Advanced Engineering Informatics 54 (2022) 101787

Contents lists available at ScienceDirect

Advanced Engineering Informatics

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

Full length article

Robotics in construction: A critical review of the reinforcement learning and imitation learning paradigms

Juan Manuel Davila Delgado *, Lukumon Oyedele

Big Data Enterprise and Artificial Intelligence Laboratory, Faculty of Business and Law, University of West of England (UWE) Bristol, Coldharbour Lane, BS16, 1QY Bristol, UK

A R T I C L E I N F O

Keywords:
Robotics
Construction
Reinforcement learning
Imitation learning
Deep reinforcement learning

A B S T R A C T

The reinforcement and imitation learning paradigms have the potential to revolutionise robotics. Many successful developments have been reported in literature; however, these approaches have not been explored widely in robotics for construction. The objective of this paper is to consolidate, structure, and summarise research knowledge at the intersection of robotics, reinforcement learning, and construction. A two-strand approach to literature review was employed. A bottom-up approach to analyse in detail a selected number of relevant publications, and a top-down approach in which a large number of papers were analysed to identify common relevant themes and research trends. This study found that research on robotics for construction has not increased significantly since the 1980s, in terms of number of publications. Also, robotics for construction lacks the development of dedicated systems, which limits their effectiveness. Moreover, unlike manufacturing, construction's unstructured and dynamic characteristics are a major challenge for reinforcement and imitation learning approaches. This paper provides a very useful starting point to understating research on robotics for construction by (i) identifying the strengths and limitations of the reinforcement and imitation learning approaches, and (ii) by contextualising the construction robotics problem; both of which will aid to kick-start research on the subject or boost existing research efforts.

1. Introduction

The idea that robotics would revolutionise the construction industry has been around for many years. However, it has not materialised. [36] noted that despite numerous attempts to develop robotic systems for construction field operations, few practical applications could be found in construction sites. They emphasised that the promises of robotics remained unfulfilled and the attempts to transfer robotic technologies from manufacturing have not succeeded. While some commercial robotic solutions exist nowadays, the overall situation has not changed in a meaningful manner in the last three decades.

In recent years, reinforcement learning (RL) —a machine learning (ML) paradigm— has been considered as an approach with the potential to facilitate and widen the applicability of robotics to many other fields besides the traditional ones, such as automotive and advanced manufacturing. RL is thought to be well-suited for robotics task planning and control; because instead of meticulously programming high-dimensional robot movements step-by-step in a manual manner, RL can be used to train robot behaviour autonomously [91]. Moreover, research indicates that RL methods could be readily applicable to robotics by carefully adapting existing RL implementations form other fields [79].

The RL paradigm enables an agent to learn an optimal sequence of tasks to reach an objective by interacting with an environment that provides feedback by showing the effects of the agent's actions on the environment and a corresponding reward (Fig. 1). This approach is advantageous compared with traditional control methods, because only the objectives and rewards need to be specified and not the actual tasks, which is a very laborious job. The main disadvantage of RL approaches is that the agent needs to explore a large space of actions and their corresponding environment states to find a successful set of actions that accomplish the desired objective. Thus, RL approaches are computationally intensive; and, the processing requirements increase massively with marginal increments in the environment's complexity, i.e. the state-action spaces that need to be explored. The Imitation Learning (IL) paradigm intends to address the aforementioned issue of the exponential explosion of state-action mappings occurring in complex environments, for which evaluating every action-state mapping becomes unfeasible. In

* Corresponding author.
E-mail addresses: manuel.daviladelgado@uwe.ac.uk (J. Manuel Davila Delgado), l.oyedele@uwe.ac.uk (L. Oyedele).

https://doi.org/10.1016/j.aei.2022.101787
Received 30 October 2020; Received in revised form 28 May 2022; Accepted 16 October 2022
Available online 29 October 2022

IL, the agent is provided in advance with demonstrations of high-reward actions to follow in given situations by an expert, thus during the learning process the exploration is focused and the learning time reduced (Fig. 1). In IL, the agent can also ask feedback directly to the expert in situations where the best action is not clear, and the expert can interact with the environment as well. Note that IL can be considered a subtype of RL even though there are significant differences between RL and IL.

Given the potential benefits of RL and IL approaches for robotics, they could represent a perfect opportunity to improve development efforts in robotics for construction. For instance, RL and IL could potentially enable the use of robotics in unstructured environments such as construction sites and enable robots with low sensitivity to constant changes in the environment. However, there is a lack of structured knowledge in this area. Even though RL and IL have been reported in literature for many years, they have not been employed for robotics in construction widely. The lack of structure limits the transfer of knowledge being produced in computer science and manufacturing fields to construction.

The primary objective of this paper is to consolidate, structure, and summarise research knowledge on RL-based robotics for construction. Also, this paper seeks to explore the potential effects that RL and IL approaches might have on advancing the development of robotics for construction. More specifically, this paper explores the state-of-the-art RL and IL approaches being used to improve robotics, outlines the strengths and limitations that could make adoption feasible, and characterises the construction situation in terms of RL-based robotics. In sum, it provides an overview of the research landscape at the intersection of robotics, reinforcement learning, and construction.

The paper is structured as follows: in the next section the research method employed is presented, then section 3 presents an overview of the RL paradigm. The state-of-research, the characterisation of research, and research challenges at the intersection of robotics, reinforcement learning, and construction are presented in sections 4, 5, and 6 respectively. Lastly, the discussion and conclusions are provided in sections 7 and 8.

2. Research method

The research method consists of three phases (Fig. 2):

(I) Initial Exploration. A preliminary literature review was conducted to identify potential impacts that the RL and IL paradigms might have on research related to robotics for construction. Publications concerning this topic in specific were not found; therefore, publications that address how RL and IL impact research in robotics were used to identify relevant themes for a further more detailed investigation on the subject, e.g. [59,102,7,87,83]. Because this topic has not been systematically investigated before, a "two-strand" approach for literature review was employed. The first strand is a *bottom-up* approach, in which the aforementioned papers were analysed in detail to characterise the RL and IL paradigms to then identify manners in which they can impact research on robotics for construction. The second strand is a *top-down* approach, in which themes identified in the papers above were used to search for relevant publications in academic databases. Note that both strands were carried out in parallel.

(II) Bottom-up approach. This first strand consists of three steps: (A) *Characterising RL and IL* (section 3), in which first an overall description of the RL paradigm was outlined; then, a coarse categorisation of RL and IL was delineated to facilitate the understanding of the existing approaches; and lastly, the strengths and limitations of the approaches were identified. (B) *Challenge Identification* (section 6), in which RL and IL challenges for robotics were compiled from literature and categorised for easier understanding; (C) *Challenge translation* (section 6), in which construction was formulated as an RL problem in terms of its site and tasks characteristics. Also, the known RL and IL challenges were translated to construction-specific challenges and put in the context of the construction sector characteristics.

(III) Top-down approach. The second strand has four steps: (1) *Literature review* (section 4.1); in this step, several literature searches were carried out using terms identified in the other two phases above. The Scopus database and the database storing all the publications presented at the International Symposium on Automation and Robotics in Construction (ISARC) were used in this study. These two sources provide a broad search scope and variety. The four levels of queries are (i) high-level searches in Scopus and ISARC, (ii) mid-level searches in Scopus, (iii) mid-level searches in ISARC, and (iv) granular searches in Scopus. See Fig. 2 for details on the search terms used and the number of resulting publications. For the granular searches in Scopus, a title and abstract screening was carried out to identify relevant papers to the construction sector and a detailed screening to identify papers that address robotics for construction specifically. The resulting papers were used in step number four "research characterisation" described below. (2) *General thematic analyses* (section 4.2), in which three analyses were carried out: (i) Keyword analysis, which investigated the top-20 most common keywords used in 795 publications on robotics, construction, and machine learning listed in Scopus. The most used keywords were defined by using the Scopus's keyword ranking and by condensing different keywords that refer to the same term, e.g., "neural networks" and "artificial neural networks". (ii) ML methods analysis, which identified the most used ML methods in robotics publications. In this case, the keywords indexed in 40,574 publications were analysed and condensed to identify the most used methods. (iii) Term co-occurrence, in which the VOS Viewer software [35] was used to analyse the terms in titles, abstracts, and keywords from 795 publications on robotics,

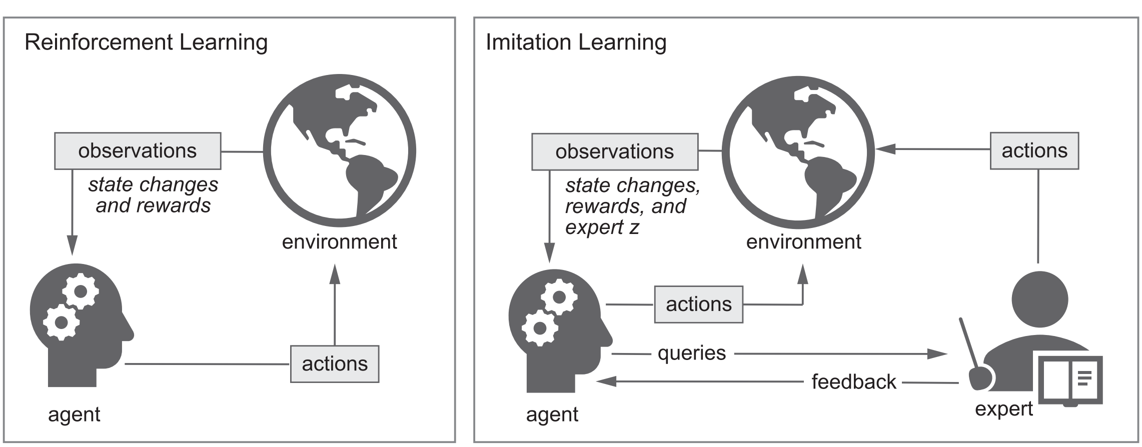

Fig. 1. The reinforcement and imitation learning paradigms.

