Integrating Augmented Reality and Artificial Intelligence in Assembly Tasks: A Review of Strategies, Tools, and Challenges

Ana Ester Garcia de Paiva Pinheiro
Department of Architecture and Construction, Universidade Estadual de Campinas (UNICAMP), Rua Saturnino de Brito, 224, Cidade Universitária Zeferino Vaz. Campinas, São Paulo, CEP 13083-889, Brazil
ORCID: 0009-0004-8352-0577
E-mail address: a188159@dac.unicamp.br

Ana Regina Mizrahy Cuperschmid
Department of Architecture and Construction, Universidade Estadual de Campinas (UNICAMP), Rua Saturnino de Brito, 224, Cidade Universitária Zeferino Vaz. Campinas, São Paulo, CEP 13083-889, Brazil
ORCID: 0000-0002-6792-174X
E-mail address: cuper@unicamp.br

Abstract
The integration of Augmented Reality (AR) and Artificial Intelligence (AI) is a growing subject in the technological field, especially when applied to benefit assembly tasks. This paper presents a Systematic Literature Review to explore the benefits, challenges, methods and tools in the utilization of AR and AI to assembly tasks applications. The study selected 27 relevant publications from the period between 2019 and January 2025 to identify strategies reported in the literature to implement AR and AI to assembly processes. The results show that the integration of these technologies was used mainly in the sectors of industry and manufacturing and the AI was employed mostly to object detection through deep learning models. This review highlights the possibility of utilization of integrating AR and AI for a diversity of fields and the necessity to implement automatic real time fault detection for error minimization and productivity enhancement.
Keywords: Extended Reality; Machine Learning; Mounting; Mixed Reality; Industry 4.0

Introduction

Despite technological advances throughout the time, such as machinery and digital manuals, instead of printed ones, the human force is still needed in many factories. Although necessary, human involvement in the process of assembly strongly influences several factors such as feasibility, working comfort, financial costs, operation safety and product quality (Santhi et al., 2015).

The incidence of human errors is mostly related to workplace conditions, environment and equipment (Saptari; Jia Xin Leau; Mohamad, 2015). The human limitations, such as cognitive load and fatigue, can impact performance, demanding technological assistance to improve efficiency, reduce mental strain and minimize errors (Stork; Schubö, 2010). Considering the increase of complexity in assembly processes, rise of decision-making challenges and product quality problem possibility, it is necessary to implement measures to optimize the assembly process such as immersive technologies that innovate assemblies and mitigate the mentioned risks (Panagou et al., 2023). This article provides a comprehensive Systematic Literature Review (SLR) and a discussion on adopted strategies reported in the literature for the integration of Augmented Reality (AR) and Artificial Intelligence (AI) applied to assembly tasks. For this purpose, the key questions that motivated this research were: (i) What was the main purpose of using AI? (ii) Which AI models were used the most? (iii) Which software and hardware were most used? (iv) How can AR and AI solutions be adapted to different assembly scenarios? (v) What were the common challenges and limitations in implementing AR and AI in assembly processes? (vi) What were the recommendations for future research or development in the application of AR and AI in assembly processes? This paper addresses techniques developed for this niche, presents recent trends and advancements in the field, and provides a comprehensive overview of contributions already made by the academic community. Further exploration into this subject can benefit AR projects aiming to adopt AI for assembly processes, given the potential that this association must facilitate these processes.

Augmented Reality integrated with Artificial Intelligence

The concept of AI can be defined as a subfield of computer science focused in enabling computers for human-performed tasks like interpretation, learning, knowledge representation, problem-solving, among others (Zhang et al., 2021). With the advancement of AI, more specialized approaches have emerged, such as machine learning, a subset of AI (Pradhan; Dinesh Kumar, 2019), and deep learning, a type of machine learning (Sasikala et al., 2021). Machine learning carries out knowledge acquisition through experience and improves its performance throughout

Commented [A1]: Mais um

the process (Sasikala et al., 2021). Within machine learning, deep learning stands out, as a branch of machine learning, assists computers learn from extensive amounts of information for tasks like image and speech recognition and object detection (Soori; Arezoo; Dastres, 2023).

AR can benefit from AI by implementing data interpretation (Uma, 2019), real-time feedback (Naqvi et al., 2024), object, voice and gesture recognition, environmental understanding, input data delivery (Ştefan et al., 2023), personalization and major interactivity (Lampropoulos, 2025). Due to these attributes, this integration can be applied to a variety of specific sectors such as training and education for healthcare professionals (Battineni et al., 2024), pharmacy (Roosan, 2023) and general medical applications (Naqvi et al., 2024; Okeh, 2024).

The most frequent object detection techniques identified in this systematic review can be categorized into classical computer vision, ensemble learning and deep learning approaches. Among the computer vision methods, it is worth mentioning the Scale Invariant Feature Transform (SIFT) as a technique that extracts distinctive features from images, making them invariant to scale changes and robust for object recognition (Lowe, 2004). ORB, an alternative to SIFT, is a binary descriptor that is rotation-invariant and highly efficient, making it suitable for real-time applications such as object detection on mobile devices (Rublee et al., 2011). The RANSAC algorithm is another technique used in object detection, particularly for geometric shape recognition, as it effectively eliminates outliers and fits models with minimal data (Jenkins; Goodwin; Talafha, 2024)

Ensemble learning is another method used to enhance object detection accuracy by combining multiple models integrating multiple base learners to improve performance capabilities. The main goal behind ensemble learning is recognizing that individual models have inherent limitations and can promote errors; thus, by leveraging multiple models, the approach achieves better classification performance and robustness (Mienye; Sun, 2022).

Deep learning-based object detection models, particularly those leveraging convolutional neural networks (CNNs), have transformed the field. CNNs are extensively used for processing and analyzing images, playing a vital role in object detection (Object Detection Using Deep Learning, CNNs and Vision Transformers: A Review). You Only Look Once (YOLO) is a real-time object detection framework derived from CNN (Carcellar; Tychuaco; Yumang, 2024) that trains using complete images and optimizes detection performance in a simple and fast way (Redmon et al., 2016). Region-based CNN (R-CNN) increases accuracy by applying a deep network to classify object proposals(Girshick, 2015). CenterNet proposes a low-cost solution that focuses on the geometric center of objects (Duan et al., 2019), while RetinaNet is a single-stage detector that uses deep CNNs to strengthen detection performance (Afif et al., 2020).

Modern advancements present Vision Transformer (ViT), a flexible, sequence-based approach, with big efficiency in handling a variety of input resolutions and ability to capture long-range dependencies and global context through self-attention mechanisms, being superior to the already consolidated CNNs (Elharrouss et al., 2025). Additionally, recurrent neural networks (RNNs) and variants, such as bidirectional long short-term memory (BiLSTM), have been investigated for object detection for its ability to model sequential dependencies, initially applied in text processing but with emerging uses in visual tasks (G. Liu & Guo, 2019). Finally, there are other advancements like deep neural networks (DNNs), known for their depth and width, have demonstrated strong performance across different datasets (Ciresan; Meier; Schmidhuber, 2012) and Spatio-Temporal Prompting Network (STPN), particularly useful in video-based object detection (Sun et al., 2023).

Methodology

The present SLR was carried out based on the PRISMA protocols (Tricco et al., 2018). The following databases were selected to perform the search, due to their importance within the research field: Engineering Village, Web of Science and Scopus. The search terms were segmented into three categories, Figure 1. To compose the search strings, the Boolean operators "AND" between categories and "OR" between terms were used (Fig. 1). The database search was based on journal articles that contained the selected strings within the title, abstract or keywords, in English. The selected timeframe was between January 2019 and January 2025, aiming to evaluate what is most up to date in this research area.

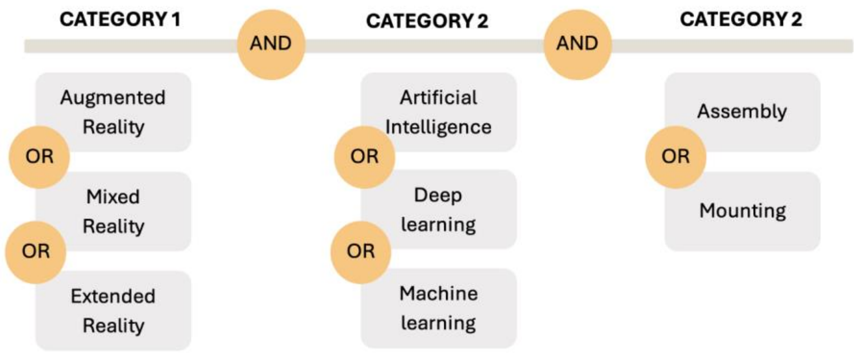

Figure 1. Selection of search strings

Inclusion and exclusion criteria

Inclusion and exclusion criteria were defined to help restrict the search to publications that addressed the subject in question. Among the inclusion criteria, the presentation of a case study dealing with AR associated with AI for assembly tasks and the descriptions of application development.

Exclusion criteria aimed to eliminate:

(i) literature review articles, to avoid duplicating or overlapping results;
(ii) off-topic articles;
(iii) articles that do not approach assembly tasks case studies, aiming specifically practical solutions;
(iv) articles that used human-robot collaboration, since including studies involving human-robot collaboration would introduce variables related to robotic assistance, which fall outside the scope of this research.

From the 192 publications obtained from the Scopus (41), Web of Science (86) and Engineering Village (65) databases, a final number of 27 publications was reached after applying the inclusion and exclusion criteria (Fig. 2). Initially, 88 duplicate articles were removed. Subsequently, it was observed that some articles lacked the search strings in the title, abstract, or keywords. This occurred because Web of Science include additional keywords in the automated search, such as *Keyword Plus*, which are generated by the Clarivate algorithm to expand search possibilities for articles (Clarivate Support, 2025). Through manual verification of the search strings, it was found that 52 articles did not meet the established criteria. It is worth noting that these articles, in addition to lacking the search strings, were found to be irrelevant to the topic upon reviewing their abstracts. All 27 resulting articles were successfully retrieved from the mentioned sources.

