and synchronization remain unresolved technical challenges, particularly as systems scale to distributed and cloud-based architectures [59]. The limited adoption of haptic and multimodal systems reflects persistent barriers in hardware complexity and cost, while the emerging gap in explainability highlights the need to address operator trust as AI-driven components become more prevalent in digital twin environments [73],[70]. Collectively, these gaps define the primary directions for future research in human-centered digital twin interaction for robotic arms.

As shown in Table 7, VR-based interaction is the most consistently evaluated modality, with objective metrics such as task completion time and positional accuracy frequently reported alongside subjective measures including NASA-TLX and SUS. However, even within VR studies, the specific metrics used and experimental setups vary considerably, limiting direct cross-study comparison. Gesture-based systems show significant variation in latency depending on the underlying network protocol, with 5G-enabled architectures reporting substantially lower delays (20–60 ms) compared to WiFi-based systems (80–300 ms) [52],[54],[76]. Haptic feedback systems operate under the most stringent latency requirements, typically needing sub-20 ms response times to maintain stable force feedback loops, which constrains their deployment in cloud-based digital twin architectures [59],[77]. The absence of a unified evaluation framework across modalities particularly for hybrid and multimodal systems represents a critical gap that impedes the comparability of findings and the development of human-centered benchmarks for digital twin interaction research.

Overall, Table 7 highlights a clear disparity in evaluation maturity across modalities. While VR-based systems benefit from relatively well-established metrics, gesture-based, haptic, and hybrid approaches lack consistent evaluation standards, with latency requirements varying by an order of magnitude depending on the modality and deployment context [52],[54],[59]. This disparity reinforces the need for a unified human-centered evaluation framework that accommodates the full range of interaction techniques reviewed in this paper.

**Table 7.** Comparison of Evaluation Metrics Across Interaction Modalities

Modality	Objective Metrics	Subjective Metrics	Latency Range	Standardisation Level	Ref.
VR-Based Interaction	Task completion time, positional accuracy, trajectory smoothness	NASA-TLX, SUS, user satisfaction surveys	50–200 ms (local); 100–400 ms (cloud)	Moderate — commonly used but inconsistently applied	[63],[67],[61]
Gesture-Based Interfaces	Recognition accuracy, response time, false positive rate	SUS, ease-of-use ratings	80–300 ms (WiFi); 20–60 ms (5G)	Low — no agreed benchmark or gesture set	[10],[68]
Haptic Feedback	Force accuracy, contact stability, trajectory error	NASA-TLX, perceived realism ratings	1–20 ms (stable haptic loop)	Low — hardware-dependent; metrics vary widely	[65],[70]
Hybrid / Multimodal	Combined task completion time, accuracy across modalities	NASA-TLX, SUS, workload comparison	Varies by modality combination	Very Low — no unified framework exists	[67],[57]
Traditional (Joystick)	Task completion time, positional accuracy	SUS, subjective preference	≤10 ms (local, near real-time)	Moderate — established but rarely benchmarked against immersive systems	[57],[58]

## 5. CONCLUSIONS

This review makes a significant contribution to the growing field of human-centered digital twin interaction by providing the first cross-modal, structured synthesis of user interaction techniques for robotic arm systems published between 2020 and 2025. Through the systematic analysis of 77 peer-reviewed studies retrieved from IEEE Xplore, Scopus, and Web of Science, this paper classifies interaction modalities, critically evaluates evaluation methodologies, and identifies key research gaps in digital twin-based robotic interaction systems. The analysis identifies four primary interaction categories: virtual reality (VR)-based interaction, gesture-based interfaces, haptic feedback systems, and hybrid or multimodal approaches. VR-based interaction emerges as the most dominant modality, consistently reported across studies for its immersive visualization, intuitive spatial control, and effectiveness in teleoperation and human-robot collaboration tasks [63],[40]. However, it is important to acknowledge that this dominance may partly reflect the maturity and availability of VR development tools and research infrastructure rather than an inherent superiority over other modalities. Gesture-based interfaces offer natural and contactless control but remain constrained by recognition accuracy and environmental sensitivity [10], while haptic feedback systems enhance precision and realism yet face significant barriers in hardware complexity and cost [65]. Hybrid and multimodal approaches show strong potential for combining the strengths of individual techniques, though their system complexity and limited standardization remain open challenges [67].