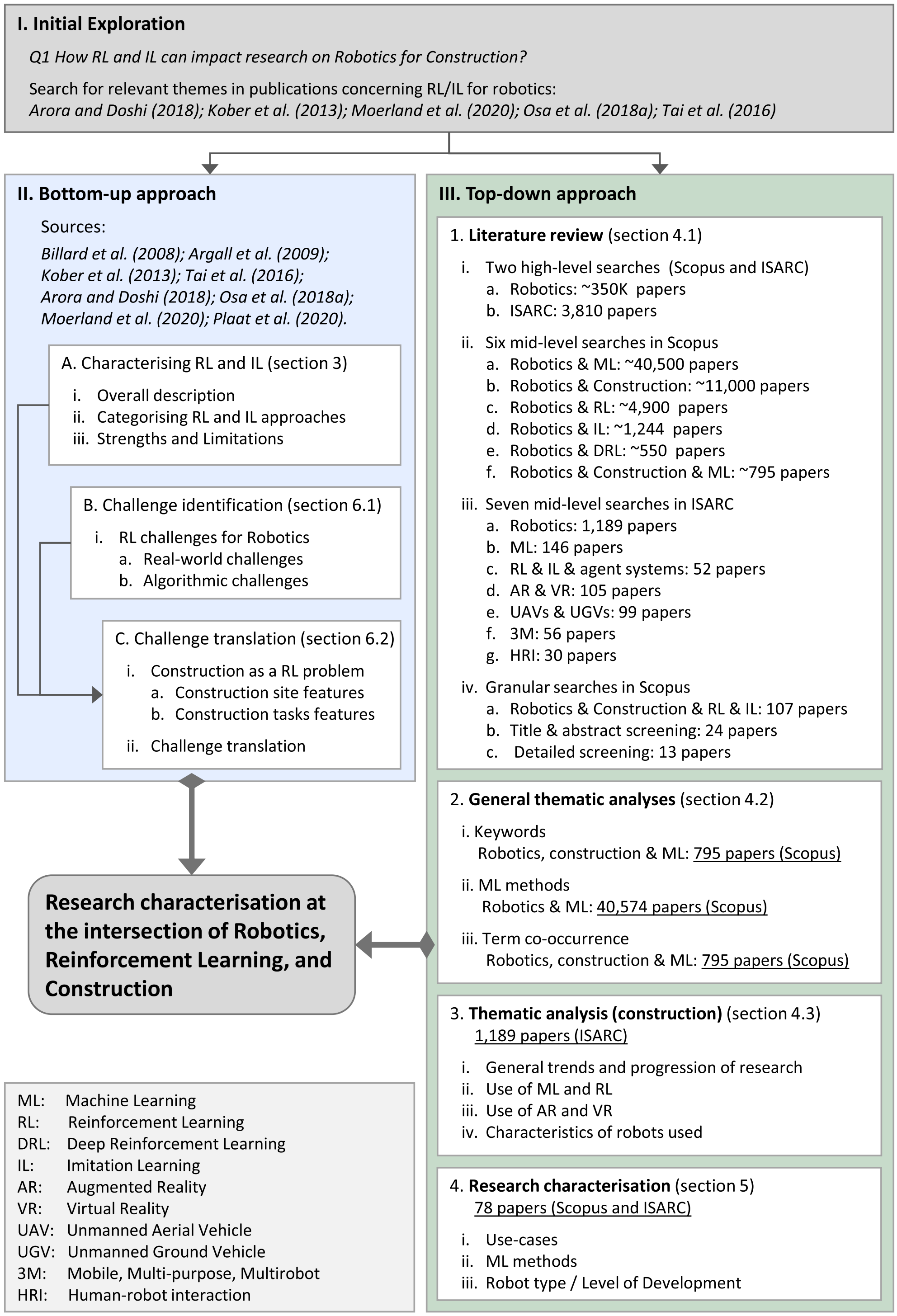

Fig. 2. Diagram illustrating the three phases of the research method used.

construction, and machine learning. (3) *Thematic analysis in construction publications* (section 4.3), in which 1,189 publications related to robotics published at ISARC were analysed. The following aspects were investigated: general trends and the progression of publications; the use of ML, RL, and IL; the use of augmented reality (AR) and virtual reality (VR) in conjunction with robotics; and the characteristics of the employed robots. (4) *Research characterisation* (section 5), in which publications at the intersection of robotics, reinforcement learning, and construction were analysed in detail. Seventy-eight publications were selected to carry out a publication-level analysis in terms of use-cases, ML methods, types of robots, and the level of development of the research reported in those publications. The papers analysed were selected from the mid-level searches in ISARC and the granular searches in Scopus (Fig. 2).

Note that for the analyses using papers resulting from the mid-level searches in ISARC and the granular search in Scopus, publications related to off-site construction, additive manufacturing, 3D printing, and construction process automation were not considered. Note that were taken into account only publications written in English, published in conferences and journals in the areas of computer science, engineering, and mathematics. Also, different terms that refer to the same types of approaches as imitation learning were used to broaden the scope of the query, i.e., behavioural cloning, inverse reinforcement learning, and learning from demonstration.

3. The reinforcement learning paradigm

RL is a field of study within ML, as well as a type of problem and a class of solution methods. The generic RL solution method maps environment states to actions to maximise a numerical reward signal. The learning is carried out by finding what actions yield the most reward through trial and error and incentivising a higher but delayed reward [101]. RL is very different from the other two main ML strands, i.e., *supervised learning*, in which learning is achieved by providing a set of correct predictions or mappings between a description of a situation and a label of a correct action. Supervised learning solutions aim to generalise its responses to select the correct mapping for situations that are not present in the dataset used for learning. Conversely, in RL the correct answer is not known, and learning is achieved through exploration. The other strand is *unsupervised learning,* in which the objective is to find hidden structures within unlabelled datasets. Its primary aims are clustering analysis, summarising data, and feature explanation [43]. Table 1 presents the main differences between supervised, unsupervised, and RL. Note that IL is not included in the table because it can be regarded as an extended version of IL as most of IL approaches use RL techniques in one way or another.

3.1. The general RL approach

The main components are (i) *agents, environments, actions, and states*. An agent is an information construct that explores an environment by taking actions (a_t) and receives feedback on the outcome of the choices made. The environment is the simulated environment in which the agent interacts. At every step of the interaction (t), the agent observes a partial state of the environment (s_t) and takes an action. The environment reacts to the actions by changing the state and can change the state on its own as well. The action space is the set of all possible actions in an environment. There are discrete action spaces in which a finite number of discrete actions can be executed; and continuous action spaces in which a finite number of continuous actions can be taken.

(ii) *Rewards*. In addition to the state, at every time step, the agent is also provided with a reward (r_t), which is a scalar value indicating the level of success or failure to reach the desired objective every time the agent takes an action. The total reward (R_t), is the summation of all the rewards (r_t). (iii) The *Value Function* specifies what actions will increase rewards in the long run. The value (Q) of a state is the total reward that an agent can expect to amass in the future starting from that given state. Rewards determine immediate successes, while value indicates long-term accumulated success. There are two main types of value functions: (a) Discounted reward, that incentivises the quick accumulation of rewards by including a discounting factor lambda and multiplying it by all of the individual rewards obtained ($\gamma^t r_t$). The formulation describes the expected total future reward that an agent in a given state can achieve by executing an action, as follows $Q_{discounted}(s_t, a_t) = \mathrm{E}\{\gamma^t R_t | s_t, a_t\}$. The second one is (b) Average Reward, in this case, the function is formulated to maximise the average reward across time so that high rewards nearby and distant in time are equally preferred. In this case, an average reward (μ) is subtracted to each reward. This differential reward represents how much more reward the agent will receive from the current state in action compared to the average reward. The average reward formulation is then $Q_{average}(s_t, a_t) = \mathrm{E}\{R_t - \mu^\pi | s_t, a_t\}$, in which ($\pi$) represents the selected policy, which is further explained below.

(iv) *Policy* is a rule that the agent uses to decide what actions to take. More formally, a policy (π) is a state-action mapping that maximises the cumulative expected reward, i.e., $\pi^*(a) = \max_a Q(s, a)$. There are deterministic policies in which the same action is always taken for a given state, i.e., $a = \pi(s)$; and probabilistic policies in which an action is drawn from a distribution of actions, i.e. $a \sim \pi(s, a) = P(a|s)$. Contrary to supervised learning, the agent does not have information about high-reward actions and must first explore the environment to find successful policies. Once that the agent has identified high rewards, it has to decide whether to stick with those or try new actions to find even higher rewarded actions. In this sense, two types of exploration methods exist (a) *Off-policy* methods, in which the exploration strategy is codified outside the selected policy and can be employed during the learning process, and (b) *On-policy* methods that collect sample information about the environment using the employed policy, thus the exploration strategy is included into the policy itself. Note that this is only a brief overview of the RL approach. Please refer to literature for more detailed definitions and formulations, e.g., [59,101].

Table 1
Types of machine learning approaches.

Type	Supervised Learning	Unsupervised Learning	Reinforcement Learning
Data	Labelled data:(x, y) x is data, y is the label	Data:(x) x is unlabelled data	Data:(s_t, a_t) s_t is state, a_t is action
Goal	Learn function to map: $x \rightarrow \hat{y}$	Learn an underlying structure to find relationships: $P_{model}(x) \cong P_{data}(x)$	Maximise future reward over many time steps through interaction with the environment or with an expert
Problems	Classification, Regression	Clustering, feature and dimensionality reduction	Optimal Control
Examples	Deep Neural Networks Convolutional Neural Networks Recurrent Neural Networks	Autoencoders Variational Autoencoders	Q-Learning Policy Optimisation

3.2. Deep reinforcement learning

Deep Learning (DL) is a term that represents a type of neural network architecture. The term “deep” denotes the relatively large number of hidden layers in a neural network. Note that DL architectures can be used for both in supervised and unsupervised learning. Regarding Deep Reinforcement Learning (DRL), the term “deep” denotes that the agent uses deep neural networks to find successful state-action mappings; while in traditional RL the agent uses a relational table of mappings between states and actions, called Q table. For most of the real-life problems, traditional RL approaches are not suitable because the size and complexity of the problems lead to extremely large Q tables that are practically impossible to compute. Thus, replacing the Q table with a deep neural network enables to estimate a state-action mapping by approximating the complexity of the function describing the environment. Fig. 3 presents the RL and DRL general architectures. Note that while in the traditional RL architecture a state-action pair maps to a single expected reward, in DRL each state is mapped to all the expected rewards for all possible actions.