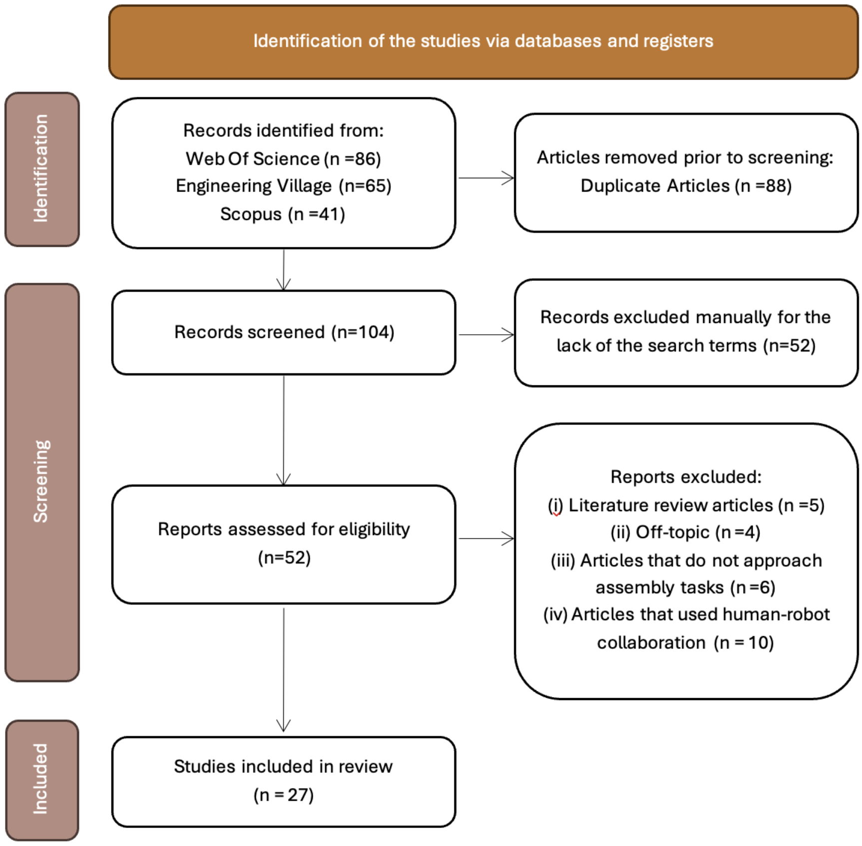

Figure 2. PRISMA flow diagram of this SLR

Furthermore, four articles were excluded as they did not align with the objectives of this study, focusing instead on topics from unrelated scientific domains, including neurology (Qin; Bulbul, 2023), livestock (Arıkan et al., 2023), labor law (Jain, 2019), and biomedicine (Mertz, 2023).

The articles which have already conducted a literature review were excluded, as the objective is to analyze the adopted strategies in articles presenting a case study using AR and AI applied to assembly tasks. The focus of this SLR is the identification, classification and of tools and methods used in the implementation of AR and AI in assembly tasks, in addition to providing an update of the period considered. It is pertinent to mention the generality of the assembly tasks addressed in this work, in contrast to the specific sectors such as construction (Hajirasouli et al., 2022), manufacture (Sahu; Young; Rai, 2021; Szajna; Kostrzewski, 2022; Lăzăroiu et al., 2024) and industry (Devagiri et al., 2022). These articles were excluded from this analysis also to avoid duplicating or overlapping results.

Six articles did not approach assembly tasks in the study (Azzam et al., 2023; Du; Kong; Zhao, 2023; Liu; Zhang; Gowda, 2023; Luo et al., 2024; Ma et al., 2024; Yue et al., 2024) and for this reason, they were excluded. Ten articles employed human-robot collaboration (Mueller et al., 2019; Wang et al., 2019; Dimitropoulos et al., 2021; ZHAO et al., 2021; Chu; Liu, 2023; Zhang et al., 2023a; Li et al., 2024a; Tolenov; Omarov, 2024; Xie et al., 2024; Zheng et al., 2024). They were excluded to ensure that the analysis remains focused on understanding the direct impact of AI and AR on human-performed tasks without confounding factors introduced by robotic interaction or automation.

Bibliometric Analysis
This section aims to analyze the bibliometric data related to the 27 resulting articles, summarized in Table 1.

Table 1. Bibliometric analysis of the studied articles.

RESEARCH	YEAR	JOURNAL	NATIONALITY
[69]	2022	INTERNATIONAL JOURNAL OF ADVANCED MANUFACTURING TECHNOLOGY	Italy
[70]	2023	INTERNATIONAL JOURNAL OF ADVANCED MANUFACTURING TECHNOLOGY	China
[71]	2020	SENSORS	South Korea
[72]	2024	BUILDINGS	China
[73]	2020	VIRTUAL REALITY AND INTELLIGENT HARDWARE	China
[74]	2021	IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT	China
[75]	2021	IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS	China
[76]	2019	SYMMETRY-BASEL	Slovakia
[77]	2023	ROBOTICS AND COMPUTER-INTEGRATED MANUFACTURING	China
[78]	2024	JOURNAL ON MULTIMODAL USER INTERFACES	India
[79]	2024	AUTOMATION IN CONSTRUCTION	Australia
[80]	2024	INTERNATIONAL JOURNAL OF HUMAN-COMPUTER INTERACTION	Switzerland
[81]	2021	JOURNAL OF AMBIENT INTELLIGENCE AND HUMANIZED COMPUTING	Singapore
[82]	2024	JOURNAL OF MANUFACTURING SYSTEMS	China
[83]	2023	ADVANCED ENGINEERING INFORMATICS	China
[84]	2022	APPLIED SCIENCES-BASEL	Italy
[85]	2024	ADVANCED ENGINEERING INFORMATICS	Taiwan
[86]	2023	IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS	China
(Filipescu et al., 2024)	2024	SENSORS	Romania
(Zhao et al., 2023)	2023	EXPERT SYSTEMS WITH APPLICATIONS	China
(Zhao et al., 2022)	2022	MEASUREMENT SCIENCE AND TECHNOLOGY	China
(Cramer et al., 2024)	2024	INTERNATIONAL JOURNAL OF COMPUTER ASSISTED RADIOLOGY AND SURGERY	Germany
(Lai et al., 2020)	2020	JOURNAL OF MANUFACTURING SYSTEMS	United States
(Aiken et al., 2024)	2024	IEEE ACCESS	Canada
(Geng et al., 2025)	2025	JOURNAL OF MANUFACTURING SYSTEMS	China
(Li; Chen, 2022)	2022	FRONTIERS IN PSYCHOLOGY	China
(Grappiolo et al., 2021b)	2021	INTERNATIONAL JOURNAL OF ADVANCED MANUFACTURING TECHNOLOGY	Netherlands

In the final selection of 27 articles considered in this study, the 14 studies were found in all the three used databases (Fig. 3).

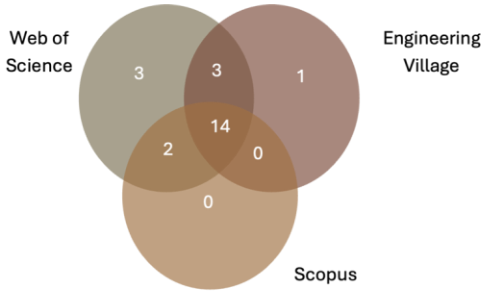

Figure 3. Articles found in each database

The selected articles were published between 2019 and September 2024, with a noticeable increase in publications from 2020 onwards (Fig. 4). Although 2020 represented a record number of publications on the topic, with a decline occurring afterwards, the numbers observed in subsequent years are still significant and constant, indicating academic interest in the topic.

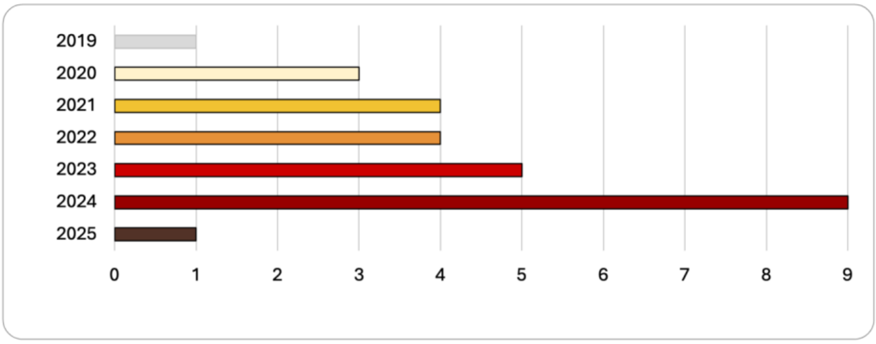

Figure 4. Publications per year.

The co-occurrence of keywords in the publications was analyzed using the VOSviewer tool. This analysis highlights the diverse range of approaches within the subject. This is illustrated by the scattered and weakly connected clusters, which are visualized in a word cloud. A total of 81 unique keywords were detected, considering author and index keywords.

To achieve a more cohesive grouping of concepts and a clearer visualization, a grouping of the keywords in the bibliography for a word cloud was made. This aimed to reduce unnecessary variations of terms related to the same concept, such as "artificial vision" and "machine vision" transforming into "computer vision". Due to many variations of assembly, the specific terms for example "cable assembly", "assembly inspection", "assembly assistant", among others, were summarized to "assembly". The same process was repeated to terms such as "augmented reality", "intelligent manufacturing", "neural networks" and "object detection". This process is useful to identify patterns and concentrate on key themes, preventing similar but slightly different terms from fragmenting the visualization. This process optimizes data representation and allows for a more straightforward and interpretable analysis (Table 2).

Table 2. Original keywords used and generalized keywords

Original keyword	Generalized keyword
AI AI surgery	artificial intelligence
Cable assembly	assembly
Assembly inspection	

Assembly assistant Assembly guidance	
Co-assembly	
Assembly elements	
Civil aircraft assembly	
Mechanical assembly	
Assembly Sequence Planning Disassembly	
AR	augmented reality
Wearable Augmented Reality Augmented Reality marker	
Artificial vision	computer vision
Machine vision	
Intelligent manufacturing workcell	Intelligent manufacturing
Recurrent neural networks	neural networks
Convolutional neural networks	
Region-based Convolutional Neural Networks (R-CNN)	
Redundant objects detection	object detection
Education 4.0/5.0	Industry 4.0

The network of connections was elaborated with the following parameters: the most frequent words are represented by larger circles and the number of connections is proportional to the line thickness (Fig. 5). The most recurrent keywords were deep learning, with 13 occurrences, followed by “augmented reality” and “assembly”, both with 11 occurrences.