A critical finding of this review is the significant inconsistency in evaluation methodologies across the reviewed studies. While objective metrics such as task completion time, positional accuracy, and system latency are widely used, their application varies considerably in terms of experimental design, task scenarios, and participant selection. Latency, identified as one of the most critical challenges in digital twin environments, ranges from 50 to 200 milliseconds for local processing architectures and 100 to 400 milliseconds for cloud-based systems, with haptic feedback requiring sub-20 millisecond response times to maintain stable interaction loops [59]. Subjective measures such as NASA-TLX and SUS are applied inconsistently, and many studies prioritize technical performance over human-centered usability assessment. The absence of a unified evaluation framework can be attributed to disciplinary fragmentation between the robotics and human-computer interaction communities, the emerging nature of digital twin interaction as a research domain, and the lack of agreed benchmark tasks or standardized test environments across studies [57],[71]. Beyond evaluation, this review identifies several persistent research gaps. Human-centered design remains underexplored, with many studies focusing on system architecture and technical performance while giving limited attention to operator workload, cognitive load, and long-term usability [71]. The integration of explainable AI within digital twin interaction systems is an emerging but underdeveloped area, critical for building operator trust and system transparency in high-stakes industrial and remote operation contexts [73]. These gaps persist due to a combination of technical constraints, resource-intensive hardware requirements for haptic and multimodal systems, and the absence of shared datasets or evaluation benchmarks that would enable reproducible cross-study comparisons. For system designers and industry practitioners, these findings underscore the importance of adopting human-in-the-loop design principles, investing in low-latency communication infrastructure such as 5G and edge computing, and building modular interaction architectures that can accommodate multiple modalities.

The theoretical contribution of this paper is threefold. First, it provides a structured taxonomy of user interaction techniques for robotic arms in digital twin environments, enabling systematic comparison across modalities. Second, it critically evaluates existing evaluation practices and highlights the specific dimensions metrics, experimental design, and participant selection where standardization is most urgently needed. Third, it synthesizes research gaps and frames them within a human-centered design perspective that bridges technical and usability considerations. The review is subject to certain limitations: the search was restricted to English-language publications across three databases, which may exclude relevant studies published in other languages or indexed elsewhere. Additionally, the inclusion criteria focused specifically on peer-reviewed studies with real-time digital twin synchronization, which, while ensuring rigor, may exclude valuable early-stage or exploratory works. Based on the identified gaps, three specific high-priority future research directions are proposed. First, the development of a standardized human-robot interaction (HRI) evaluation protocol specifically designed for digital twin environments, incorporating both objective performance metrics and human-centered usability measures, is urgently needed to enable meaningful cross-study benchmarking. Second, investigation into latency-aware and predictive synchronization strategies leveraging 5G connectivity and edge AI should be prioritized to close the gap between theoretical DT architectures and practical real-time deployment requirements. Third, the design and validation of hybrid multimodal interaction systems that integrate VR, gesture, and haptic feedback within a unified, explainable framework represents a compelling direction for advancing both the usability and reliability of digital twin-based robotic systems. As digital twin technology continues to mature, advancing human-centered interaction will be essential to realizing its full potential in intelligent manufacturing, remote operations, and collaborative robotics domains where the quality of the human system interface directly determines operational safety, efficiency, and adaptability.

## DECLARATION

### Author Contribution

All authors contributed equally to the main contributor to this paper. All authors read and approved the final paper.

### Funding

There are no publication fees for this journal.

### Acknowledgement

Thanks are extended to Fakulti Teknologi Maklumat Dan Komunikasi (FTMK), Universiti Teknikal Malaysia Melaka (UTeM), for their technical support and research resources. Appreciation is also given to the Centre for Research and Innovation Management (CRIM), UTeM, for funding this research through the PJP Perspektif 2024 grant. And gratitude also expresses to the Ministry of Higher Education, Malaysia (KPT), for supporting this study through the Fundamental Research Grant Scheme Exploratory Research Consortium (FRGS-EC) FRGS-EC/1/2024/ICT09/UTEM/03/2.

### Conflicts of Interest

The authors declare no conflict of interest.