3.3. A categorisation of approaches for the RL and IL paradigms

Fig. 4 presents a high-level categorisation of approaches for the RL and IL paradigms, while Table 2 and 3 presents the main benefits, disadvantages, and literature surveys. RL approaches can be classified into two major categories depending on whether the agent has access or can learn a model of the environment, i.e., *model-free* and *model-based*. On the other hand, according to (Osa et al., 2018a), IL approaches can be

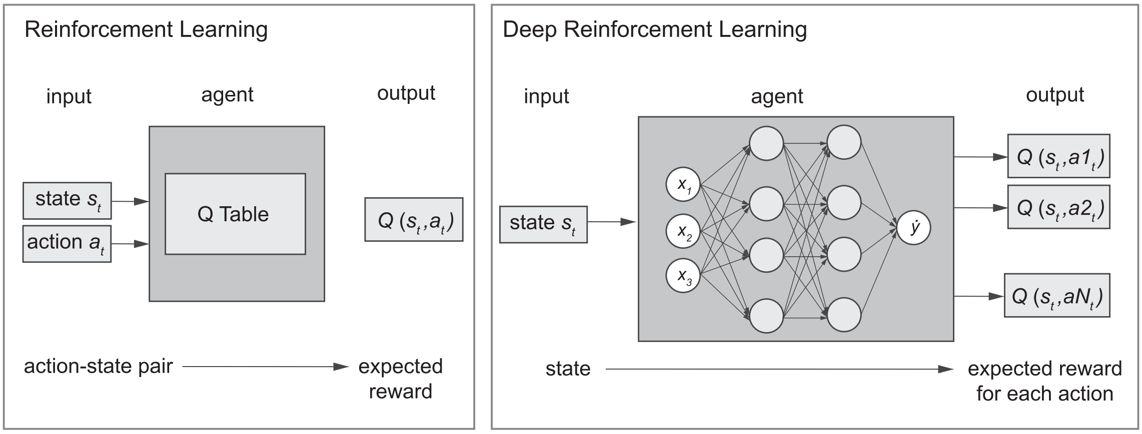

Fig. 3. Main differences between RL and DRL.

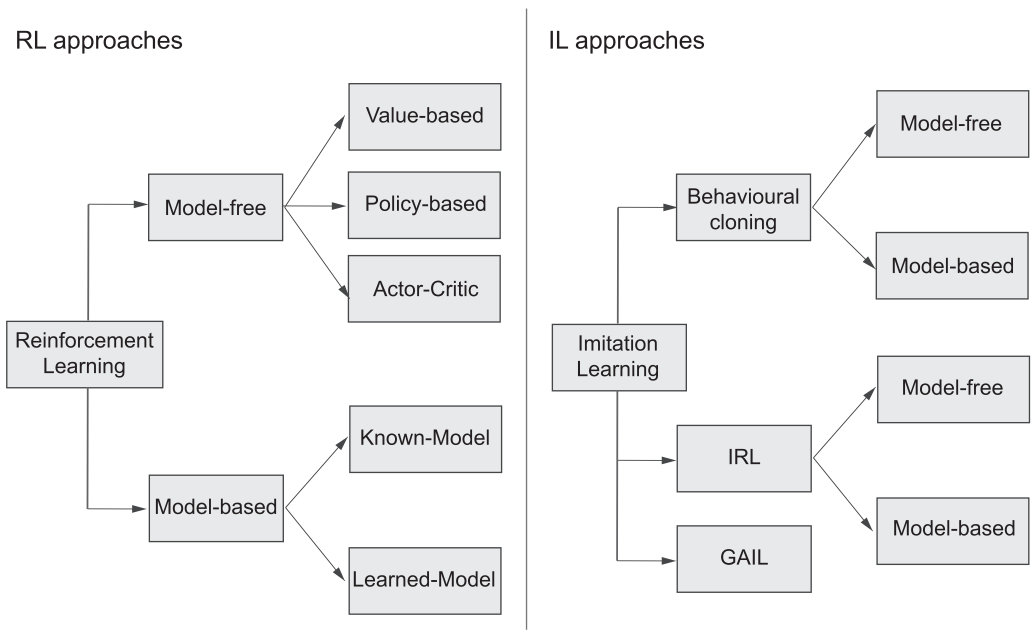

Fig. 4. A categorisation of RL and IL approaches.

classified in categories depending on whether the agent derives the optimal policy —observed from the expert— by using supervised learning, i.e., *behavioural cloning*, or a reinforcement learning approach, i.e. *inverse reinforcement learning* (IRL). [102] also include a *Generative Adversarial Imitation Learning* (GAIL) as another category which combines generative adversarial neural networks and model-free IRL [37]. A more granular analysis of the RL and IL approaches is presented in the two subsequent sections. Note that this categorisation is not extensive or definitive. Due to the modularity of algorithms, it is difficult to represent the RL and IL fields neatly in a hierarchical diagram. Some state-of-the-art approaches combine algorithms form different categories and mix learning strategies. Nevertheless, Fig. 4 and Table 2 highlight the most fundamental differences and helps to identify trade-offs among the most common approaches.

3.4. Reinforcement Learning: model-based and model-free

A common way to classify RL approaches is *model-free* approaches, in which the agent learns a value function or a policy by exploring through interactions with the environment; and *model-based*, in which the agent has access to a model of the environment, i.e., a function that describes the state transitions and rewards.

Model-Based RL. The main advantage of model-based approaches is that they can potentially learn optimal policies in less time. By having access to a model of the environment, the agents can simulate state changes for a range of actions and select the best option; thus, facilitating to find optimal policies in fewer interactions with the environment. This feature is usually called sample efficiency. The difference in the number of interactions required to learn a policy between model-based and model-free approaches can be up to three orders of magnitude higher. This is very important for robotics because learning time is very expensive in physical environments compared with computer-simulated environments. The main disadvantage of model-based approaches is that for most real-life cases, it is very unlikely to have access to a sufficiently detailed model of the environment. Even when a model is available, errors or inaccuracies in the environment models result in policies exploiting the model deficiencies. This issue is called model-bias and can be identified when the agent performs well in the model of the environment but sub-optimally in the real environment.

Model-based approaches can be divided into approaches for which the model is given, i.e. *known-model*, e.g. [96], and in which the model needs to be learned, i.e. *learned-model*, e.g. [110,41], see Fig. 4. The

Table 2
Benefits and disadvantages of RL approaches extracted from the referred surveys.

	Benefits	Disadvantages	Use-case examples	Surveys
Model-Based	- Sample efficient (fast learning) Safer exploration approaches Targeted exploration Potential for transfer learning Explainable	- Difficult to generate accurate models Model-bias Additional computation to learn the model Instability due to uncertainties and model errors Large number of tuneable parameters	- Navigation Robotics control Maze solving Object transportation	[102 83 91 59]
Known-model	- Learns only the policy	- Known models exist only for limited environment (e.g. board games)		
Learned-model	- Model learned using supervised learning approaches	- Learns the model and the policy		
Model-Free	- Easy to implement Fast computation Low tuning requirements Relatively stable	- Sample inefficient (slow learning) Potential unsafe exploration Unexplainable in some cases Low level of generalisation	- 2D simplified navigation 2D object transportation	[102 59]
Value-based	- Effective in discrete action spaces Effective for exhaustive action space searches Relatively easy problem to compute	- Inefficient for large action spaces Inefficient for continuous action spaces Requires total coverage of the state space Difficult to scale		
Policy-based	- Effective in continuous action spaces Effective in high-dimensional action spaces Only requires limited coverage of state space Relatively easy to scale	- Significantly harder problem to compute Usually only local optima are found		

Table 3
Benefits and disadvantages of IL approaches extracted from the referred surveys.

	Benefits	Disadvantages	Use-case examples	Surveys
Behavioural cloning	- Simple implementation Efficient for short-term environments Effective for small state spaces	- Accumulation of errors Lead to potential unknown states Requires almost total coverage of the state space	- Trajectory learning Trajectory transfer Probabilistic movement	[102 88]
Inverse Reinforced Learning	- No need to specify reward function Limited manual task specification Improved generalisation Robust against changes in the environment	- Ill-posed problem Potential ambiguous solutions Low scalability Very high computational costs for relatively small state and action spaces	- Complex drone maneuvers Multi-robot patrolling	[7 6 15 59]
Model-based	- Data efficient learning process Learned policy satisfies system dynamics	- Learning the model is a difficult task Computationally expensive		
Model-free	- Policy is learned directly	- Difficult policy estimation of long-term trajectories System dynamics is only implied		

known-model approaches are essentially planning algorithms in which the agent only needs to learn a policy given the model and a particular state [83]; while in the learned-model approaches the agent needs to learn both the model and the policy. The learned-model approaches usually interact with the model using a base policy and then use supervised learning approaches to learn an optimal policy. Note that model-based learned-model approaches are more sample efficient than model-free approaches because the model is a simplification of the environment.

Model-free RL. Contrary to model-based approaches, model-free approaches can be regarded as pure trial-and-error algorithms. Model-free approaches are relatively easy to implement and to fine-tune, they are by far the most reported in literature, but they are very sample inefficient. The main disadvantage of model-free approaches is that their high-sample requirements limit their application almost uniquely to simulated domains [108]. Model-free approaches lack data efficiency, targeted exploration, transfer learning, safety, and are not explainable; but require less computation and memory requirements and are far easier to implement with significantly fewer parameters to tune. Most model-free approaches can be divided into *value-based* and *policy-based* methods. Other terms that refer to different types of value-based methods include value function, Q-learning, critic-only, value learning, among others; while terms referring to policy-based methods include policy search, policy optimisation, actor-only, and policy learning.

Value-based methods focus on estimating the optimal values in given state-action mappings and then deriving an optimal policy from the estimated values. Value-based methods are very useful for discrete action spaces and where a complete search of the action space is possible, but for large and continuous spaces they are less effective, e.g., [14]). In the case of robotics, value-based methods cannot effectively deal with high-dimensional spaces and an approximation for the value function is usually required. Conversely, policy-based methods focus on optimising policies directly. In these methods, the policy is parameterised, and optimal parameters that maximise the policy's objective function are estimated, e.g., [94]. Policy-based methods are more effective when dealing with high-dimensional or continuous action spaces [102]; and they can exploit both deterministic and stochastic policies. Policy-based methods present significant benefits for robotics as could allow for straightforward integration of expert knowledge, allow for a domain-appropriate pre-structuring of the policy, and can be scaled up relatively easy [59]). The main disadvantage of policy-based methods is that it is not guaranteed to find a global optimum. There are other methods referred to as Actor-Critic methods that seek to address the limitations of both value-based and policy-based methods. These methods are a combination of value-based and policy-based methods. They use explicit representations of values, referred as the critic, and policy estimations, referred as the actor, e.g., [63,42].

3.5. Imitation learning: Behavioural cloning and inverse reinforcement learning

For complex environments and tasks, specifying a reward function or learning an optimal policy is very hard to accomplish. Thus, IL approaches exploit the fact that for human experts is easier to demonstrate a desired task or behaviour than to specify it in sufficient detail for replication. The general idea is that an expert (a human or agent) demonstrates how to perform a task and the agent learns a policy equivalent to the demonstrated task. Some of the main differences with RL is that in IL a transition model is used that describes the probability that an action a in state s leads to a subsequent state s', i.e., $P(s'|s,a)$. In IL the reward function $R(s,a)$ is unknown, but the agent has access to the expert's demonstration, known as trajectory, which lists a series of state-action pairs $\tau = ([s1,a1],[s2,a2]\cdots)$ representing an "optimal" policy π^*. The main disadvantage of IL approaches is that is usually unfeasible for the expert to demonstrate every possible state, and that the quality of the demonstrations can be ambiguous or suboptimal in certain areas of the state space [6]. There are two main types of IL approaches, i.e., *behavioural cloning* and *inverse reinforcement learning* (IRL). Note that there are model-based and model-free variants for both behavioural cloning and IRL as well.