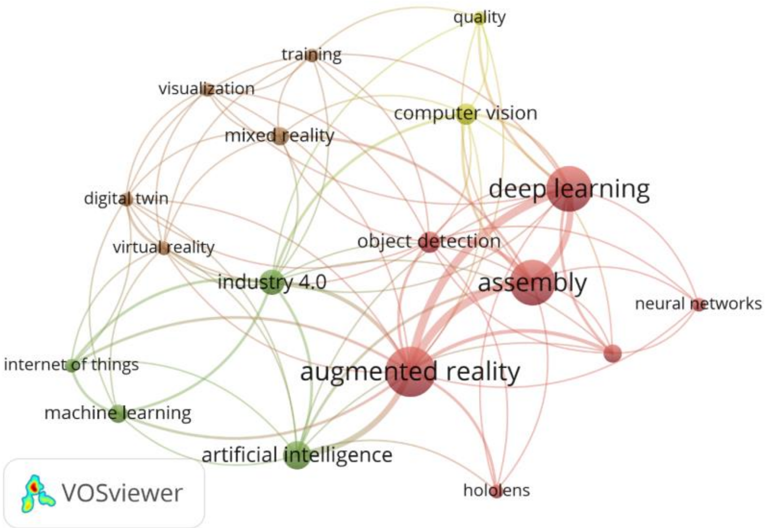

Figure 5. Keyword cloud.

The proceedings that published the most articles in the field were from International Journal of Advanced Manufacturing Technology and Journal of Manufacturing Systems with 3 publications each, while the IEEE Transactions on Industrial Informatics, Advanced Engineering Informatics and Sensors had two publications each. The remaining articles were published only once in different venues (Fig. 6).

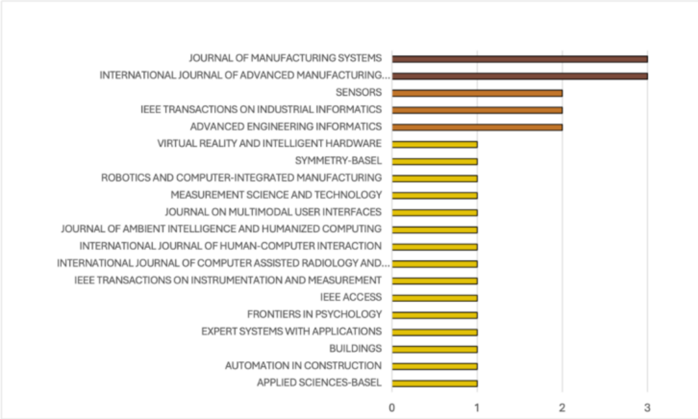

Figure 6. Publications per journal.

In the nationality segmentation analysis (Fig. 7), the research institution of the first author of each article was considered. It is evident that China accounts more than half of the selected content for this study, with 12 out of 27 articles. Following, Italy and Romania with two publications each. The analysis of publications per country serves as a geographic indication of locations where this has been more academically developed (Fig. 8).

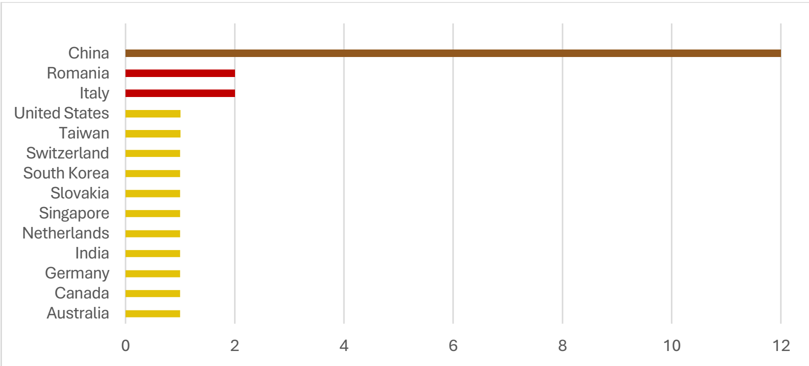

Figure 7. Publications per country.

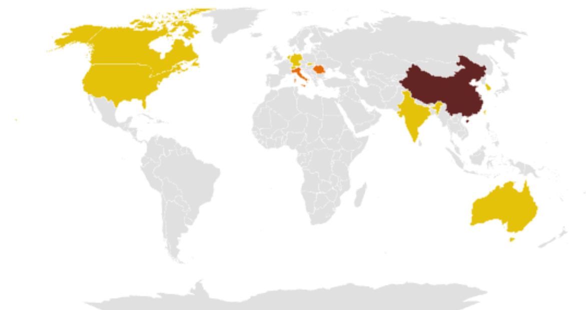

Figure 8. Locations of the publications reviewed.

Data classification
Table 3 summarizes the main analyzed topics, highlighting the key analyzed categories and their corresponding attributes. This approach enables a structured understanding of the data and facilitates further exploration of the patterns identified within the study.

Table 3. Data classification in each research.

RESEARCH	PURPOSE OF AI USAGE	IA MODEL	ASSEMBLY OBJECT	SOFTWARE	HARDWARE	ERROR DETECTION
(Generosi et al., 2022)	body tracking	CMU Deep Learning Model; Google Mediapipe Hand Tracking Model.	manufacturing work operations	-	iPhone	-
(Zhang et al., 2023b)	body tracking	YOLO	connectors of eletromechanical products	Unreal Engine;	-	✓
(Choi; Seo, 2020)	progress measurement	BAGGING METHOD	ship block	-		-
(Dzeng; Cheng; Cheng, 2024)	quality inspection	YOLO	scaffolding	Roboflow, Unity, Vuforia, MRTK	Hololens	✓
(Zheng et al., 2020)	object detection	CNN	aircraft cable	-	-	-
(Wang; Yan, 2021)	angle measurement; action recognition	YOLO; FASTER R-CNN; CNN	complex screws with a screwdriver	-	web cameras	✓
(Li; Zheng; Zheng, 2021)	object detection	RETINANET	aviation connectors	-	HiAR G200	-
(Zidek et al., 2019)	object recognition	R-CNN	standart parts (nuts, screws, washes)	Blender	Smartphone Samsung S7; Epson Moverio BT350	-
(Hu et al., 2023)	object recognition	MASK R-CNN	cable brackets	Vuforia	Iphone	-
(Raj et al., 2024)	object detection; body tracking; quality inspection	YOLO	pneumatic cylinder	Unity + MRTK + SolidWorks	Hololens	✓
(Chen et al., 2024)	vision-language understanding and generation	VIT	construction scenarios	Unity + C# Scripts;	Hololens	-
(Seeliger; Weibel; Feuerriegel, 2024)	visual navigation prediction	DNN	indoor environments	Unity + MRTK	Hololens	-
(Ho et al., 2021)	progress measurement	R-CNN	hybrid medical device	Unity; Vuforia	Iphone7; Logitech cameras; Epson Moverio	-
(Li et al., 2024b)	object detection; quality inspection	FCNs, YOLO	Non specified	Pytorch 1.7.1; OpenCV 4.5.1; Coin3D	camera monocular	-

RESEARCH	PURPOSE OF AI USAGE	IA MODEL	ASSEMBLY OBJECT	SOFTWARE	HARDWARE	ERROR DETECTION
(Wang et al., 2023)	body tracking	YOLO	aircraft component exhaust vent	Unity + MRTK	Hololens/HTC Vive	-
(De Feudis et al., 2022)	body tracking	OPEN POSE; YOLO	mechanical/industrial parts	SPSS Software	HTC Vive	-
(Chu; Chen; Chen, 2024)	object recognition	RANSAC	flange (mechanical part)	Apple ARKit	iPhone 12.	-
[86]	object detection	CENTERNET, CNN	reducer (mechanical part)	-	-	-
(Filipescu et al., 2024)		STPN	two workpieces, each consisting of five components (base, body, top and two cylinders)	SCADA and SIEMENS Tia Portal; MatLab; OpenCV; Node-RED; VNC Viewer	IFM Cameras 3D O3R222	-
(Zhao et al., 2023)	object detection	YOLO	screwdriver, cable piler, connector, etc	Unity, Visual Studio, MRTK, Open Neural Network Exchange	Hololens	-
(Zhao et al., 2022)	object detection	CENTERNET; YOLO	aircraft parts	-	-	-
(Cramer et al., 2024)	object recognition	-	surgical procedures	-	Microsoft HoloLens 2	-
(Lai et al., 2020)	object detection	FASTER R-CNN	tools, allen key, drill, screwdriver	Unity	Logitech camera	-
(Aiken et al., 2024)	object detection	YOLO; CNN	hydraulic fracturing valve	Blender-Proc2	Hololens; Matterport Pro	-
(Geng et al., 2025)	3d tracking, visual recognition, virtual-real matching, and rule reasoning are combined to complete, sequence, and locate the assembly targets	Deep Learning Recognition Model	complex electrical connectors	OpenIAI, using OpenCV, Unity, and Visual Studio	Custom made portable integrated desktop AR hardware named SNIO (Smart Nav-igator of Industrial Operations)	-
(Li; Chen, 2022)	object detection	Oriented FAST and Rotated BRIEF (ORB)	spindle	Vuforia	Hololens	-
(Grappiolo et al., 2021b)	progress measurement	RETINA NET	lumminaire	Freecad; Unity	Logitech camera	-

DISCUSSION
This section aims to focus on answering the questions established in the introduction based on the information acquired through the SLR.

What is the purpose of using AI?
Among the applications reported in the reviewed articles, it was possible to identify several uses for AI with a primary focus on object detection and recognition. Many studies implemented AI-driven object detection to identify components and parts during assembly processes (Lai et al., 2020; Zheng et al., 2020; Li; Zheng; Zheng, 2021; Zhao et al., 2022; Li et al., 2023; Aiken et al., 2024). Additionally, object recognition is employed to distinguish between different parts and tools, improving efficiency in assembly operations (Zidek et al., 2019; Hu et al., 2023; Chu; Chen; Chen, 2024; Cramer et al., 2024).

These applications allow AR systems to provide real-time guidance and error detection, ensuring accuracy and reducing the risk of assembly errors. The distinction between object detection and object recognition was identified. This separation is important in this analysis since the recognition of objects includes the classification of these objects based on their features and the detection is a task that consists of locating and identifying the objects within an image (Wang, 2016).

Beyond object detection, many articles focus on other uses for AI, such as tracking and measuring assembly progress through AI techniques. Body tracking is commonly used to observe assembler movements, enhancing ergonomics and task performance in assembly environments [69], [70], [78], [83], [84].