## REFERENCES

- [1] D. Jones, C. Snider, A. Nassehi, J. Yon, and B. Hicks, "Characterising the digital twin: A systematic literature review," *CIRP Journal of Manufacturing Science and Technology*, vol. 29, pp. 36–52, 2020, <https://doi.org/10.1016/j.cirpj.2020.02.002>.
- [2] Y. Lu, C. Liu, I. K. Wang, H. Huang, and X. Xu, "Digital twin-driven smart manufacturing: Connotation, reference model, applications and research issues," *Robotics and Computer-Integrated Manufacturing*, vol. 61, p. 101837, 2020, <https://doi.org/10.1016/j.rcim.2019.101837>.
- [3] A. Mazumder, M. Sahed, Z. Tasneem, M. S. Kaiser, and M. Mahmud, "Towards next generation digital twin in robotics: Trends, scopes, challenges and future," *Heliyon*, vol. 9, no. 6, p. e16889, 2023, <https://doi.org/10.1016/j.heliyon.2023.e16889>.
- [4] A. P. Burghardt, J. Szybicki, P. Gierlak, K. Musial, P. Pietruś, R. Cygan, "Programming of Industrial Robots Using Virtual Reality and Digital Twins," *Applied Sciences*, vol. 10, no. 2, p. 486, 2020, <https://doi.org/10.3390/app10020486>.
- [5] W. Wu, J. Liu, H. Zhang, and Y. Li, "Research on guidance methods of digital twin robotic arms," *Sensors*, vol. 23, no. 7, p. 3521, 2023, <https://doi.org/10.3390/s23073521>.
- [6] D. Mourtzis, J. Angelopoulos, N. Panopoulos, "Smart Manufacturing and Tactile Internet Based on 5G in Industry 4.0: Challenges, Applications and New Trends," *Electronics*, vol. 10, no. 24, p. 3175, 2021, <https://doi.org/10.3390/electronics10243175>.