Behavioural cloning. These approaches leverage supervised learning to learn the expert's policy. The policy that reproduces the demonstrated behaviour is obtained by directly mapping the agents' input to the expert trajectory; then a supervised learning approach is used to learn the expert's policy, e.g. [77,87]. The main challenge of behavioural cloning is that the state-action pairs do not hold the independent and identically distribution assumption required for supervised learning approaches; thus, errors in different states accumulate and can lead to unknown or never-trained states. Some behavioural cloning approaches mitigate this issue by enabling access to a demonstrator (a human or agent) during the learning process, thus accumulating more training data and purging potential errors [2].

Inverse Reinforcement Learning. In these approaches the objective is to learn the reward function –instead of the policy– directly from an expert's trajectory, and then find the optimal policy using an RL approach, e.g. [1,60]. The main IRL challenge is that, as with many inverse problems, finding a reward function associated with an expert's trajectory is an ill-posed problem. In the IRL case, many reward functions, even sub-optimal ones, can explain the expert's trajectories [85], and two or more very similar reward functions may yield very different policies. Also, IRL approaches tend to grow disproportionately in complexity with the problem size. See [7] for a detailed explanation of the limitations of IRL approaches.

Similar to RL approaches, there are model-based and model-free IL approaches. In this case, the difference resides in whether the agent has access to a model of the environment. For IL-model-based approaches, the learning process is data-efficient, and the learned policy is ensured to satisfy the system dynamics. However, learning the model is a complicated and computationally expensive task. For IL-model-free approaches, the policy can be learned directly, but policy estimation, especially in long-term trajectories is more difficult (see Table 3). Note that the policy learned by an IRL approach is valid as long as the estimated reward function represents the desired trajectory correctly, while a policy learned by a behavioural cloning approach is valid as long as the learned state-action mapping is valid [87].

3.6. A comparison between RL and IL approaches

The main difference between RL and IL approaches is that in RL approaches the exploration is completely unrestricted, while in IL the exploration is guided by expert demonstrations.

Table 2 presents a comparison among RL approaches. A first distinction between model-based and model-free methods is that model-based methods enable fast and safe learning because exploration is constrained by the model. However, the drawback is that the models are usually very difficult and time consuming to generate and might not reflect reality accurately. In model-based methods, a model can be given, for example an agent learning to play a board game (known-model), or the model can be learned in conjunction with the policy. On the other hand, model-free approaches are relatively easy to implement, but the learning can be very time-consuming and the learned policies unsafe or not relevant to the desired behaviours. Model-free approaches can be categorised in value-based and policy-based. Value-based approaches are commonly used for relatively simple problems in discrete environments and action spaces; while policy-based approaches are used for more complex problems requiring continuous environments and action spaces.

Table 3 presents a comparison among IL spaces. The main distinction between behavioural cloning and IRL is that in behavioural cloning the reward functions need to be defined explicitly while in IRL it is not needed. Behavioural cloning is better suited for small state spaces and short-term environments; while IRL is better suited for changing and noisy environments, but it is an ill-posed problem that can lead to ambiguous solutions. The same distinctions between model-based and model-free approaches apply here. Lastly, the major drawback for all RL and IL approaches is the limitation on the size of state and action spaces, usually around 20 [7], which limits its applications for real-world applications.

4. State-of-research at the intersection of robotics, reinforcement learning, and construction

This section presents three analyses that seek to increase the understanding of the research carried out at the intersection of robotics, reinforcement learning, and construction. First is presented an overview of the evolution in the number of publications related to robotics, reinforcement learning, and construction. Secondly, a thematic analysis investigates machine learning methods, most common keywords, and co-occurrence of terms. Lastly, a thematic analysis is presented in the papers presented at the International Symposium on Automation and Robotics in Construction (ISARC) from 1984 to 2019.

4.1. Evolution of research publications in robotics, reinforcement learning, and construction

Systematic and institutionalised research efforts on robotics for construction started in the mid-1980s [17], which is evidenced by the increasing number of publications on the subject. Fig. 5 presents two graphs that provide an overview of the evolution of research on robotics and construction by mapping the progression in the number of publications loosely related to robotics and construction.

Fig. 5a presents the number of publications listed in Scopus related to variations of the terms (i) "robotics" and (ii) "robotics" and "construction" from 1980 to 2020 (see details in Fig. 2). The first striking point is that research efforts on robotics and construction represent only a tiny fraction of the total research on robotics, as there are about 30 times more publications about robotics in general than for robotics and construction. Secondly, a slight quasi-linear increase in the number of publications on robotics and construction can be appreciated, but this trend is very different from the robotics' trend in general. A noticeable increase in the number of publications related to robotics starts at the beginning of the 1980s; then, it plateaus at the end of the 1980s and during the 1990s. However, at the beginning of the 2000′s the number of publications massively increases until the end of that decade, only to experience another huge increase by the end of the 2010s.

Fig. 5b presents the progression in the number of publications presented at ISARC from its first edition in 1984 to 2019. ISARC is probably the most reputable and robust international conference on robotics for construction, and it represents a valuable source of information to

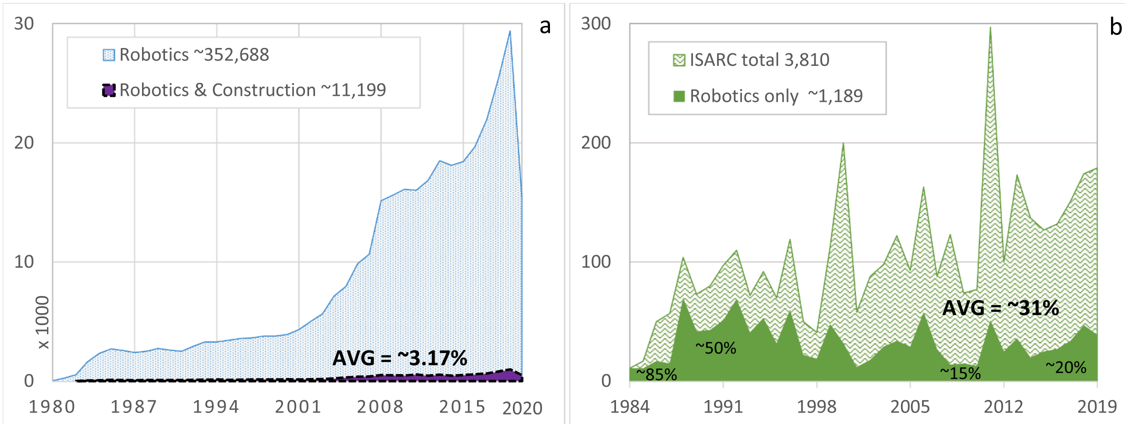

Fig. 5. (a) Progression in the number of publications on Scopus relating to “robotics” in general and to “robotics” and “construction”. (b) Progression in the number of papers presented at the ISARC conferences highlighting the papers focusing on “robotics”.

understand research on robotics for construction. The first ISARC was held in the US in 1984, since then the symposium has been held every year. Fig. 5b differentiates the publications that contain variations of the terms “robotics” and “drone” in their titles, keywords, and abstracts from all the other publications presented at ISARC. Note that publications referring to off-site construction, additive manufacturing, and process automation are not differentiated. From the 3,810 papers presented at ISARC since 1984, only around 31 % are in some way related to robotics. More importantly, the ratio of papers relating to robotics presented at ISARC has been decreasing. In the first editions, the percentage of robotics-related papers was around 85 %, quickly decreasing to 50 % in the 1990s. The ratio decreased to $\sim$ 15 % by the mid-2000s, its lower level, and in recent years has increased to approximately 20 %. This trend is significantly different from the increasing trend of the total number of publications presented at ISARC.

Fig. 6 presents a more granular analysis of the evolution in the number of publications given relevant search terms for this study. The intention is to provide an overview of the research interests in the intersecting areas of robotics, reinforcement learning, and construction. Six sets of terms were defined and searched for in the Scopus database (see details in Fig. 2). The number of publications per year for the six sets are presented in Fig. 6. The sets of search terms are: (1) “robotics”, (2) “robotics” and “construction”, (3) “robotics” and “reinforcement learning”, (4) “robotics” and “deep reinforcement learning”, (5) “robotics” and “imitation learning”, and (6) “robotics” and “construction” and “machine learning”. Note that the term “robotics” was included to provide a reference of the larger field, and that the numbers per year have been multiplied by a scaling factor so that detail in the other trends is not lost.

Fig. 6 indicates that the number of publications on robotics and construction has trailed the overall trend in robotics publications in general. Publications for robotics and construction started in the mid-1980s, which aligns with the seminal activities in Japan in this field that kick-started massive research efforts in robotics for construction, as

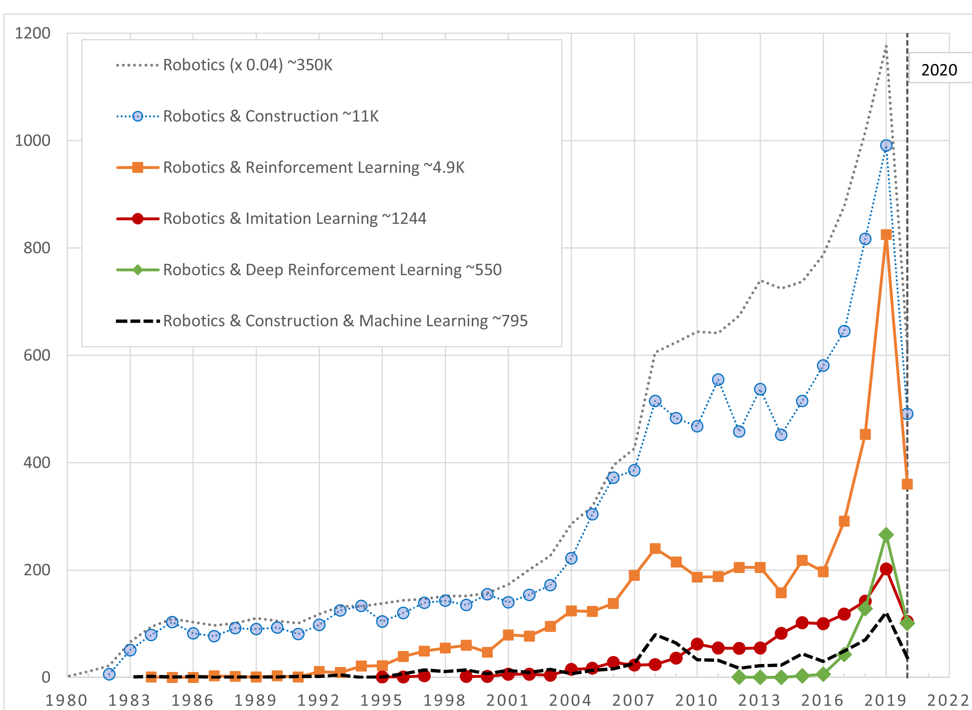

Fig. 6. Progression of the number of relevant publications. An approximate number of the total number of publications per series is presented in the legend. Note that for clarity the yearly numbers of “robotics” publications have been multiply by a scaling factor equal to 0.04.

various initiatives were created including university research groups, research institutes, conferences, and industry technology centres [17]. The second major increase was at the beginning of the 2000s, which aligns with the massive increase in publications on machine learning [26].