Progress measurement, which monitors the stages of assembly tasks, is also a crucial AI use (Choi; Seo, 2020; Grappiolo et al., 2021a; Ho et al., 2021). Additionally, AI-driven angle measurement and action recognition can give precise feedback on task execution, improving the accuracy and efficiency of manual assembly processes [74], ensuring that assembly tasks are performed in a correct and optimized way.

Some studies have explored AI applications in more functionalities, such as quality inspection, utilized in the identification of errors and ensure compliance with assembly standards [72], [78], [82]. Other studies also explored the role of AI in vision-language understanding for human-machine interaction [79] and visual navigation prediction for guiding assembly in complex environments [80]. Additionally, some researchers integrated AI-driven techniques like 3D tracking, virtual-real matching, rule reasoning, and target location to enhance precision in assembly tasks (Geng et al., 2025) (Fig. 9).

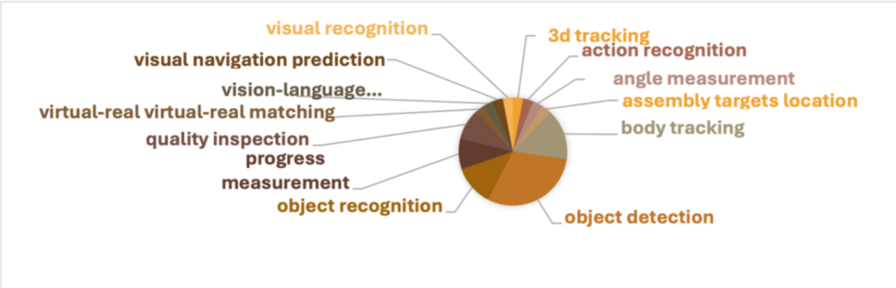

Figure 9. Purposes for AI usage.

Error analysis appears in many of the articles, often conducted by the researchers on the collected data to assess precision in the applications (Lai et al., 2020; Li et al., 2023, 2024b; Zhao et al., 2023). But one step ahead in the integration of AR and AI is the automation of the error detection, providing instant feedback for the assembler. This feature is observed in 4 of the 27 articles (Wang; Yan, 2021; Zhang et al., 2023b; Dzeng; Cheng; Cheng, 2024; Raj et al., 2024). They presented different approaches to automating error detection in assembly processes, leveraging advanced AI techniques and real-time monitoring systems.

(Zhang et al., 2023b) developed a two-stage assembly posture detection method that calculates matching errors for rotational and translational movements. Rotational errors are based on orientation discrepancies from rotation matrices, while translational errors use the intersection rate between bounding boxes estimated by the YOLO-6D network and virtual guidance models.

(Dzeng; Cheng; Cheng, 2024) trained a deep machine learning model with real images to automatically recognize deficiencies in scaffoldings, such as missing crosstie rods, lower-tie rods, and footboards. Their dataset included 4,240 images for training and over 1,000 for validation and testing, ensuring robust error detection.

(Wang; Yan, 2021) introduced a real-time angle-monitoring system for GPU assembly, designed to detect and warn operators of improper fastening angles with electric screwdrivers. The system prevents unacceptable quality errors by providing immediate feedback to the operator when deviations occur.

(Raj et al., 2024) implemented a fault detection algorithm integrated with HoloLens2, which uses hand-tracking capabilities to ensure all assembly steps are completed correctly based on the position of components, the user's thumb and index finger positions, and the distance between them, tailored to specific object sizes.

An analysis of the applications identified reveals diverse purposes for the implementation of AI in this context including object detection and recognition, body tracking, quality inspection, and fault detection. These automations optimize workflows by reducing errors, minimizing assembly time, and assisting users in the mistakes identification and recovery once they make a mistake. Collectively, these advantages contribute to the overall improvement of product quality.

Which AI models were used?
The analyzed studies utilized a variety of AI models, with object detection being the most prevalent, highlighting the CNN and its derivate YOLO. YOLO was majorly chosen due to its high-speed real-time detection capabilities for identifying assembly parts (Wang; Yan, 2021; De Feudis et al., 2022; Zhao et al., 2022, 2023; Wang et al., 2023; Zhang et al., 2023b; Aiken et al., 2024; Dzeng; Cheng; Cheng, 2024; Li et al., 2024b; Raj et al., 2024) Several studies implemented CNNs, a consolidated model for object detection and recognition, uniquely (Zheng et al., 2020) and integrated with other models as Faster R-CNN (Wang; Yan, 2021), CenterNet (Li et al., 2023) and YOLO (Aiken et al., 2024) for improved image processing.

More specialized deep learning architectures were utilized for distinct purposes, beyond object detection and recognition; for example, Fully Convolutional Networks (FCNs) are applied for semantic segmentation (Li et al., 2024b), and ViT are employed for vision-language understanding (Chen et al., 2024). These deep learning models benefit AR applications by providing more detail to the object recognition task and contextual understanding for assembly processes.

Variations of R-CNN and its derivatives were consistently utilized for object detection and segmentation tasks, such as Faster R-CNN, which improves detection speed (Lai et al., 2020; Wang; Yan, 2021), and Mask R-CNN, which adds instance segmentation capabilities (Hu et al., 2023). Other studies used RetinaNet, known for its ability to handle class imbalance through focal loss, for precise object detection (Grappiolo et al., 2021a; Li; Zheng; Zheng, 2021).

Beyond the deep learning architectures, a variety of works focused on body and pose tracking models, which are crucial for tracking worker movements and offering real-time feedback such as Google Mediapipe Hand Tracking and CMU Deep Learning Models were used for hand and body pose detection (Generosi et al., 2022). Similarly, OpenPose, which detects human body joints, is employed for body tracking in industrial assembly scenarios (De Feudis et al., 2022). Additionally, DNNs are applied for visual navigation and movement prediction in complex environments (Seeliger; Weibel; Feuerriegel, 2024), ensuring enhancement in worker interaction with AR interfaces and guaranteeing correct assembly techniques.

Other AI models emphasize geometric recognition and feature matching techniques such as RANSAC, applied for detecting geometric features (Chu; Chen; Chen, 2024) and ORB, applied for identifying objects from different angles (Li; Chen, 2022). STPN were also used for analysis of sequential assembly processes and progress tracking (Filipescu et al., 2024).

Apart from neural networks, ensemble learning was utilized by Choi & Seo (Choi; Seo, 2020) applying bagging and boosting methods to improve model accuracy and robustness. Moreover, a deep learning recognition model integrating multiple recognition techniques is employed for complex component identification by (Geng et al., 2025) in a custom implementation.

Considering the diverse AI techniques, a classification of the AI models utilized in the analyzed articles was developed (Fig. 10). This classification aims to enhance the understanding of the strategies employed in each study to integrate AI and AR for assembly tasks. Notably, YOLO emerged as the most frequently used model.

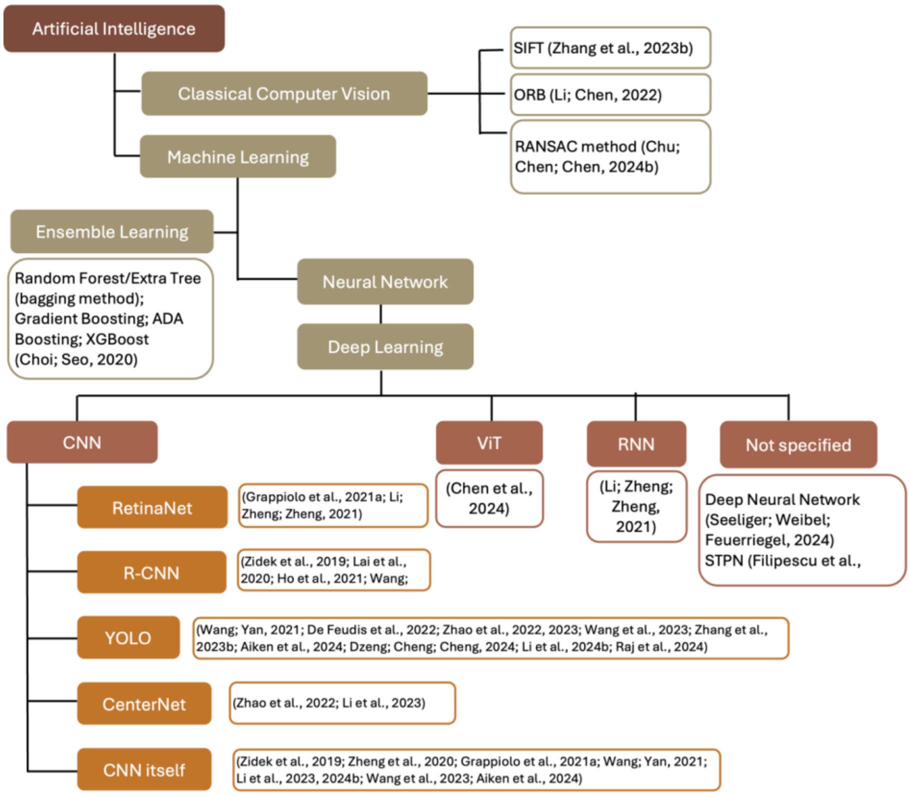

Figure 10. AI models identified in the studies.

AI models such as YOLO, CNNs, Mask R-CNN, ViT, and Faster R-CNN were employed in the studied assembly tasks, each catering to specific technical needs and specific purposes. YOLO is the most popular choice for real-time object detection, due to its speed and accuracy, making it ideal for identifying and locating assembly components. The CNNs were widely used for analyzing visual information, especially in object recognition and defect detection, as a model that covers not just the aforementioned YOLO but the other used models. Mask R-CNN and RetinaNet provided for the experimental studies in question precise segmentation and pixel-level accuracy, crucial for tasks requiring detailed component analysis.

It can be inferred that modern approaches like ViT are suitable for complex assembly scenarios, although they still do not represent an option that has been so widely implemented in studies over the last 6 years. Alternative methods such as OpenPose enable human pose estimation, enhancing safety and ergonomics during assembly but serves for more limited, serving specific analyses rather than comprehensive assembly task evaluation.

Which software and hardware were most used?

The reviewed studies utilized a range of software, with Unity being the most frequently used AR development engine, often combined with the Mixed Reality Toolkit (MRTK) for development of immersive AR environments (Wang et al., 2023; Zhao et al., 2023; Dzeng; Cheng; Cheng, 2024; Raj et al., 2024; Seeliger; Weibel; Feuerriegel, 2024). Many studies also employed Vuforia, a well-known AR Software Development Kit (SDK) associated to the aforementioned Unity (Ho et al., 2021; Li; Chen, 2022; Hu et al., 2023). Additionally, Unreal Engine was used for AR simulations by L. Zhang et al., (2023), while Apple ARKit was chosen for iOS-based AR applications by Chu et al., (2024).