- [7] J. J. Lopez-Huanca, C. Flores-Uriazar, A. Ramos-Cabezas, "Augmented Reality for Human–Robot Interaction in Collaborative Industrial Environments," *Sensors*, vol. 22, no. 3, P. 808, 2022, <https://doi.org/10.3390/s22030808>.
- [8] V. Villani, F. Pini, F. Leali, and C. Secchi, "Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications," *Mechatronics*, vol. 55, pp. 248–266, 2020, <https://doi.org/10.1016/j.mechatronics.2018.02.009>.
- [9] K. Wan, Y. Wang, J. Zhang, and X. Li, "A virtual reality-based immersive teleoperation system for robotic manipulation," *Journal of Manufacturing Systems*, vol. 72, pp. 230–243, 2024, <https://doi.org/10.1016/j.jmsy.2023.12.004>.
- [10] A. Vysocký, M. Novák, and P. Bartoš, "Hand gesture interface for robot path definition in collaborative workspaces," *Sensors*, vol. 23, no. 5, p. 2417, 2023, <https://doi.org/10.3390/s23052417>.
- [11] C. Savur, *A Physiological Computing System to Improve Human-Robot Collaboration by Using Human Comfort Index*. Rochester Institute of Technology, 2022, <https://doi.org/10.3390/machines11050536>.
- [12] S. Hopko, J. Wang, and R. Mehta, "Human factors considerations and metrics in shared space human-robot collaboration: A systematic review," *Frontiers in Robotics and AI*, vol. 9, p. 799522, 2022, <https://doi.org/10.3389/frobt.2022.799522>.
- [13] E. Artetxe, O. Barambones, I. Calvo, P. Fernández-Bustamante, I. Martín, and J. Uralde, "Wireless technologies for industry 4.0 applications," *Energies*, vol. 16, no. 3, p. 1349, 2023, <https://doi.org/10.3390/en16031349>.
- [14] U. Asad, M. Khan, A. Khalid, and W. A. Lughmani, "Human-centric digital twins in industry: A comprehensive review of enabling technologies and implementation strategies," *Sensors*, vol. 23, no. 8, p. 3938, 2023, <https://doi.org/10.3390/s23083938>.
- [15] S. Mancin, M. Sguanci, D. Andreoli, F. Soekeland, G. Anastasi, M. Piredda, and M. G. De Marinis, "Systematic review of clinical practice guidelines and systematic reviews: a method for conducting comprehensive analysis," *MethodsX*, vol. 12, p. 102532, 2024, <https://doi.org/10.1016/j.mex.2023.102532>.
- [16] B. Kitchenham, D. Budgen, and O. P. Brereton, "Evidence-based software engineering and systematic reviews," *IEEE Software*, vol. 28, no. 1, pp. 28–32, 2011, <https://doi.org/10.1109/MS.2010.122>.
- [17] T. Arksey and L. O'Malley, "Scoping studies: Towards a methodological framework," *International Journal of Social Research Methodology*, vol. 8, no. 1, pp. 19–32, 2005, <https://doi.org/10.1080/1364557032000119616>.
- [18] P. Runeson and M. Höst, "Guidelines for conducting and reporting case study research in software engineering," *IEEE Transactions on Software Engineering*, vol. 35, no. 2, pp. 131–164, 2009, <https://doi.org/10.1109/TSE.2009.6>.
- [19] D. Page *et al.*, "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," *BMJ*, vol. 372, p. n71, 2021, <https://doi.org/10.1136/bmj.n71>.
- [20] M. L. Rethlefsen *et al.*, "PRISMA-S: An extension to the PRISMA statement for reporting literature searches in systematic reviews," *Journal of the Medical Library Association*, vol. 109, no. 2, pp. 174–200, 2021, <https://doi.org/10.5195/jmla.2021.962>.
- [21] M. Gusenbauer and N. R. Haddaway, "Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources," *PLOS ONE*, vol. 15, no. 7, p. e0238237, 2020, <https://doi.org/10.1371/journal.pone.0238237>.
- [22] H. Snyder, "Literature review as a research methodology: An overview and guidelines," *Journal of Business Research*, vol. 104, pp. 333–339, 2019, <https://doi.org/10.1016/j.jbusres.2019.07.039>.
- [23] A. Carrera-Rivera, W. Ochoa, F. Larrinaga, and G. Lasar, "How-to conduct a systematic literature review: A quick guide for computer science research," *MethodsX*, vol. 9, p. 101895, 2022, <https://doi.org/10.1016/j.mex.2022.101895>.
- [24] J. Wohlin *et al.*, "Guidelines for snowballing in systematic literature studies and a replication in software engineering," *Information and Software Technology*, vol. 64, pp. 1–12, 2019, <https://doi.org/10.1016/j.infsof.2015.03.006>.
- [25] D. Moher, A. Liberati, J. Tetzlaff, and D. G. Altman, "Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement," *PLoS Medicine*, vol. 6, no. 7, p. e1000097, 2009, <https://doi.org/10.1371/journal.pmed.1000097>.
- [26] S. Castillo and P. Grbovic, "The APISSE Methodology for Systematic Literature Reviews in Engineering," in *IEEE Access*, vol. 10, pp. 23700–23707, 2022, <https://doi.org/10.1109/ACCESS.2022.3148206>.
- [27] B. Kitchenham, D. Budgen, and O. P. Brereton, "Using mapping studies as the basis for further research – A participant-observer case study," *Information and Software Technology*, vol. 53, no. 6, pp. 638–651, 2011, <https://doi.org/10.1016/j.infsof.2010.12.011>.
- [28] K. Petersen, R. Feldt, S. Mujtaba, and M. Mattsson, "Systematic mapping studies in software engineering," *Proceedings of the 12th International Conference on Evaluation and Assessment in Software Engineering*, pp. 68–77, 2008, <https://doi.org/10.14236/ewic/EASE2008.8>.
- [29] S. A. Perrig, L. F. Aeschbach, N. Scharowski, N. von Felten, K. Opwis, and F. Brühlmann, "Measurement practices in user experience (UX) research: a systematic quantitative literature review," *Frontiers in Computer Science*, vol. 6, p. 1368860, 2024, <https://doi.org/10.3389/fcomp.2024.1368860>.
- [30] E. Negri, L. Fumagalli, and M. Macchi, "A review of the roles of digital twin in CPS-based production systems," *Procedia Manufacturing*, vol. 51, pp. 939–948, 2020, <https://doi.org/10.1016/j.promfg.2020.10.131>.
- [31] F. Tao and Q. Qi, "Make more digital twins," *Nature*, vol. 573, pp. 490–491, 2019, <https://doi.org/10.1038/d41586-019-02849-1>.