Research on robotics and RL also follows a similar trend than for robotics in general. Nevertheless, in this case, the trend starts a continuous increase at the beginning of the 1990s and then plateaus in the late mid-2000s only to increase sharply in the mid-2010s. Compared with IL, the total number of RL publications is almost four times more than IL publications. IL publications started almost a decade later around the mid-1990s; and it was until the beginning of the 2000s that a noticeable continuous increase occurred, a decade later than RL publications as well. The overall robotics and IL trend is different as well. IL publications show a continuous constant increase, while RL presents a massive peak in publications in the last few years increasing the difference in yearly publication greatly.

DRL publications have only recently appeared and have had a huge increase of publications following the RL trend. The number of DRL publications surpassed the yearly publications on IL in 2019, while having started almost 20 decades later. It is noteworthy that research on IL and robotics has not followed the other trends closely, but there is an apparent uptick in publications from the mid-2010s. The DRL trend could be explained by the recent explosion of interest in deep learning and the research successes on computer vision, which have ample applicability on robotics. On the other hand, these successes have less of an impact for IL, which traces its origins to planning algorithms and control theory [88].

Note that the progression of publications on "Robotics & Construction & Machine Learning" is plotted as well in Fig. 6 to provide a baseline and to identify whether specific approaches to RL have influenced research interests in robotics for construction. In this sense, it could be argued that RL and DRL could have potentially influenced research on robotics for construction much more than IL. More importantly, note that the "Robotics & Construction & Machine Learning" trend might be a more accurate representation of the overall research on robotics for construction than the "Robotics & Construction" trend. This is because unrelated publications cannot be effectively filtered out due to limitations on the database searches. Also, it fits better with the trends observed in both charts presented in Fig. 5.

4.2. General thematic analyses

In this section, results from three thematic analyses on research publications are presented, which analyse (i) the most common keywords, (ii) the most-used machine learning methods in robotics, and (iii) the co-occurrence of relevant terms in titles, keywords, and abstracts.

4.2.1. Most common keywords in robotics, construction, and machine learning publications

Fig. 7 illustrates the most common top-20 keywords used in 795 publications on robotics, construction, and machine learning (see section 2). The keywords were grouped based on their similarity and ordered based on the number of instances of each keyword in all the papers. These groups of keywords indicate important research topics for robotics, machine learning, and construction. The first group clusters keywords referring to special types or robots or agents. The relevant terms here are "mobile", "multi", and "autonomous", which characterise ideal requirements in robotics for construction. For example, robots in construction need to be mobile, multi-purpose, autonomous, and should be able to collaborate with other robots.

The second group refers to the ability of robots to sense and interpret their environment, which in most cases is done visually using cameras and computer vision techniques. Examples abound for this type of research. For example, Weng et al. [111] presented an approach based on convolutional neural networks (CNN) for supporting robot grasp detection using computer vision. The approach used CNN detection in two phases to estimate the objects pose and the picking angle. Small object detection is an unsolved issue in computer vision, and in robotics is a key limitations in various situations. In this sense, Bai et al. [9] presented a so called "single shot multi-box detector" for detecting small objects; and, Gao, Liu and Ju [38] presented and approach to detect hand gestures where the detection target is very small and far from the camera. Regarding construction, Huang et al. [49] presented an

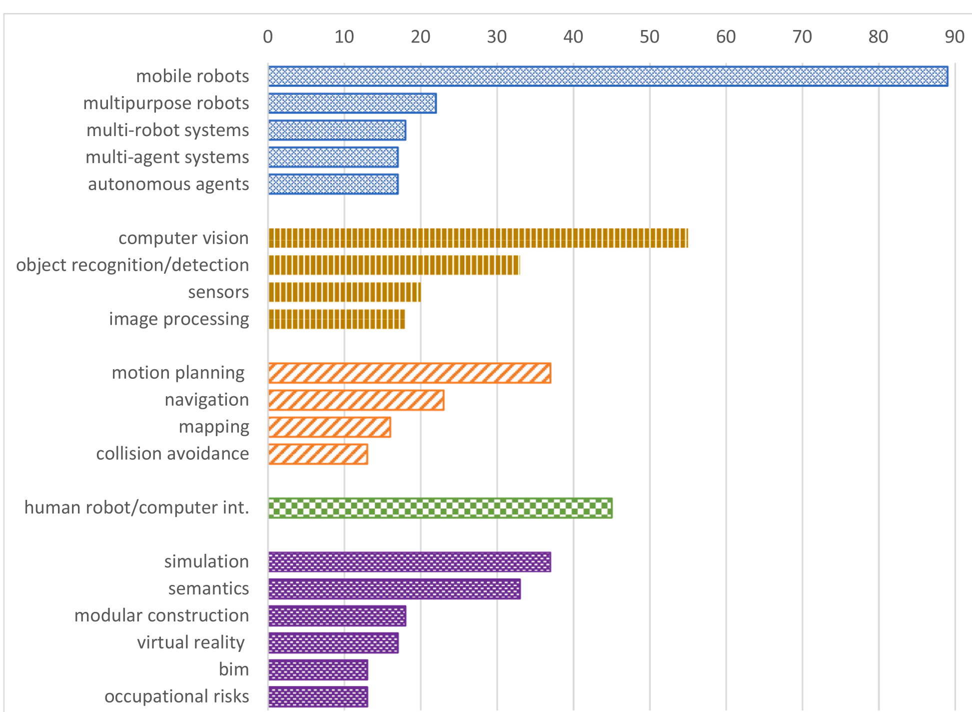

Fig. 7. The most common 20 keywords used in publications on robotics, construction, and machine learning.

approach to detect whether workers are wearing a helmet. The authors note that this approach increases the feature map scale and improves the loss function convergence.

The third group relates to the robot's ability to navigate the environment. There have been sizable research efforts in this aspect as well. For instance, Zhang et al. [114] presents an approach for trajectory planning that improves the kinematic and optimization algorithms of traditional robot navigation approaches. Research in this area incorporates various methods used for optimisation and leverages simulation and synthetic data. For example, X. Liu et al. [74,71,73] present an approach that leverages genetic algorithms for path planning and calibrates and optimises the path by combining virtual and real data. Detection and navigation approaches are combined as well. Yuting Liu et al. [74,71,73] present an approach that combines target detection, searching, localisation, and navigation using images and depth data.

The fourth group considers the interaction between robots and humans; both crucial for the development of effective robotics systems in construction sites. The last group clusters varied keywords with different but important themes, i.e., simulation, semantics, modular construction, virtual reality (VR), Building Information Modelling (BIM), and occupational risks. It is clear the relevance of modular construction and BIM for robotics, as modular components will facilitate assembly and BIM models are indispensable for automating construction tasks. The term occupational risk is also very relevant, as robots represent large safety implications and will increase risks to an already hazardous sector. The term simulation refers to the simulated agents and environment required for RL. In contrast, the term semantics refers to the desired ability of robots to move from literal interpretations to meaningful and context-aware interpretations when interacting with humans or learning from expert demonstrations [40]. Lastly, the relevance of VR for robotics was identified some time ago [25], but recently the most extensive research efforts of robotics and VR is surgery e.g. [19] and medical rehabilitation, e.g. [30].

4.2.2. *Most-used machine learning methods in robotics*

Fig. 8 presents the most mentioned algorithms in approximately 40,500 publications on robotics and machine learning published in the last four decades that are listed in Scopus (see section 2). The methods have been grouped based on their similarity, and the groups have been ordered based on the highest number of instances for each individual term. Neural networks and all their most common variants populate the first group. Reinforcement Learning is in the second group, with a very significant number of mentions. Imitation Learning and its two main approaches are in third place with a significantly lower number of mentions. Evolutionary and classical machine learning approaches are in fourth and fifth places, respectively. Lastly, Bayesian networks complete the group in sixth place. The prominence of the first group can be explained because for almost every machine learning approach for robotics a neural network is needed. For example, to capture states and rewards from real-life or simulated environments, to find optimal policies in behavioural cloning, or to capture expert's trajectories in imitation learning. Fig. 8 also shows the prominence of RL over IL. A reason for this disparity could be that implementing an RL approach is significantly easier than implementing an IL approach, and it has considerably fewer requirements. For example, IL approaches require the implementation of the expert demonstrations, and in some cases, an additional supervised learning implementation, alongside the IL algorithm.

4.2.3. *Analysis of term co-occurrence*

Fig. 9 shows a co-occurrence diagram indicating the most relevant terms found in titles, abstracts, and keyword lists in 795 research publications on robotics, construction, and machine learning. The publications are the same as the ones used for the keyword analysis presented in Fig. 7. The diagram was generated using the software called VOS Viewer [35]. Only the terms occurring more than ten times in the papers are included in the diagram. The terms are arranged within the diagram based on the co-occurrences in the titles, abstracts, and keywords using the mapping technique called visualisation of similarities[34], in which the terms with the higher the number of co-occurrences are placed closer together in the map. The size of the circles represents the number of instances of each term. Lines link terms that appear in the same paper. The thickness of the line indicates how often the keywords appear together in different papers. Lastly, the keywords are clustered into groups using a clustering technique presented by [105].