Considering the focus of this analysis in the integration of AR and AI, many studies employed software for computer vision and AI model integration such as OpenCV, an open-source computer vision library, used for image processing and object detection (Filipescu et al., 2024; Li et al., 2024b; Geng et al., 2025). Microsoft Visual Studio was also utilized for development environment for scripting and integrating AR with AI (Zhao et al., 2023; Chen et al., 2024; Geng et al., 2025). Additionally, Roboflow, a platform for computer vision dataset management, was employed for object detection models training (Dzeng; Cheng; Cheng, 2024).

Some studies also incorporate tools such as Blender and Blender-Proc2, frequently used for the development of 3D models for AR simulations (Zidek et al., 2019; Aiken et al., 2024), as well as SolidWorks (Raj et al., 2024) and FreeCAD (Grappiolo et al., 2021a), which were employed for generating 3D models used in AR assembly tutorials. In industrial settings, SCADA and SIEMENS TIA Portal were implemented for real-time monitoring and control of systems (Filipescu et al., 2024). Additionally, a statistical analysis software, SPSS, was applied for evaluating experimental results (De Feudis et al., 2022).

A diversity of hardware devices were highlighted in this review, with Head Mounted Displays (HMDs) being the most prominent for immersive experiences with Microsoft HoloLens being the most repeatedly used headset, appearing in 8 out of the total of 27 studies (Li; Chen, 2022; Zhao et al., 2023; Aiken et al., 2024; Chen et al., 2024; Cramer et al., 2024; Dzeng; Cheng; Cheng, 2024; Raj et al., 2024; Seeliger; Weibel; Feuerriegel, 2024). Additionally, HTC Vive, another commercial HMD for mixed reality experiences, was used by De Feudis et al., (2022) and Z. Wang et al., (2023). Some studies employed specialized AR glasses, such as HiAR G200 (Li; Zheng; Zheng, 2021) and Epson Moverio BT350 (Zidek et al., 2019; Ho et al., 2021).

Smartphones and cameras were also common categories of hardware for image capturing and processing. iPhones (Ho et al., 2021; Chu; Chen; Chen, 2024) and other unspecified smartphones (Generosi et al., 2022; Hu et al., 2023), were frequently used due to their embedded camera, mobility, practicality and ARKit support, the latter being exclusive to iPhones. Logitech cameras were repeatedly used for monitoring and recording assembly processes (Lai et al., 2020; Grappiolo et al., 2021a; Ho et al., 2021). Other non-specified webcams were used (Wang; Yan, 2021) as well as monocular cameras (Li et al., 2024b) and 3D IFM Cameras (Filipescu et al., 2024) for depth capture and object recognition data in various experimental setups.

Additionally, custom and specialized hardware platforms were utilized in some of the reviewed articles to support the integration of AR and AI assembly tasks in a more personalized way. For example, (Geng et al., 2025) developed a proprietary portable integrated desktop AR hardware platform named Smart Navigator of Industrial Operations (SNIO), developed for intelligent industrial assembly operations.

Out of the 27 articles analyzed, there was a particular emphasis on Unity (9 studies), Vuforia (4 studies) and MRTK (5 studies). Unity was the most widely used platform for AR development. It was often combined with MRTK and Vuforia to enhance AR functionalities, providing robust object recognition and AR overlay capabilities. It was noted that 7 studies do not mention any specific commercial software, which can suggest that there is still a need for the maturation or creation of software tailored to meet the demands of this sector.

The most used hardware in reviewed articles were HMDs, with Microsoft HoloLens being the most frequently utilized, used in 8 studies. Another important mention is the HTC Vive, employed for Mixed Reality experiences. Apart from HMDs, smartphones and cameras were also prevalent, highlighting the employment of iPhones and Logitech cameras, employed for monitoring and recording assembly processes. It is worth mentioning that the Microsoft Hololens 2 was discontinued in 2023, related to the company's current focus on its tool for military use of the U.S. Army, the Integrated Visual Augmentation System (IVAS) (Seiler, 2023).

How were AR and AI solutions adapted to different assembly scenarios?

The reviewed articles explore a diversity of assembly objects across different industries. Studies examined more complex connectors specific to the fields of aviation (Li; Zheng; Zheng, 2021), electromechanics (Zhang et al., 2023b) and electric (Geng et al., 2025), as well as mechanical parts like flanges (Chu; Chen; Chen, 2024). Specifically in the aviation sector, studies show efficacy of this application on aircraft parts, including cables (Zheng et al., 2020), exhaust vents (Li et al., 2023), and general aircraft components (Zhao et al., 2022), demonstrate the effectiveness of these technologies in high-precision and safety-critical environments.

In addition to mechanical components, several studies emphasize the assembly of electromechanical and structural elements, often used in aviation, maritime, and industrial applications. For instance, ship block assembly (Choi; Seo, 2020) and scaffolding structures (Dzeng; Cheng; Cheng, 2024) emphasizing the value of using AR and AI in large-scale assembly tasks that require precise positioning and alignment.

Another relevant category of assembly objects in this analysis included tools, brackets, and industrial systems that require AI-driven object recognition for improved guidance. For instance, studies on the assembly of cable brackets (Hu et al., 2023) and pneumatic cylinders (Raj et al., 2024) illustrate the application of AR-AI systems in the industrial context. The recognition and use of conventional tools such as screws, pliers, Allen keys, and drills (Lai et al., 2020; Wang; Yan, 2021; De Feudis et al., 2022; Zhao et al., 2023) accentuates how these technologies support floor workers in selecting and using appropriate instruments during assembly. Furthermore, research on specialized equipment, such as hydraulic fracturing valves (Aiken et al., 2024) luminaires (Grappiolo et al., 2021a), workpieces (Filipescu et al., 2024) and spindles (Li; Chen, 2022) revealing the adaptability of AR-AI integration across diverse engineering fields.

Some articles do not mention any specific assembly object but approaches the integration of AR and AI in general sectors, for example: manufacturing work operations (Generosi et al., 2022), construction scenarios (Chen et al., 2024), indoor environments (Seeliger; Weibel; Feuerriegel, 2024), hybrid medical device (Ho et al., 2021), that does not directly mention any specific assembly object but describes the use of assembly parts and related tasks, such as identifying screws, nuts, and other small components, used as well in (Zidek et al., 2019). Moreover, studies related to processes in surgical procedures (Cramer et al., 2024) showcase the potential of AR-AI integration in assisting with delicate operations.

AR and AI solutions were adapted to a variety of assembly scenarios such as aviation, industrial settings, structures and electromechanics. Even inside these sectors, it is visible that the integration of these technologies is applicable to bigger and robust assemblies (scaffoldings, ship blocks, etc.), but also to millimetric tasks (connectors, cables, etc.). Although there is a greater predominance of the integration of these technologies in manufacturing and industrial applications, it is possible to visualize its potential for versatility in other sectors such as surgical procedures and the medical field in general.

What were the common challenges and limitations in implementing AR and AI in assembly processes?

Despite the advantages of integrating AR and AI in assembly processes, there were several challenges in its application. Hardware constraints remain a significant barrier, as the use of HMDs and quality cameras can limit accessibility due to technical limitations such as lack of interoperability beyond specific platforms like HoloLens 2 and Unity (Li; Chen, 2022) and AR glasses usability problems like can be heaviness and fatigue, restricting long-term use (Zheng et al., 2020; Chen et al., 2024).

The computational complexity of AI-models for object detection and recognition requires substantial processing power, which can interfere in real-time responsiveness reported in slower processing times (Li; Zheng; Zheng, 2021) and recognition delays (Zidek et al., 2019). The authors also recognized the insufficiency of datasets in the deep learning models training, needing large and more diverse datasets. The data set limitations affected the system's ability to generalize and adapt to different assembly scenarios (Zheng et al., 2020; Li et al., 2024b). Besides that, (Li; Chen, 2022) also mention the lack of interoperability beyond specific platforms like HoloLens 2 and Unity.

Environmental factors were a common challenge for many studies that identified problems with variations in lighting and occlusions in the assembly space (De Feudis et al., 2022; Zhang et al., 2023b), which consequently affects sensors performance, causing limitations in the captured data. Among these difficulties it may be mentioned gloved hand detection (Generosi et al., 2022) and AR markers recognition (Choi; Seo, 2020). Addressing the user experiments authors mention the limitations of its participants study such as controlled environments, where participants lack industry experience, limiting real-world applicability (Seeliger; Weibel; Feuerriegel, 2024) and real-world conditions, such as weather and equipment availability, introducing unpredictability, further complicating industrial adoption (Chen et al., 2024).

What were the recommendations for future research or development in the application of AR and AI in assembly processes?

To advance the adoption and effectiveness of AR-AI integration, future research should address several key areas. The reviewed studies mentioned many improvements for future work, highlighting the enhancement of detection accuracy and robustness such as improving real-time tracking accuracy (Zhang et al., 2023b; Dzeng; Cheng; Cheng, 2024; Raj et al., 2024), dealing with occlusions, variations in lighting, and ensuring better hand/pose detection even under different circumstances (gloves, scaffolding obstruction, and non-standard scenarios) (Generosi et al., 2022; Dzeng; Cheng; Cheng, 2024; Filipescu et al., 2024).

Regarding errors during the assembly process, it was mentioned the necessity of implementation of fault correction and quality inspection during assembly processes could further streamline industrial applications (Geng et al.,

2025) and enhance operator training, allowing systems to predict assembler errors (Grappiolo et al., 2021a). This necessity is aligned with the results found in this review, since only four of the studies employed error detection in the developed applications.

Expanding AI applications to predictive analytics and dynamic scheduling optimization is another common recommendation, particularly for industries (Choi; Seo, 2020; Filipescu et al., 2024). Cloud-edge computing solutions could also improve inspection services, while adaptive learning models could improve object recognition and positioning accuracy without reliance on markers (Li; Zheng; Zheng, 2021; Li; Chen, 2022). Future research should also explore the enhancement of training methods, such as progressive learning systems for AR-based assembly instructions (Aiken et al., 2024) and studying human factors in technology adoption for better usability evaluation (Chen et al., 2024).