- [32] Q. Qi, F. Tao, T. Hu, N. Anwer, A. Liu, Y. Wei, L. Wang, A. Y. C. Nee, "Enabling technologies and tools for digital twin," *Journal of Manufacturing Systems*, vol. 58, pp. 3–21, 2021, <https://doi.org/10.1016/j.jmsy.2019.10.001>.
- [33] J. Liu, H. Zhang, and Y. Li, "Digital twin-driven robotic system for intelligent manufacturing," *IEEE Access*, vol. 9, pp. 135482–135493, 2021, <https://doi.org/10.1109/ACCESS.2021.3116114>.
- [34] C. Li, P. Zheng, S. Li, Y. Pang, and C. K. Lee, "AR-assisted digital twin-enabled robot collaborative manufacturing system with human-in-the-loop," *Robotics and Computer-Integrated Manufacturing*, 76, 102321, 2022, <https://doi.org/10.1016/j.rcim.2022.102321>.
- [35] S. Boschert and R. Rosen, "Digital twin—The simulation aspect," in *Mechatronic Futures*, pp. 59–74, 2021, [https://doi.org/10.1007/978-3-030-65726-0\\_5](https://doi.org/10.1007/978-3-030-65726-0_5).
- [36] R. Fuller, L. Fan, C. Day, C. Barlow, "Digital Twin: Enabling Technologies, Challenges and Open Research," *IEEE Access*, vol. 8, pp. 108952–108971, 2020, <https://doi.org/10.1109/ACCESS.2020.2998358>.
- [37] A. Rasheed, O. San, T. Kvamsdal, "Digital Twin: Values, Challenges and Enablers From a Modeling Perspective," *IEEE Access*, vol. 8, pp. 21980–22012, 2020, <https://doi.org/10.1109/ACCESS.2020.2970143>.
- [38] T. Klauber, S. Zander, J. Straub, "Virtual Reality as an Interface for Human–Robot Collaboration in Precision Manufacturing," *Robotics*, vol. 11, no. 5, P. 94, 2022, <https://doi.org/10.3390/robotics11050094>.
- [39] S. Kim, J. Lee, and K. Kim, "Immersive virtual reality-based teleoperation of a robotic manipulator," *IEEE Access*, vol. 8, pp. 174019–174030, 2020, <https://doi.org/10.1109/ACCESS.2020.3025778>.
- [40] B. R. Galarza, P. Ayala, S. Manzano, and M. V. Garcia, "Virtual reality teleoperation system for mobile robot manipulation," *Robotics*, vol. 12, no. 6, p. 163, 2023, <https://doi.org/10.3390/robotics12060163>.
- [41] P. Suárez-Fernández, J. Albo-Canals, C. Martell, "Interaction Design Principles for Teleoperated Robotic Systems in Industrial Applications," *Applied Sciences*, vol. 12, no. 7, p. 3502, 2022, <https://doi.org/10.3390/app12073502>.
- [42] A. Ajoudani *et al.*, "Progress and prospects of the human–robot collaboration," *Autonomous Robots*, vol. 44, pp. 957–975, 2020, <https://doi.org/10.1007/s10514-019-09914-1>.
- [43] J. Rosen, M. Brand, M. B. Fuchs, and M. Arcan, "A myosignal-based powered exoskeleton system," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 31, no. 3, pp. 210–222, 2020, <https://doi.org/10.1109/3468.925661>.
- [44] H. Li *et al.*, "Towards the Design and Optimization of a Local Teleoperation Cockpit for Customs Remote Inspection," *2025 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pp. 2105–2110, 2025, <https://doi.org/10.1109/ROBIO66223.2025.11377237>.
- [45] J. L. Raheja, K. Das, A. Chaudhary, "Gesture Control of Industrial Robots Using Convolutional Neural Networks," *Computers and Electrical Engineering*, vol. 91, P. 107043, 2021, <https://doi.org/10.1016/j.compeleceng.2021.107043>.