Five clusters have been defined and named by the term with the most

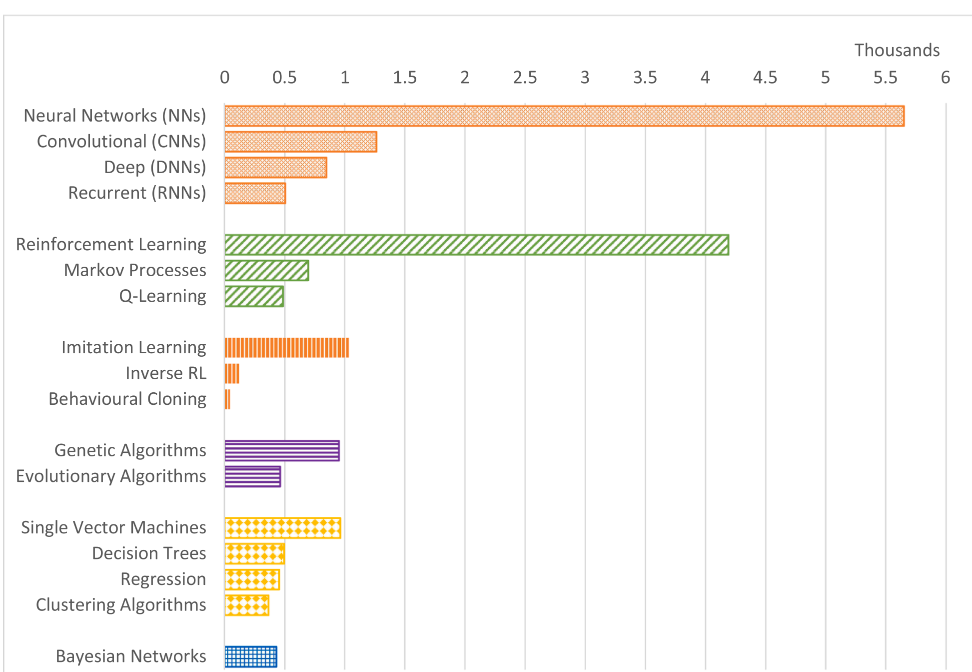

Fig. 8. The most mentioned algorithms in publications on robotics and machine learning.

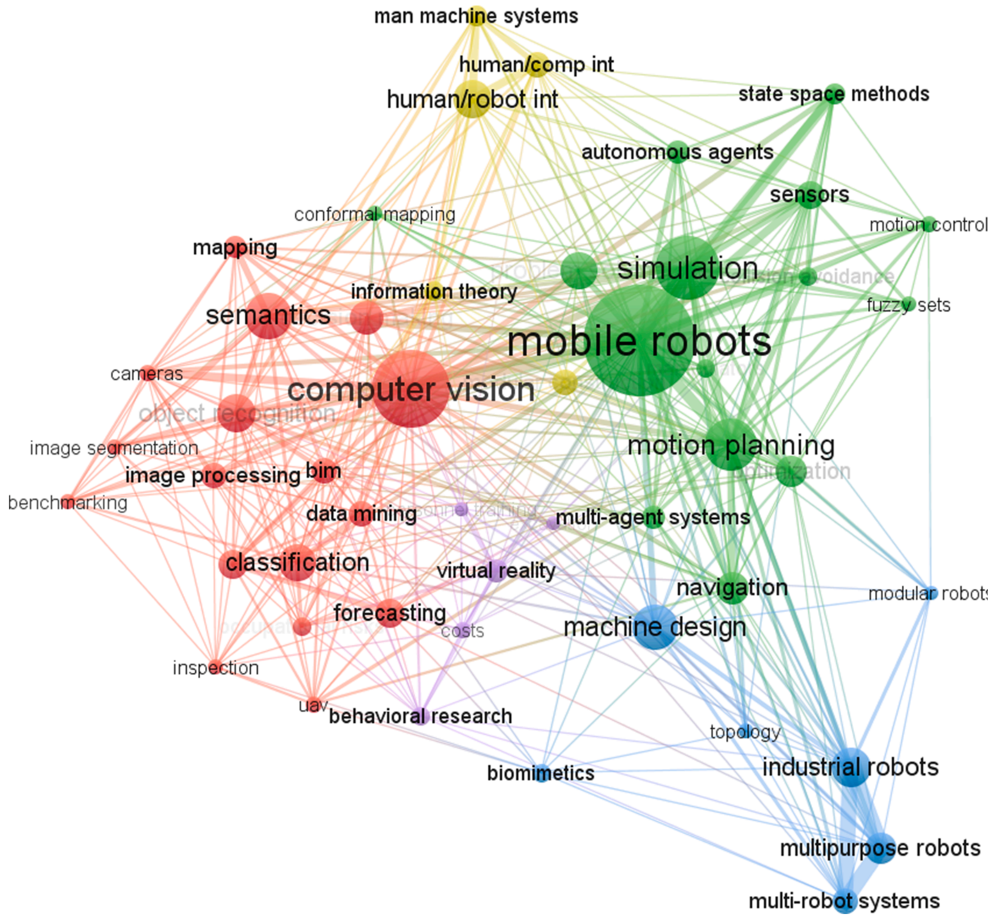

Fig. 9. A co-occurrence diagram showing the most used terms in the titles, abstracts, and keywords in robotics, construction, and machine learning publications.

instances for each cluster (Fig. 9), as follows: (a) Mobile robots, which include terms related to mobility such as motion planning, navigation, motion control, and collision avoidance. It also includes terms such as simulation and optimisation. (b) Computer vision, which agglutinates terms relating to computer vision methods such as object recognition, image processing and segmentation, classification; but also, terms were included referring to hardware, e.g., cameras and UAV, and others such as inspection and occupational risks. (c) Machine design, which clusters terms relating to types of robots, i.e., industrial robots, multi-purpose robots, multi-robot systems, and modular robots. (d) Human-robot interaction, which groups terms relating almost solely to the interaction between humans and robots, computers, and machines. (e) Virtual reality, which includes the terms behavioural research, costs, personnel training, and excavation.

Note that the co-occurrence diagram was generated by an ensemble of machine learning methods, while the categorisation in Fig. 7 was done manually. Terms identified in both analyses and the categories and clusters were used to carry out a more detail exploration of those terms in all the ISARC papers. Note as well that "search terms" used for the database queries, e.g. "robotics", "robots", "construction", among others, were omitted from the analyses presented in Fig. 7 and Fig. 9; because the intention is to identify the prominence of other keywords and terms related to the "search terms" used in the queries.

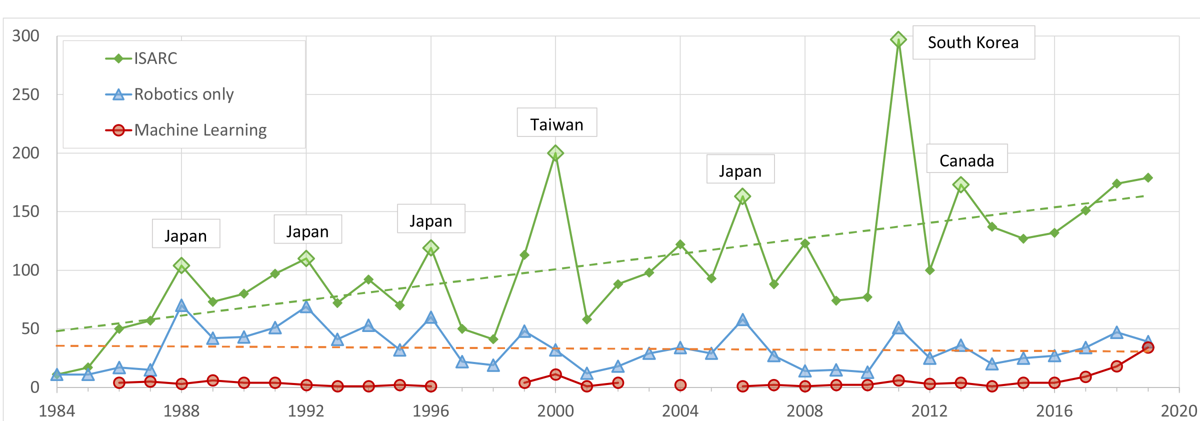

Fig. 10. Publications related to Robotics and Machine Learning presented at ISARC from1984 to 2020.

4.3. Thematic analysis of publications on robotics for construction

This section presents a thematic analysis of the publications presented at ISARC, using some of the terms and categories identified in the previous section. Fig. 10 presents a graph plotting all the papers presented at ISARC, which have been increasing moderately to ~ 150–170 papers presented each year. The number of papers related to robotics only and related to ML are plotted as well. There are symposium editions with a significantly higher number of papers. These peaks correspond to conferences held in Japan, Taiwan, South Korea, and Canada. However, only in the editions held in Japan the number of papers related to robotics increase significantly as well. In the other cases the increase is relatively smaller, and in the case of Taiwan there is a slight decrease. Note that the symposium has been organised in the US eight times; while, in Japan four times and in Germany-three times.

Fig. 10 indicates a huge increase in publications presented at ISARC in the late 1980s reaching ~ 104 publications out of which ~ 70 publications were related to robotics, almost 70 %. After that, the number of publications related to robotics presented at ISARC has been decreasing. Recently, the average is ~ 35 publications per year, representing only about 20 % of the total publications presented at ISARC. Publications related to ML started at ISARC in the mid-1980s; and have remained constant until the mid-2000s, with some exceptions in which no papers were presented. Since 2016, a constant and significant increase in publications related to ML can be observed.

Fig. 11a plots all the publications related to ML presented at ISARC, including papers relating to agent-based modelling and multi-agent systems, which are computational approaches that simulate the interaction of autonomous agents. Papers concerning RL and IL specifically, are identified as well. A massive recent increase in ML publications is evident in this plot that aligns with the massive increase in ML publications in general. Multi-agent systems started in ISARC in the mid-1990s and have remained somewhat constant ever since. However, RL and IL have been hardly addressed in robotics for construction. There are only four instances of papers that address RL directly, and only one instance that addresses IL. All of which are discussed in more detail in the next section.

Fig. 11b presents ISARC papers related to Augmented Reality (AR) and Virtual Reality (VR). The papers that address both terms in conjunction with robotics are highlighted as well. VR papers started in the mid-1990s and AR papers at the beginning of the 2000s; both presenting a significant ongoing increase from the 2010s. However, AR and VR, in conjunction with robotics, has been explored less. There are only five papers addressing VR and robotics, and five addressing AR and robotics. Papers on VR and robotics include a model of a VR-based approach to programming construction robots [84]; a VR-based method to predict humanoid-robots movements [86]; a robotic excavating system coupled with VR-based simulations [64]; a survey on safety indicators in which robotics and VR impacts to safety are discussed [66]; and an approach to simulate crane movements through a robotic arm in a VR environment [27]). Papers on AR and robotics include a concept for an AR-based robotic construction manager [16]; an AR-equipped indoor inspection robot [109]; a system architecture for an AR-based robotic teleoperation solution [100]; an AR-based workflow for human-robot interaction [62]; and an AR-based system to support collaboration among multiple robots [112].