CONCLUSIONS

This SLR aimed to show the current state of research integrating AR and AI applied to assembly tasks objectifying the better understanding of the implementation and development of these technologies, as well as the challenges involved in the employment of those novelties. The existing research demonstrated a diversity of approaches, with no standardized method for effectively combining these technologies in assembly settings. This absence of a unified methodology results in fragmented implementations, making it difficult to assess and compare results. Future studies could establish a structured approach to guide the integration of AI and AR in assembly processes, ensuring consistency and efficiency across different scenarios.

A significant concern is the ephemerality of AR hardware considering an intent of long-term integration with AI. As a relatively new technology, AR devices are subject to rapid evolution, that often leads to short product life cycles and discontinuation of hardware, which was the case with Microsoft HoloLens. This instability makes it difficult for industries to invest in long-term solutions without the risk of obsolescence.

The durability and practicality of AR equipment in environments of industry and manufacture pose concerns since assembly tasks often involve physically demanding activities, which increase the risk of device damage and user exhaustion due to prolonged usage. The available AR devices did not present the practicality and resistance required for those environments, necessitating improvements in device design to enhance durability and usability. Moreover, real-time error detection and instant feedback for assemblers remain possible features that need further development, significantly improving the accuracy and efficiency of assembly tasks, minimizing errors.

Despite the challenges and possible improvements, the integration of AI and AR in manufacturing and industrial assembly is already well established, with significant space for expansion into other domains, such as medicine and surgical procedures. Considering the use of AI models, CNNs and YOLO continue to dominate, despite the emergence of modern models like ViTs that may offer promising improvements.

ACKNOWLEDGEMENTS

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES)

REFERENCES

AFIF, M.; AYACHI, R.; SAID, Y.; PISSALOUX, E.; ATRI, M. An Evaluation of RetinaNet on Indoor Object Detection for Blind and Visually Impaired Persons Assistance Navigation. Neural Processing Letters, v. 51, n. 3, p. 2265–2279, 23 jun. 2020. ISSN 1370-4621. DOI 10.1007/s11063-020-10197-9.

AIKEN, W.; CARDEN, L.; BHABHRAWALA, A.; BRANCO, P.; JOURDAN, G.-V.; BERG, A. Strategic Digitalization in Oil and Gas: A Case Study on Mixed Reality and Digital Twins. IEEE ACCESS, v. 12, p. 87248–87267, 2024. DOI 10.1109/ACCESS.2024.3417391.

ARIKAN, İ.; AYAV, T.; SEÇKIN, A. Ç.; SOYGAZI, F. Estrus Detection and Dairy Cow Identification with Cascade Deep Learning for Augmented Reality-Ready Livestock Farming. Sensors, v. 23, n. 24, 1 dez. 2023. ISSN 14248220. DOI 10.3390/s23249795.

AZZAM, I.; PATE, K.; BREIDI, F.; CHOI, M.; JIANG, Y.; MOUSAS, C. Mixed Reality: A Tool for Investigating the Complex Design and Mechanisms of a Mechanically Actuated Digital Pump. Actuators, v. 12, n. 11, 1 nov. 2023. ISSN 20760825. DOI 10.3390/act12110419.

BATTINENI, G.; CHINTALAPUDI, N.; RICCI, G.; RUOCCO, C.; AMENTA, F. Exploring the integration of artificial intelligence (AI) and augmented reality (AR) in maritime medicine. Artificial Intelligence Review, v. 57, n. 4, p. 100, 27 mar. 2024. ISSN 1573-7462. DOI 10.1007/s10462-024-10735-0.

CARCELLAR, M. T. I. M.; TYCHUACO, C. J. S.; YUMANG, A. N. Self-Applicable Eye Strain Detection Through the Measurement of Blink Rate Using Raspberry Pi. Em: 2024 IEEE International Conference

on Artificial Intelligence in Engineering and Technology (IICAIET), Anais[...]IEEE, 26 ago. 2024. DOI 10.1109/IICAIET62352.2024.10730733.
CHEN, H.; HOU, L.; WU, S.; ZHANG, G.; ZOU, Y.; MOON, S.; BHUIYAN, M. Augmented Reality, Deep Learning and Vision-Language Query System for Construction Worker Safety. Automation in Construction, v. 157, 2024. ISSN 09265805. Disponível em: http://dx.doi.org/10.1016/j.autcon.2023.105158
CHOI, T.; SEO, Y. A Real-Time Physical Progress Measurement Method for Schedule Performance Control Using Vision, an AR Marker and Machine Learning in a Ship Block Assembly Process. SENSORS, v. 20, n. 18, 2020. DOI 10.3390/s20185386.
CHU, C. H.; CHEN, Y. R.; CHEN, S. M. Identification of flange specification in real industrial settings with human reasoning assisted by augmented reality. Advanced Engineering Informatics, v. 62, 1 out. 2024. ISSN 14740346. DOI 10.1016/j.aei.2024.102882.
CHU, C. H.; LIU, Y. L. Augmented reality user interface design and experimental evaluation for human-robot collaborative assembly. Journal of Manufacturing Systems, v. 68, p. 313–324, 1 jun. 2023. ISSN 02786125. DOI 10.1016/j.jmsy.2023.04.007.
CIRESAN, D.; MEIER, U.; SCHMIDHUBER, J. Multi-column deep neural networks for image classification. Em: 2012 IEEE Conference on Computer Vision and Pattern Recognition, Anais[...]IEEE, jun. 2012. DOI 10.1109/CVPR.2012.6248110.
CLARIVATE SUPPORT. KeyWords Plus generation, creation, and changes. Disponível em: https://support.clarivate.com/ScientificandAcademicResearch/s/article/KeyWords-Plus-generation-creation-and-changes?language=en_US. Acesso em: 12 fev. 2025.
CRAMER, E.; KUCHARSKI, A. B.; KREIMEIER, J.; ANDRESS, S.; LI, S.; WALK, C.; MERKL, F.; HÖGL, J.; WUCHERER, P.; STEFAN, P.; VON EISENHART-ROTHE, R.; ENSTE, P.; ROTH, D. Requirement analysis for an AI-based AR assistance system for surgical tools in the operating room: stakeholder requirements and technical perspectives. International Journal of Computer Assisted Radiology and Surgery, 1 nov. 2024. ISSN 18616429. DOI 10.1007/s11548-024-03193-0.
DE FEUDIS, I.; BUONGIORNO, D.; GROSSI, S.; LOSITO, G.; BRUNETTI, A.; LONGO, N.; DI STEFANO, G.; BEVILACQUA, V. Evaluation of Vision-Based Hand Tool Tracking Methods for Quality Assessment and Training in Human-Centered Industry 4.0. APPLIED SCIENCES-BASEL, v. 12, n. 4, 2022. DOI 10.3390/app12041796.
DEVAGIRI, J. S.; PAHEDING, S.; NIYAZ, Q.; YANG, X.; SMITH, S. Augmented Reality and Artificial Intelligence in Industry: Trends, Tools, and Future Challenges. Expert Systems with Applications, v. 207, 2022. ISSN 09574174. Disponível em: http://dx.doi.org/10.1016/j.eswa.2022.118002
DIMITROPOULOS, N.; TOGIAS, T.; ZACHARAKI, N.; MICHALOS, G.; MAKRIS, S. Seamless human–robot collaborative assembly using artificial intelligence and wearable devices. Applied Sciences (Switzerland), v. 11, n. 12, 2 jun. 2021. ISSN 20763417. DOI 10.3390/app11125699.
DU, F.; KONG, F.; ZHAO, D. A Knowledge Transfer Method for Unsupervised Pose Keypoint Detection on Domain Adaptation and CAD Models. ADVANCED INTELLIGENT SYSTEMS, v. 5, n. 2, fev. 2023. DOI 10.1002/aisy.202200214.
DUAN, K.; BAI, S.; XIE, L.; QI, H.; HUANG, Q.; TIAN, Q. CenterNet: Keypoint Triplets for Object Detection. Em: 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Anais[...]IEEE, out. 2019. DOI 10.1109/ICCV.2019.00667.
DZENG, R.-J.; CHENG, C.-W.; CHENG, C.-Y. A Scaffolding Assembly Deficiency Detection System with Deep Learning and Augmented Reality. BUILDINGS, v. 14, n. 2, fev. 2024. DOI 10.3390/buildings14020385.
ELHARROUSS, O.; HIMEUR, Y.; MAHMOOD, Y.; ALRABAEE, S.; OUAMANE, A.; BENSAALI, F.; BECHQITO, Y.; CHOUCHANE, A. ViTs as backbones: Leveraging vision transformers for feature extraction. Information Fusion, v. 118, p. 102951, jun. 2025. ISSN 15662535. DOI 10.1016/j.inffus.2025.102951.
FILIPESCU, A.; SIMION, G.; IONESCU, D.; FILIPESCU, A. IoT-Cloud, VPN, and Digital Twin-Based Remote Monitoring and Control of a Multifunctional Robotic Cell in the Context of AI, Industry, and Education 4.0 and 5.0. Sensors, v. 24, n. 23, 1 dez. 2024. ISSN 14248220. DOI 10.3390/s24237451.
GENEROSI, A.; AGOSTINELLI, T.; CECCACCI, S.; MENGONI, M. A novel platform to enable the future human-centered factory. International Journal of Advanced Manufacturing Technology, v. 122, n. 11–12, p. 4221–4233, 2022. Disponível em: http://dx.doi.org/10.1007/s00170-022-09880-z
GENG, J.; WANG, Y.; CHENG, Y.; ZHANG, X.; XU, Y.; LV, W. Systematic AR-based assembly guidance for small-scale, high-density industrial components. Journal of Manufacturing Systems, v. 79, p. 86–100, 1 abr. 2025. ISSN 02786125. DOI 10.1016/j.jmsy.2025.01.002.
GIRSHICK, R. Fast R-CNN. Em: 2015 IEEE International Conference on Computer Vision (ICCV), Anais[...]IEEE, dez. 2015. DOI 10.1109/ICCV.2015.169.

GRAPPIOLO, C.; PRUIM, R.; FAETH, M.; DE HEER, P. ViTroVo: in vitro assembly search for in vivo adaptive operator guidance An artificial intelligence framework for highly customised manufacturing. INTERNATIONAL JOURNAL OF ADVANCED MANUFACTURING TECHNOLOGY, v. 117, n. 11–12, p. 3873–3893, 2021a. DOI 10.1007/s00170-021-07824-7.