- [46] A. Fang, H. Zhong, P. Zhao, "A Real-Time Hand Gesture Recognition and Human–Robot Interaction System," *Sensors*, vol. 21, no. 22, p. 7556, 2021, <https://doi.org/10.3390/s21227556>.
- [47] V. F. de Oliveira, G. Matioli, C. J. B. Júnior, R. Gaspar and R. G. Lins, "Digital Twin and Cyber-Physical System Integration in Commercial Vehicles: Latest Concepts, Challenges and Opportunities," in *IEEE Transactions on Intelligent Vehicles*, vol. 9, no. 4, pp. 4804–4819, 2024, <https://doi.org/10.1109/TIV.2024.3378579>.
- [48] M. Peral, A. Sanfeliu and A. Garrell, "Efficient Hand Gesture Recognition for Human–Robot Interaction," in *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 10272–10279, 2022, <https://doi.org/10.1109/LRA.2022.3193251>.
- [49] H. Lee, S. D. Kim, and M. A. U. Al Amin, "Control framework for collaborative robot using imitation learning-based teleoperation from human digital twin to robot digital twin," *Mechatronics*, vol. 85, p. 102833, 2022, <https://doi.org/10.1016/j.mechatronics.2022.102833>.
- [50] D. -N. Song, W. -C. Tang, Y. -N. Zhao, Y. -G. Zhong and J. -W. Ma, "Convolution-Based Velocity-Smoothing Principle and Its Application to Real-Time Parametric Curve Interpolation," in *IEEE Transactions on Automation Science and Engineering*, vol. 22, pp. 23443–23454, 2025, <https://doi.org/10.1109/TASE.2025.3625244>.
- [51] K. Foteinos, M. Linardakis, P. Radoglou-Grammatikis, V. Argyriou, P. Sarigiannidis, I. Varlamis, and G. T. Papadopoulos, "Visual hand gesture recognition with deep learning: A comprehensive review of methods, datasets, challenges and future research directions," *arXiv preprint arXiv:2507.04465*, 2025, <https://doi.org/10.48550/arXiv.2507.04465>.
- [52] A. Khalil and J. Kwon, "Nonlinear Performance Degradation of Vision-Based Teleoperation under Network Latency," *arXiv preprint arXiv:2603.06850*, 2026, <https://doi.org/10.48550/arXiv.2603.06850>.
- [53] M. Zhu, S. He, T. Chen, and C. Lee, "Wearable Intelligent Human–Machine Interfaces Ready for Sustainable Edge Computing Systems," *AI Sensors*, vol. 1, no. 2, p. 9, 2025, <https://doi.org/10.3390/aisens1020009>.
- [54] R. E. Kondo *et al.*, "An industrial edge computing architecture for Local Digital Twin," *Computers & Industrial Engineering*, vol. 193, p. 110257, 2024, <https://doi.org/10.1016/j.cie.2024.110257>.
- [55] H. Culbertson, S. B. Schorr, A. M. Okamura, "Haptics: The Present and Future of Artificial Touch Sensation," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 1, pp. 385–409, 2018, <https://doi.org/10.1146/annurev-control-060117-105043>.
- [56] M. S. Nawaz, S. U. R. Khan, S. Hussain, and J. Iqbal, "A study on application programming interface recommendation: state-of-the-art techniques, challenges and future directions," *Library Hi Tech*, vol. 41, no. 2, pp. 355–385, 2023, <https://doi.org/10.1108/LHT-02-2022-0103>.
- [57] A. Baratta, A. Cimino, F. Longo, and L. Nicoletti, "Digital twin for human-robot collaboration enhancement in manufacturing systems: Literature review and direction for future developments," *Computers & Industrial Engineering*, vol. 187, p. 109764, 2024, <https://doi.org/10.1016/j.cie.2023.109764>.