Fig. 12 presents two charts analysing the papers presented in ISARC related to three sets of terms, i.e.: (i) mobile robots, multi-purpose robots, and multi-robots referred as “3M”; (ii) human-robot interaction referred as “HRI”; and (iii) unmanned aerial vehicles (UAV) and unmanned ground vehicles (UGV) referred as “drones”. Fig. 12a presents a chart illustrating the ratio between publications related to 3 M, HRI, and drones and publications related to robotics in general. Publications on 3 M topics have been somewhat constant; while publications on HRI have been presented from the late 1980s to the late 2000s but have decreased significantly since then. Publications on drones started in the late 1980s as well but have increased considerably starting from the beginning of the 2010s. In the last three editions analysed, papers on drones —on average– account for more than 30 % of all the papers on robotics presented at ISARC; while in 1988 accounted for only ~5 %. Fig. 12b presents a more detailed mapping of papers relating to 3 M, HRI, and drones. In the last three editions analysed there has been a massive increase in publications related to UAVs and UGVs more than doubling the average of the previous years. Only eight papers have been presented on multi-purpose robots. The last one was presented in 1996 and the first one in 1988. Seven papers have been presented on multi-robot systems, the first one in 1994 and then the next ones from 2007 to 2019. Regarding mobile-robots, 49 papers have been presented from 1986 to 2019. Regarding HRI, only three publications have been presented in the last eight years.

5. Characterising research at the intersection of robotics, reinforcement learning, and construction

In this section, a more granular analysis is presented that provides a detailed characterisation of recent research at the intersection of robotics, reinforcement learning, and construction. Seventy-eight publications published from 2015 onwards were selected from the ISARC database and from the Scopus search result on robotics, construction, and reinforcement learning (see details in section 2). All the papers describe research efforts to develop robotics systems for the construction industry, and in some cases, they employ ML and RL approaches.

First, an analysis of the type of use-cases addressed in the publications is presented. The use-cases have been grouped in seven categories (Fig. 13), i.e.: (1) *surveying*, which groups research the uses robotics to determine three-dimensional attributes and relationship in construction elements and its environment. For instance, publications on a robot that marks positions on the ground to install pedestals for free access floors

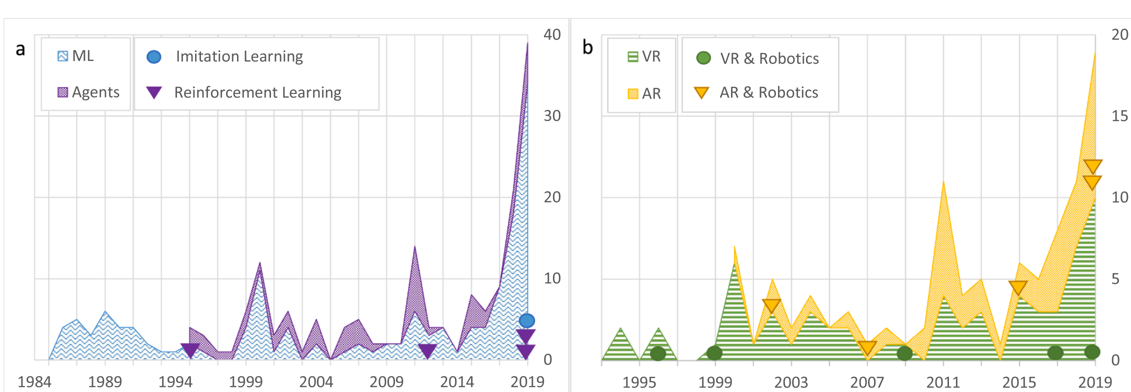

Fig. 11. Publications related to (a) ML and Agent systems; and (b) AR and VR presented at ISARC from1984 to 2020.

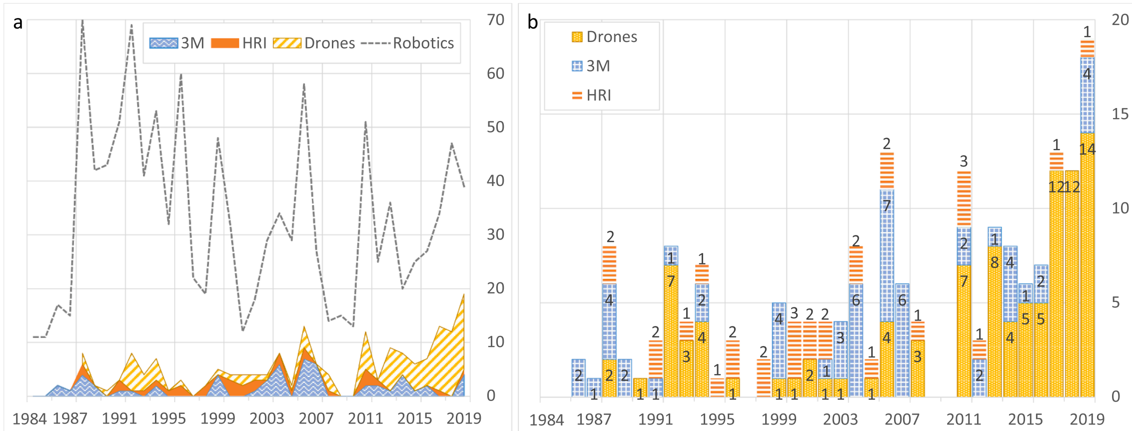

Fig. 12. Publications related to “3M”, “HRI”, and “drones” presented at ISARC from1984 to 2020. (a) Highlights the ratio among the publications related to robotics in general and related to each topic in specific. (b) Presents the number of publications per each topic.

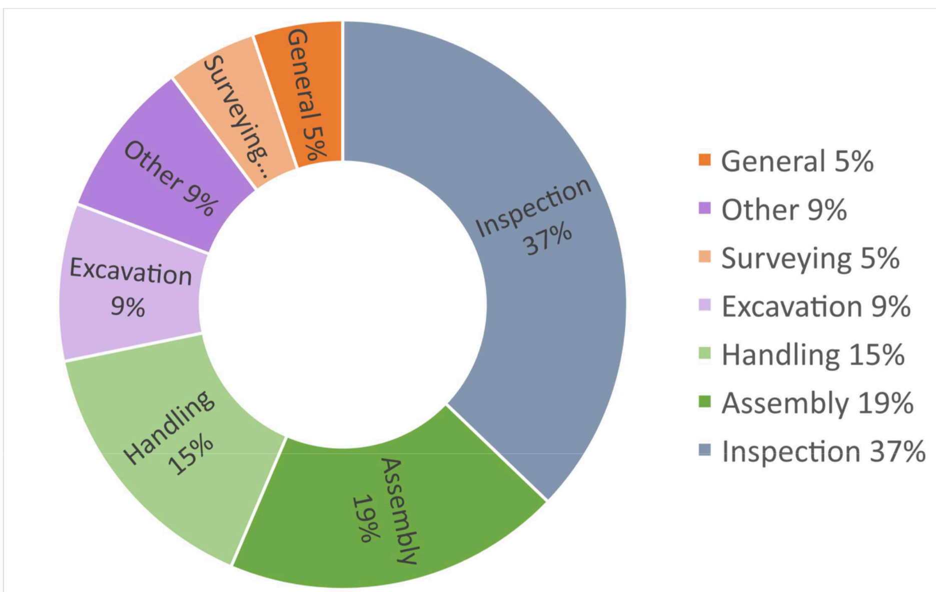

Fig. 13. Distribution of construction use-cases addressed in selected publications.

[103–104]). (2) *Excavation and earth moving*, e.g. publications on approaches that enable using traditional excavation machinery autonomously or remotely with minimum intervention [65], or approaches that require a complete overhaul of the hydraulic systems and the installation of various types of sensors for precise electrical control [45]. (3) *Handling building components*, which group research that enables robotic system to handle and move building components and materials. For instance research that explores the advantages of using UAV support to enable a robotised crane to move prefabricated concrete components from the ground to a higher floor for installation [23]; or research that enables accurate real-time tracking of the positions and poses of objects grasped by robotic arms [52]. (4) *Assembling and installation of building components*, which group papers that deal with robotic control that enable building component installation. Such as research on a robotic arm mounted on a mobile platform to install floor tiles [70], or research on developing a cable-based robotic system to install curtain wall modules. (5) *Monitoring and inspection*, which agglutinates robotic solutions to support environment mapping, progress monitoring, and maintenance inspections. For example, research on a UGV that monitors the quality (density and moisture content) of embankments [55], research exploring the advantages of combining UAVs and UGVs to capture geometric data [57], or the development of robotic systems to carry out condition inspections of tunnel linings [50]. (6) *General construction tasks*, which groups publications presenting robotic systems for various types of tasks commonly carried out in construction sites such as drilling [113], concrete chipping [28], or painting [21]. Lastly, (7) *other*, which agglutinates research closely related to construction; e.g., research that explores low-cost alternatives for accurate robot positioning and navigation such as fiducial markers [80], or the generation of floorplans suitable for robot-assisted living [53].

The analysed publications were grouped in four tiers according to the number of times that a use-case was addressed in the publications (indicated by different colours in Fig. 13). The first tier with the most numerous use-case is monitoring and inspection, representing 37 % of the papers analysed. A probable reason for this is that UGVs, and particularly UAVs, commonly used for inspection are easier and less expensive to acquire and to deploy than other types of robots. For example, robotic arms used in research cost around 50,000 USD [54], heavy-duty UGVs ~ 20,000 USD [81], research-oriented four-wheel-drive rovers range from approximately 5,000 to 12,000 USD, and UAVs could range from 1000 USD [57]to 200 USD [61]. Also, overall inspection is a relatively easier task than the other use-cases. Assembly and handling follow in a second-tier with significantly lower percentages, 19 % and 15 % respectively. These use cases are favoured because they will enable a streamlined workflow from off-site manufacturing to automated installation. Automated excavation and others are in a third tier, both with 9 %. While in tier 4 with the least number of publications are surveying and generic works with 5 % each. Note that surveying is a significantly more complex task than inspection because it requires precise measuring and marking; while inspection has been focusing only on capturing data using sensors and cameras.