GRAPPIOLO, C.; PRUIM, R.; FAETH, M.; DE HEER, P. ViTroVo: In Vitro Assembly Search for in Vivo Adaptive Operator guidance An Artificial Intelligence Framework for Highly Customised Manufacturing. INTERNATIONAL JOURNAL OF ADVANCED MANUFACTURING TECHNOLOGY, v. 117, n. 11–12, p. 3873–3893, dez. 2021b. ISSN 0268-3768. DOI 10.1007/s00170-021-07824-7.

HAJIRASOULI, A.; BANIHASHEMI, S.; DROGEMULLER, R.; FAZELI, A.; MOHANDES, S. R. Augmented reality in design and construction: thematic analysis and conceptual frameworksConstruction InnovationEmerald Group Holdings Ltd., , 8 jun. 2022. . ISSN 14770857. DOI 10.1108/CI-01-2022-0007.

HO, N.; WONG, P.-M.; HOANG, N.-S.; KOH, D.-K.; CHUA, M. C. H.; CHUI, C.-K. CPS-Based Manufacturing Workcell for the Production of Hybrid Medical. JOURNAL OF AMBIENT INTELLIGENCE AND HUMANIZED COMPUTING, v. 12, n. 12, p. 10865–10879, dez. 2021. ISSN 1868-5137. DOI 10.1007/s12652-020-02798-y.

HU, J.; ZHAO, G.; XIAO, W.; LI, R. AR-Based Deep Learning for Real-Time Inspection of Cable Brackets In. ROBOTICS AND COMPUTER-INTEGRATED MANUFACTURING, v. 83, out. 2023. ISSN 0736-5845. DOI 10.1016/j.rcim.2023.102574.

JAIN, S. A. Future Technology and Labour - Are We Heading Towards a Jobless Future? Indian Journal of Law and Justice, v. 10, n. 2, p. 1 – 15, 2019. ISSN 09763570. Disponível em: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85169144497&partnerID=40&md5=5fb75d1caa9f650f45faf71ecea24e1c

JENKINS, T.; GOODWIN, A.; TALAFHA, S. An ORSAC method for data cleaning inspired by RANSAC. International Journal of Informatics and Communication Technology (IJ-ICT), v. 13, n. 3, p. 484, 1 dez. 2024. ISSN 2722-2616. DOI 10.11591/ijict.v13i3.pp484-498.

LAI, Z.-H.; TAO, W.; LEU, M. C.; YIN, Z. Smart Augmented Reality Instructional System for Mechanical assembly towards Worker-Centered Intelligent Manufacturing. JOURNAL OF MANUFACTURING SYSTEMS, v. 55, p. 69–81, abr. 2020. ISSN 0278-6125. DOI 10.1016/j.jmsy.2020.02.010.

LAMPROPOULOS, G. Combining Artificial Intelligence with Augmented Reality and Virtual Reality in Education: Current Trends and Future Perspectives. Multimodal Technologies and Interaction, v. 9, n. 2, p. 11, 28 jan. 2025. ISSN 2414-4088. DOI 10.3390/mti9020011.

LĂZĂROIU, G.; GEDEON, T.; VALASKOVA, K.; VRBKA, J.; ŠULEŘ, P.; ZVARIKOVA, K.; KRAMAROVA, K.; ROWLAND, Z.; STEHEL, V.; GAJANOVA, L.; HORÁK, J.; GRUPAC, M.; CAHA, Z.; BLAZEK, R.; KOVALOVA, E.; NAGY, M. Cognitive digital twin-based Internet of Robotic Things, multi-sensory extended reality and simulation modeling technologies, and generative artificial intelligence and cyber–physical manufacturing systems in the immersive industrial metaverse. Equilibrium. Quarterly Journal of Economics and Economic Policy, v. 19, n. 3, p. 719–748, 27 set. 2024. ISSN 23533293. DOI 10.24136/eq.3131.

LI, C.; ZHENG, P.; ZHOU, P.; YIN, Y.; LEE, C. K. M.; WANG, L. Unleashing mixed-reality capability in Deep Reinforcement Learning-based robot motion generation towards safe human–robot collaboration. Journal of Manufacturing Systems, v. 74, p. 411–421, jun. 2024a. ISSN 02786125. DOI 10.1016/j.jmsy.2024.03.015.

LI, M.; CHEN, Y. Using Artificial Intelligence Assisted Learning Technology on Augmented Reality-Based Manufacture Workflow. Frontiers in Psychology, v. 13, 2022. DOI 10.3389/fpsyg.2022.859324. Disponível em: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85134175130&doi=10.3389%2ffpsyg.2022.859324&partnerID=40&md5=51f319f806fdb55abb77c9cfe3987962

LI, S.; ZHENG, P.; ZHENG, L. An AR-Assisted Deep Learning-Based Approach for Automatic Inspection of Aviation Connectors. IEEE Transactions on Industrial Informatics, v. 17, n. 3, p. 1721–1731, 2021. Disponível em: http://dx.doi.org/10.1109/TII.2020.3000870

LI, W.; WANG, J.; LIU, M.; ZHAO, S.; DING, X. Integrated Registration and Occlusion Handling Based on Deep Learning for Augmented-Reality-Assisted Assembly Instruction. IEEE Transactions on Industrial Informatics, v. 19, n. 5, p. 6825–6835, 2023. Disponível em: http://dx.doi.org/10.1109/TII.2022.3189428

LI, W.; XU, A.; WEI, M.; ZUO, W.; LI, R. Deep Learning-Based Augmented Reality Work Instruction Assistance system for Complex Manual Assembly. JOURNAL OF MANUFACTURING SYSTEMS, v. 73, p. 307–319, abr. 2024b. ISSN 0278-6125. DOI 10.1016/j.jmsy.2024.02.009.

LIU, G.; GUO, J. Bidirectional LSTM with attention mechanism and convolutional layer for text classification. Neurocomputing, v. 337, p. 325–338, abr. 2019. ISSN 09252312. DOI 10.1016/j.neucom.2019.01.078.
LIU, Y.; ZHANG, S.; GOWDA, M. A Practical System for 3-D Hand Pose Tracking Using EMG Wearables With Applications to Prosthetics and User Interfaces. IEEE Internet of Things Journal, v. 10, n. 4, p. 3407–3427, 2023. ISSN 23274662. Disponível em: http://dx.doi.org/10.1109/JIOT.2022.3223600
LOWE, D. G. Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision, v. 60, n. 2, p. 91–110, nov. 2004. ISSN 0920-5691. DOI 10.1023/B:VISI.0000029664.99615.94.
LUO, L.; LI, Y.; WANG, X.; JIN, X.; QIN, Z. Supply Chain Vulnerability in Prefabricated Building Projects and Digital Mitigation Technologies. IEEE TRANSACTIONS ON ENGINEERING MANAGEMENT, v. 71, p. 10686–10698, 2024. ISSN 0018-9391. DOI 10.1109/TEM.2023.3272585.
MA, R.; SHI, H.; GAO, H.; GUAN, H.; IQBAL, M.; MUMTAZ, S. CFedDT: Cross-Domain Federated Learning in Digital Twins for Metaverse Consumer Electronic Products. IEEE Transactions on Consumer Electronics, v. 70, n. 1, p. 3167–3182, 1 fev. 2024. ISSN 15584127. DOI 10.1109/TCE.2023.3327010.
MERTZ, L. New Center Primes the Extended Reality Frontier. IEEE Pulse, v. 14, n. 3, p. 7–11, 1 maio 2023. ISSN 21542317. DOI 10.1109/MPULS.2023.3294083.
MIENYE, I. D.; SUN, Y. A Survey of Ensemble Learning: Concepts, Algorithms, Applications, and Prospects. IEEE Access, v. 10, p. 99129–99149, 2022. ISSN 2169-3536. DOI 10.1109/ACCESS.2022.3207287.
MUELLER, R.; VETTE, M.; MASIAK, T.; DUPPE, B.; SCHULZ, A. Intelligent Real Time Inspection of Rivet Quality Supported by Human-Robot-Collaboration. SAE International Journal of Advances and Current Practices in Mobility, v. 2, n. 2, p. 811–817, 2019. ISSN 26419645. Disponível em: http://dx.doi.org/10.4271/2019-01-1886
NAQVI, W. M.; NAQVI, I. W.; MISHRA, G. V.; VARDHAN, V. D. The future of telerehabilitation: embracing virtual reality and augmented reality innovations. The Pan African Medical Journal, v. 47, 2024. ISSN 1937-8688. DOI 10.11604/pamj.2024.47.157.42956.
OKEH, E. Transforming Healthcare: A Comprehensive Approach to Mitigating Bias and Fostering Empathy through AI-Driven Augmented Reality. Proceedings of the AAAI Conference on Artificial Intelligence, v. 38, n. 21, p. 23753–23754, 24 mar. 2024. ISSN 2374-3468. DOI 10.1609/aaai.v38i21.30553.
PANAGOU, S.; SILEO, M.; PAPOUTSAKIS, K.; FRUGGIERO, F.; QAMMAZ, A.; ARGYROS, A. Complexity based investigation in collaborative assembly scenarios via non intrusive techniques. Procedia Computer Science, v. 217, p. 478–485, 2023. ISSN 18770509. DOI 10.1016/j.procs.2022.12.243.
PRADHAN, Manaranjan.; DINESH KUMAR, U. . Machine learning with Python. [s.l.] Wiley, 2019. ISBN 9788126579907.
QIN, Y.; BULBUL, T. Electroencephalogram-Based Mental Workload Prediction for Using Reality Head Mounted Display in Construction Assembly: A Deep Approach. AUTOMATION IN CONSTRUCTION, v. 152, ago. 2023. ISSN 0926-5805. DOI 10.1016/j.autcon.2023.104892.
RAJ, S.; MURTHY, L. R. D.; SHANMUGAM, T. A.; KUMAR, G.; CHAKRABARTI, A.; BISWAS, P. Augmented Reality and Deep Learning Based System for Assisting Assembly. JOURNAL ON MULTIMODAL USER INTERFACES, v. 18, n. 1, p. 119–133, mar. 2024. ISSN 1783-7677. DOI 10.1007/s12193-023-00428-3.
REDMON, J.; DIVVALA, S.; GIRSHICK, R.; FARHADI, A. You Only Look Once: Unified, Real-Time Object Detection. Em: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Anais[...]IEEE, jun. 2016. DOI 10.1109/CVPR.2016.91.
ROOSAN, D. Augmented Reality and Artificial Intelligence: Applications in Pharmacy. Em: [s.l: s.n.]p. 227–243. DOI 10.1007/978-3-031-27166-3_13.
RUBLEE, E.; RABAUD, V.; KONOLIGE, K.; BRADSKI, G. ORB: An efficient alternative to SIFT or SURF. Em: 2011 International Conference on Computer Vision, Anais[...]IEEE, nov. 2011. DOI 10.1109/ICCV.2011.6126544.
SAHU, C. K.; YOUNG, C.; RAI, R. Artificial intelligence (AI) in augmented reality (AR)-assisted manufacturing applications: a review. International Journal of Production Research, v. 59, n. 16, p. 4903–4959, 2021. ISSN 1366588X. DOI 10.1080/00207543.2020.1859636.
SANTHI, B.; GURUMOORTHY, B.; CHAKRABARTI, A.; SEN, D. Assessment of Cause of Difficulty in Assembly Tasks. Em: [s.l: s.n.]p. 563–572. DOI 10.1007/978-81-322-2232-3_49.
SAPTARI, A.; JIA XIN LEAU; MOHAMAD, N. A. The effect of time pressure, working position, component bin position and gender on human error in manual assembly line. Em: 2015 International Conference on Industrial Engineering and Operations Management (IEOM), Anais[...]IEEE, mar. 2015. DOI 10.1109/IEOM.2015.7093793.
SASIKALA, S.; SUBHASHINI, S. J.; ALLI, P.; JANE RUBEL ANGELINA, J. Deep Learning Applications in Medical Imaging. Em: [s.l: s.n.]p. 178–208. DOI 10.4018/978-1-7998-5071-7.ch008.