- [58] N. Mizuno, Y. Tazaki, T. Hashimoto, and Y. Yokokohji, "A comparative study of manipulator teleoperation methods for debris retrieval phase in nuclear power plant decommissioning," *Advanced Robotics*, vol. 37, no. 9, pp. 541-559, 2023, <https://doi.org/10.1080/01691864.2023.2169588>.
- [59] M. T. B. Touhid, E. Zhu, M. V. Ehteshamfara, and S. Yang, "Evaluation of digital twin synchronization in robotic assembly using YOLOv8," *The International Journal of Advanced Manufacturing Technology*, vol. 134, no. 1, pp. 871-885, 2024, <https://doi.org/10.1007/s00170-024-14182-7>.
- [60] S. Hart, "NASA Task Load Index (NASA-TLX): 20 years later," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 64, no. 1, pp. 904-908, 2020, <https://doi.org/10.1177/1071181320641236>.
- [61] Y. Kim, H. Park, and J. Lee, "Usability evaluation of immersive teleoperation interfaces for robotic manipulation," *IEEE Access*, vol. 9, pp. 147920-147931, 2021, <https://doi.org/10.1109/ACCESS.2021.3123864>.
- [62] S. Gholami, M. Lorenzini, E. De Momi and A. Ajoudani, "Quantitative Physical Ergonomics Assessment of Teleoperation Interfaces," in *IEEE Transactions on Human-Machine Systems*, vol. 52, no. 2, pp. 169-180, April 2022, <https://doi.org/10.1109/THMS.2022.3149167>.
- [63] W. Fan *et al.*, "Digital Twin-Driven Mixed Reality Framework for Immersive Teleoperation With Haptic Rendering," in *IEEE Robotics and Automation Letters*, vol. 8, no. 12, pp. 8494-8501, 2023, <https://doi.org/10.1109/LRA.2023.3325784>.
- [64] F. Chinello, M. Malvezzi, C. Pacchierotti, D. Prattichizzo, "Design and Characterization of a Kinesthetic Wearable Haptic Interface for the Hand," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 3293-3300, 2020, <https://doi.org/10.1109/LRA.2020.2975705>.
- [65] R. V. Patel, S. F. Atashzar and M. Tavakoli, "Haptic Feedback and Force-Based Teleoperation in Surgical Robotics," in *Proceedings of the IEEE*, vol. 110, no. 7, pp. 1012-1027, 2022, <https://doi.org/10.1109/JPROC.2022.3180052>.
- [66] M. Sarac, M. Solazzi, A. Frisoli, "Design Requirements of Generic Hand Exoskeletons and Survey of Hand Exoskeletons for Rehabilitation, Assistive, or Haptic Use," *IEEE Transactions on Haptics*, vol. 12, no. 4, pp. 400-413, 2019, <https://doi.org/10.1109/TOH.2019.2924881>.
- [67] M. Walker, T. Phung, T. Chakraborti, T. Williams, and D. Szafir, "Virtual, augmented, and mixed reality for human-robot interaction: A survey and virtual design element taxonomy," *ACM Transactions on Human-Robot Interaction*, vol. 12, no. 4, pp. 1-39, 2023, <https://doi.org/10.1145/3597623>.
- [68] B. Liang, N. Li, G. Li, J. Huang, Z. Yu and X. Zhao, "Sensing Technologies for Hand Gesture Recognition in Human-Robot Interaction: A Review," in *IEEE Sensors Journal*, vol. 26, no. 2, pp. 1501-1519, 2026, <https://doi.org/10.1109/JSEN.2025.3635622>.
- [69] Q. Ouyang *et al.*, "Bio-Inspired Haptic Feedback for Artificial Palpation in Robotic Surgery," in *IEEE Transactions on Biomedical Engineering*, vol. 68, no. 10, pp. 3184-3193, 2021, <https://doi.org/10.1109/TBME.2021.3076094>.
- [70] J. H. Bong, S. Choi, J. Hong, and S. Park, "Force feedback haptic interface for bilateral teleoperation of robot manipulation," *Microsystem Technologies*, vol. 28, no. 10, pp. 2381-2392, 2022, <https://doi.org/10.1007/s00542-022-05382-w>.
- [71] J. Endsley, "Human factors in automation: A review of issues and challenges," *IEEE Transactions on Human-Machine Systems*, vol. 50, no. 6, pp. 521-531, 2020, <https://doi.org/10.1109/THMS.2020.3001651>.
- [72] M. Charalambous, S. Fletcher, P. Webb, "The Development of a Scale to Evaluate Trust in Industrial Human-Robot Collaboration," *International Journal of Social Robotics*, vol. 8, no. 2, pp. 193-209, 2016, <https://doi.org/10.1007/s12369-015-0333-8>.
- [73] A. Chaddad, J. Peng, J. Xu, and A. Bouridane, "Survey of explainable AI techniques in healthcare," *Sensors*, vol. 23, no. 2, p. 634, 2023, <https://doi.org/10.3390/s23020634>.
- [74] V. Kumbhar, G. Sireesha, V. D. Jadhav, A. Barve, R. Sharma and M. E. M. Soudagar, "Integrating Explainable AI with Human-in-the-Loop Systems for Transparent Decision-Making in Autonomous Robots," *2025 International Conference on Intelligent Communication Networks and Computational Techniques (ICICNCT)*, pp. 1-6, 2025, <https://doi.org/10.1109/ICICNCT66124.2025.11232580>.
- [75] A. K. Ramasubramanian, R. Mathew, M. Kelly, V. Hargaden, and N. Papakostas, "Digital twin for human-robot collaboration in manufacturing: Review and outlook," *Applied Sciences*, vol. 12, no. 10, p. 4811, 2022, <https://doi.org/10.3390/app12104811>.
- [76] M. Simonetto, A. Montanari, T. Rossi, M. Giordani, "5G Wireless Connectivity for Industrial Robots: Key Challenges and Open Issues," *IEEE Communications Magazine*, vol. 59, no. 2, pp. 28-34, 2021, <https://doi.org/10.1109/MCOM.001.2000598>.
- [77] A. W. Yu and A. Nayak, "The Internet of Humanoids: A Survey of Technologies, Applications, and Challenges," in *IEEE Internet of Things Journal*, vol. 13, no. 6, pp. 10498-10521, 2026, <https://doi.org/10.1109/JIOT.2026.3653698>.

## AUTHOR BIOGRAPHY

**Aiman Hakim Bin Azahari** received his matriculation qualification in Information Technology from Selangor Matriculation College. He obtained his bachelor's degree in information technology (Game Technology) from Universiti Teknikal Malaysia Melaka (UTeM). He is currently pursuing a master's degree by research at Universiti Teknikal Malaysia Melaka. His research interests include human-computer interaction (HCI), simulation, and virtual reality (VR), with a particular focus on interactive systems and user-centred technologies.

---

Email: [m032410019@student.utm.edu.my](mailto:m032410019@student.utm.edu.my)

Researcher Website: None

**Mohd Khalid Bin Mokhtar** received his Diploma in Computer Science, majoring in Multimedia, from Universiti Teknologi Malaysia (UTM) in 2006. He continued his studies at the same university and obtained a bachelor's degree in computer science, specializing in Computer Graphics and Multimedia, in 2008. He later pursued a master's degree by research, which he completed in 2017. His main areas of expertise include computer graphics, visualization, augmented reality (AR), virtual reality (VR), and simulation. Currently, his research focuses on human-computer interaction (HCI).

Email: <mailto:khalid.mokhtar@utm.edu.my>

Researcher Website: Mohd Khalid Mokhtar - Scopus

**Nazreen Abdullasim** received his higher education in the field of computer science and is currently affiliated with Universiti Teknikal Malaysia Melaka (UTeM). His academic and research background focuses on computer graphics-related domains, including crowd simulation, collision detection, virtual reality (VR), augmented reality (AR), and interactive game-based systems. His main areas of expertise include simulation and immersive technologies. Currently, his research focuses on human computer interaction (HCI) and realistic virtual environment development.

Email: [nazreen.abdullasim@utm.edu.my](mailto:nazreen.abdullasim@utm.edu.my)

Researcher Website: Nazreen Abdullasim - Google Scholar

**Mohd Hafiz Bin Zakaria** received his academic training in the field of information and communication technology and is currently a faculty member at the Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka (UTeM). His scholarly work encompasses topics in computer science and interactive technologies, with contributions in areas such as simulation, virtual environments, and human-computer interaction through his publications and research activities. His research interests include advancing computational methods and immersive systems. He is actively engaged in research that bridges theoretical computing frameworks with practical applications in digital interaction and simulation technologies.

Email: [hafiz@utm.edu.my](mailto:hafiz@utm.edu.my)

Researcher Website: Mohd Hafiz Zakaria - Google Scholar

**Asniyani Nur Haidar Binti Abdullah** received her diploma in Computer Science (Multimedia) as a fast-track student from Universiti Teknologi Malaysia, Kuala Lumpur, in 2011. She then earned her B.Sc. in Computer Science (Graphic and Multimedia Software) and her Master of Philosophy (Computer Science), both from Universiti Teknologi Malaysia, Johor Bahru, in 2015 and 2017, respectively. Her research interests include graphic, machine learning, virtual reality, deep learning, computer vision, and data science.

Email: [asniyani@utm.edu.my](mailto:asniyani@utm.edu.my)

Researcher Website: Asniyani Nur Haidar Abdullah - Google Scholar

**Shafina Binti Abd Karim Ishigaki** received her bachelor's degree from Universiti Teknologi Malaysia (UTM) in 2021. She later pursued her master's degree at the same university, which she completed in 2023. She is currently a Lecturer at the Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka (UTeM). Her main areas of expertise include human-computer interaction (HCI), multimedia systems, game engine development, virtual reality (VR), augmented reality (AR), and simulation. Her research interests focus on immersive technologies, game-based learning, and interactive virtual environments.

Email: [shafina@utm.edu.my](mailto:shafina@utm.edu.my)

Researcher Website: Google Scholar – Shafina Binti Abd Karim Ishigaki

**Ikmal Faiq Albakri Mustafa Albakri** obtained his Diploma in Computer Science (Multimedia) through a fast-track programme at Universiti Teknologi Malaysia, Kuala Lumpur, in 2013. He subsequently obtained a Bachelor of Science in Computer Science (Graphic and Multimedia Software) in 2017, followed by a Master of Philosophy in Computer Science in 2021, both from Universiti Teknologi Malaysia, Johor Bahru. His research interests lie in the areas of 3D animation, extended reality, virtual production, and human-computer interaction (HCI).

Email: [ikmalfaiq@utm.edu.my](mailto:ikmalfaiq@utm.edu.my)

Researcher Website: Ikmal Faiq Albakri Mustafa Albakri - Google Scholar

**Muhamad Najib Zamri** is an Assistant Professor for Computer Science Programme at the University of Southampton Malaysia. Dr. Najib has over 10 years of experience in teaching computer science-related fields and research work. He actively engaged in research activities and has published in reputable journals, conference proceedings, and other types of articles in his field. His area of expertise is computer graphics, specifically real-time 3D visualisation and navigation.

Email: [m.n.zamri@soton.ac.uk](mailto:m.n.zamri@soton.ac.uk)

Researcher Website: Muhamad Najib Zamri - Google Scholar