SEELIGER, A.; WEIBEL, R. P.; FEUERRIEGEL, S. Context-Adaptive Visual Cues for Safe Navigation in Augmented Reality Using Machine Learning. International Journal of Human-Computer Interaction, v. 40, n. 3, p. 761–781, 2024. ISSN 10447318. Disponível em: http://dx.doi.org/10.1080/10447318.2022.2122114

SEILER, R. Microsoft committed to HoloLens 2 and Mixed Reality. Disponível em: https://techcommunity.microsoft.com/blog/themixedrealityblog/microsoft-committed-to-hololens-2-and-mixed-reality/3732405. Acesso em: 16 fev. 2025.

SOORI, M.; AREZOO, B.; DASTRES, R. Artificial intelligence, machine learning and deep learning in advanced robotics, a review. Cognitive Robotics, v. 3, p. 54–70, 2023. ISSN 26672413. DOI 10.1016/j.cogr.2023.04.001.

ŞTEFAN, L.; GHEORGHIU, D.; HODEA, M.; MOŢĂIANU, M. The History of Furniture Objects: An Intelligent Augmented Reality Application. Em: [s.l: s.n.]p. 171–190. DOI 10.1007/978-3-031-27166-3_10.

STORK, S.; SCHUBÖ, A. Cognition in Manual Assembly. KI - Künstliche Intelligenz, v. 24, n. 4, p. 305–309, 14 nov. 2010. ISSN 0933-1875. DOI 10.1007/s13218-010-0054-y.

SUN, G.; WANG, C.; ZHANG, Z.; DENG, J.; ZAFEIRIOU, S.; HUA, Y. Spatio-temporal Prompting Network for Robust Video Feature Extraction. Em: 2023 IEEE/CVF International Conference on Computer Vision (ICCV), Anais[...]IEEE, 1 out. 2023. DOI 10.1109/ICCV51070.2023.01250.

SZAJNA, A.; KOSTRZEWSKI, M. AR-AI Tools as a Response to High Employee Turnover and Shortages in Manufacturing during Regular, Pandemic, and War Times. Sustainability (Switzerland), v. 14, n. 11, 2022. ISSN 20711050. DOI 10.3390/su14116729. Disponível em: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85132298900&doi=10.3390%2fsu14116729&partnerID=40&md5=5f3b10a8622e62d205c0622b660bf8d0

TOLENOV, S.; OMAROV, B. Real-Time Self-Localization and Mapping for Autonomous Navigation of Mobile Robots in Unknown Environments. International Journal of Advanced Computer Science and Applications, v. 15, n. 10, 2024. ISSN 21565570. DOI 10.14569/IJACSA.2024.0151090.

TRICCO, A. C.; LILLIE, E.; ZARIN, W.; O'BRIEN, K. K.; COLQUHOUN, H.; LEVAC, D.; MOHER, D.; PETERS, M. D. J.; HORSLEY, T.; WEEKS, L.; HEMPEL, S.; AKL, E. A.; CHANG, C.; MCGOWAN, J.; STEWART, L.; HARTLING, L.; ALDCROFT, A.; WILSON, M. G.; GARRITTY, C.; LEWIN, S.; GODFREY, C. M.; MACDONALD, M. T.; LANGLOIS, E. V.; SOARES-WEISER, K.; MORIARTY, J.; CLIFFORD, T.; TUNÇALP, Ö.; STRAUS, S. E. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Annals of Internal Medicine, v. 169, n. 7, p. 467–473, 2 out. 2018. ISSN 0003-4819. DOI 10.7326/M18-0850.

UMA, S. Latest Research Trends and Challenges of Computational Intelligence Using Artificial Intelligence and Augmented Reality. Em: [s.l: s.n.]p. 43–59. DOI 10.1007/978-3-030-02674-5_3.

WANG, K.-J.; YAN, Y.-J. A Smart Operator Assistance System Using Deep Learning for Angle Measurement. IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT, v. 70, 2021. DOI 10.1109/TIM.2021.3124044.

WANG, L.; GAO, R.; VÁNCZA, J.; KRÜGER, J.; WANG, X. V.; MAKRIS, S.; CHRYSSOLOURIS, G. Symbiotic human-robot collaborative assembly. CIRP Annals, v. 68, n. 2, p. 701–726, 1 jan. 2019. ISSN 17260604. DOI 10.1016/j.cirp.2019.05.002.

WANG, X. Deep Learning in Object Recognition, Detection, and Segmentation. Foundations and Trends® in Signal Processing, v. 8, n. 4, p. 217–382, 2016. ISSN 1932-8346. DOI 10.1561/2000000071.

WANG, Z.; ZHANG, X.; LI, L.; ZHOU, Y.; LU, Z.; DAI, Y.; LIU, C.; SU, Z.; BAI XIAOLIANG AND BILLINGHURST, M. Evaluating visual encoding quality of a mixed reality user interface for human-machine co-assembly in complex operational terrain. ADVANCED ENGINEERING INFORMATICS, v. 58, 2023. DOI 10.1016/j.aei.2023.102171.

XIE, J.; LIU, Y.; WANG, X.; FANG, S.; LIU, S. A new XR-based human-robot collaboration assembly system based on industrial metaverse. Journal of Manufacturing Systems, v. 74, p. 949–964, 1 jun. 2024. ISSN 02786125. DOI 10.1016/j.jmsy.2024.05.001.

YUE, Z.; HAN, Z.; YANG, X.; LIU, L. DCSPose: A Dual-Channel Siamese Framework for Unseen Textureless Object Pose Estimation. Applied Sciences (Switzerland), v. 14, n. 2, 2024. ISSN 20763417. DOI 10.3390/app14020730. Disponível em: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85192465548&doi=10.3390%2fapp14020730&partnerID=40&md5=8d1c6f3d32a569db3288084965896fe7

ZHANG, C.; ZHOU, G.; MA, D.; WANG, R.; XIAO, J.; ZHAO, D. A deep learning-enabled human-cyber-physical fusion method towards human-robot collaborative assembly. Robotics and Computer-Integrated Manufacturing, v. 83, 1 out. 2023a. ISSN 07365845. DOI 10.1016/j.rcim.2023.102571.

ZHANG, L.; PAN, Y.; WU, X.; SKIBNIEWSKI, M. J. Introduction to Artificial Intelligence. Em: [s.l: s.n.]p. 1–15. DOI 10.1007/978-981-16-2842-9_1.
ZHANG, L.; XU, F.; LIU, Y.; ZHANG, D.; GUI, L.; ZUO, D. A Posture Detection Method for Augmented Reality-Aided Assembly Based on YOLO-6D. INTERNATIONAL JOURNAL OF ADVANCED MANUFACTURING TECHNOLOGY, v. 125, n. 7–8, p. 3385–3399, abr. 2023b. ISSN 0268-3768. DOI 10.1007/s00170-023-10964-7.
ZHAO, G.; FENG, P.; ZHANG, J.; YU, C.; WANG, J. Rapid offline detection and 3D annotation of assembly elements in the augmented assembly. Expert Systems with Applications, v. 222, 2023. Disponível em: http://dx.doi.org/10.1016/j.eswa.2023.119839
ZHAO, G.; HU, J.; XIAO, W.; ZOU, J. A mask R-CNN based method for inspecting cable brackets in aircraft. Chinese Journal of Aeronautics, v. 34, n. 12, p. 214–226, 1 dez. 2021. ISSN 10009361. DOI 10.1016/j.cja.2020.09.024.
ZHAO, Q.; KONG, Y.; SHENG, S.; ZHU, J. Redundant Object Detection Method for Civil Aircraft Assembly Based on machine Vision and Smart Glasses. MEASUREMENT SCIENCE AND TECHNOLOGY, v. 33, n. 10, out. 2022. ISSN 0957-0233. DOI 10.1088/1361-6501/ac7cbd.
ZHENG, L.; LIU, X.; AN, Z.; LI, S.; ZHANG, R. A smart assistance system for cable assembly by combining wearable augmented reality with portable visual inspection. Virtual Reality and Intelligent Hardware, v. 2, n. 1, p. 12–27, 2020. Disponível em: http://dx.doi.org/10.1016/j.vrih.2019.12.002
ZHENG, P.; LI, C.; FAN, J.; WANG, L. A vision-language-guided and deep reinforcement learning-enabled approach for unstructured human-robot collaborative manufacturing task fulfilment. CIRP Annals, v. 73, n. 1, p. 341–344, 1 jan. 2024. ISSN 17260604. DOI 10.1016/j.cirp.2024.04.003.
ZIDEK, K.; LAZORIK, P.; PITEL, J.; HOSOVSKY, A. An Automated Training of Deep Learning Networks by 3D Virtual Models for Object Recognition. SYMMETRY-BASEL, v. 11, n. 4, abr. 2019. DOI 10.3390/sym11040496.