Next and analysis of the most used computational approaches used is presented. Fig. 14 shows the distribution of computational approaches used in the selected papers, i.e. (a) *reinforcement learning* (RL), which includes deep reinforcement learning (DRL) and Imitation Learning (IL). These publications are discussed in detail below and outlined in Table 4. (b) *Supervised learning*, which groups papers focusing on computer vision use-cases such as detecting building mechanical components [106], and using CNNS to detect signs of degradation such as cracks [61] and delamination [81]. (c) *Unsupervised learning*, which includes papers that use unsupervised learning as support for RL. For example, the use of autoencoders to encode expert demonstrations to a representation that can be learned easily by the agent [67]. (d) *Evolutionary approaches*, which clusters publications that describe methods to automate certain parts of the robot's behaviour; e.g. the use of optimisation algorithms to plan navigation routes so that robots can evade obstacles [10]. (e) *Control*, which groups traditional robotic control approaches. Note that publications that do not employ any computational method to control robot behaviour are assigned to the "NA" category and that papers that use more than one approaches are considered in all the respective categories. From the analysed publications, the large majority, almost three fourths, employ traditional control approaches (72 %), while other more advanced methods only represent 26 %. From those, 13 % corresponds to RL approaches; traditional RL with 8 %, DRL with 4 %, and IL with 1 %. Supervised learning accounts for 8 %, evolutionary approaches for 4 %, and unsupervised learning for 1 %. This distribution aligns loosely with the keyword analysis presented in Fig. 8, albeit the position of supervised learning and RL are swapped. This difference could be explained because the former analysis is broader, and the assessment here is more specific to construction.

Fig. 15 presents a mapping of 76 papers published since 2015 on robotics in construction according to the level of development of the robotic system and the type of robot used. Note that the same research efforts reported in multiple papers since 2015, e.g., papers reporting progress updates on previous publications, are only considered once. The level of development of the reported robotic systems is categorised in five levels, i.e.: (i) theoretical studies, (ii) simulations, (iii) evaluated in a small-scale restricted environment, (iv) evaluated in a real-scale restricted environment, and (v) evaluated in a construction site semi-restricted environment. The robot type categories are (a) not defined, for generic publications that do not define the type of robot; (b) UAV/UGV for unmanned aerial vehicles and unmanned ground vehicles; (c) arms, for static robotic arms; (d) mobile arms, for robotic arms on mobile platforms or rails; (d) construction machinery, for automated excavators, dump trucks, dozers; and (e) newly developed, for robotic systems specifically developed for the construction industry. The size of the circles depends on the number of publications at the intersection of each category, which is indicated by the number at the centre of the circle. Darker shades indicate a higher level of development.

Publications on UAVs and UGVs are the most numerous accounting for 36 % of the total, followed by publications on static robotic arms with 26 %. At the same time, robotics systems specifically developed for construction represent the 17 %. Most of the research reported evaluations in small-scale restricted environments (34 %), followed by simulations (28 %) and real-scale restricted environments (18 %). Only UAVs/UGVs and static robotic arms were tested on actual construction sites or infrastructure sites. Fig. 15 suggests that research is concentrated in robotic systems that are relatively easy to acquire and to test, namely drones and small-scale robotic arms. Newly developed robotic systems show a homogeneous distribution among all levels of development but

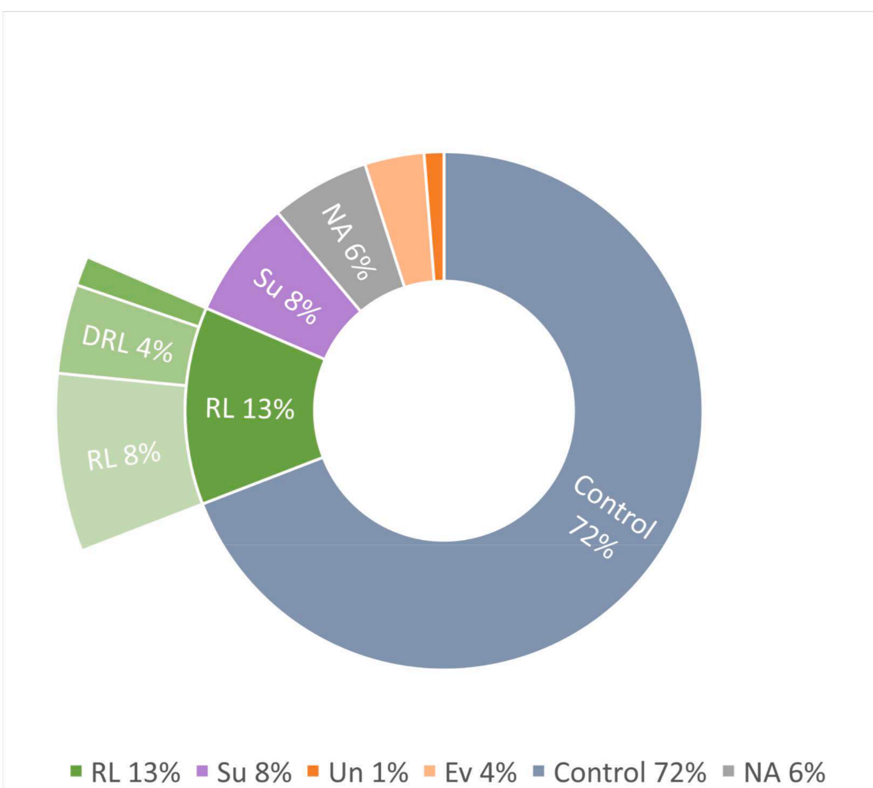

Fig. 14. Distribution of computational approaches addressed in the selected papers. Supervised Learning (Su), Unsupervised Learning (Un), Evolutionary (Ev).

Table 4
Sample of papers addressing Robotics, Construction, and RL. Small-scale restricted environment (SRE). Real-scale restricted environment (RRE).

Year	Description	Use-Case	RL-IL methods	Robot Type	Task type	Development Level	References
1995	Determining gait of a wall-climbing quadruped robot	Inspection	RL > Model-Free > Value-based and Genetic Algorithms	4-legged wall climbing	control	simulation	[24]
2012	Kinematic and dynamic modelling of a mobile robotic arm	Handling	RL is only referred to solve the optimal controls.	mobile arm	control	theory	[29]
2013 2015	Manipulate and transport parts to assemble truss-like structures	Handling	RL > Model-Free > Value-based and a heuristic search algorithm	UAV	planning	SRE	[11,12]
2014	Manipulate and transport parts to assemble truss-like structures	Handling	RL > Model-Free > Value-based and a heuristic search algorithm	multi UAV	planning	simulation	[13]
2017	Assembling block-like parts using magnets	Assembly	RL > Model-Free > Policy-based RL for choosing among subtask that are learned separately. Optimal path planning: PSO.	mobile pusher	planning	simulation	[10]
2018	Dry stacking irregular objects (2D)	Assembly	DRL > Model-Free > Value-based	arm	control	SRE	[72]
2019	Slab stone installation on walls	Assembly	DRL > Model-Free > Value-based	dual-arm	control	simulation	[69]
2019	Automated navigation and obstacle avoidance	Handling	DRL > Model-Free > Value-based	UGV	planning	simulation	[47]
2020	Tile ceiling installation	Assembly	IL > Behavioural Cloning > Model-free Autoencoders to generate easy-to-learn representations	arm	control	simulation	[67]
2020	Placement of object in a discrete environment with obstacles	Object placement	DRL > Model-Free > Value-based	mobile pusher	planning	simulation	[116]
2021	Assembly of timber joints with a robotic arm	Assembly	IL > Behavioural Cloning > Model-free	arm	control	RRE	[4]
Examples of robotic systems specifically designed for construction (no RL or IL approaches are implemented)							
2015	A 6-legged robot prototype for under-bridge inspection using electromagnets	Inspection	NA	6-legged upside-down climbing	control	SRE	[5]
2016	Cable-suspended robot for masonry wall assembly	Assembly	NA	suspended from fixed structure	control	simulation	[20]
2016	Steel-structure climbing robot for bridge inspection using magnets	Inspection	NA	magnetic-based wheeled climbing	control	SRE	[90]
2018	Stacker crane to install prefabricated façade modules	Assembly	NA	modified stacker crane	control	SRE	[51]
2018	Wall climbing marking robot to indicate installation positions	Surveying	NA	vacuum-based wheeled climbing	control	RRE	[58]

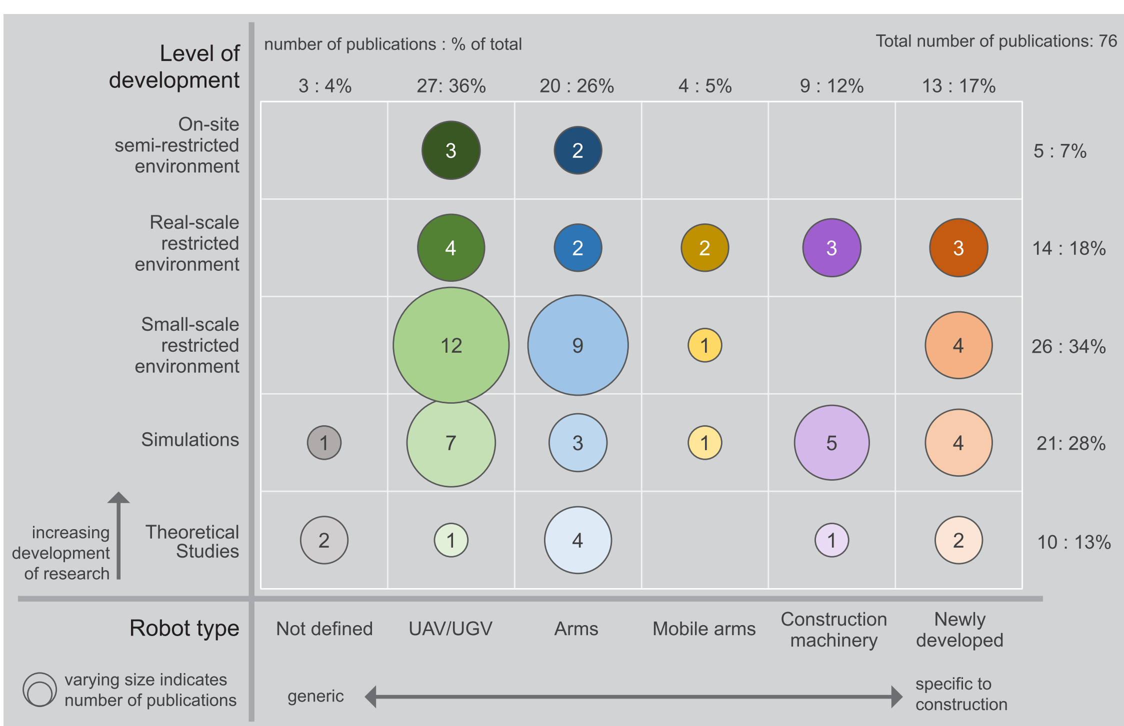

Fig. 15. Diagram mapping publications on robotics for construction according to their level of development and the type of robots used.

have not been tested on real site conditions. Publications on automated construction machinery are slightly fewer than newly developed and have not been tested on sites either. This could be due to the high safety measures that need to be in place to test such systems. Publications on mobile arms are the least numerous, and their respective theoretical studies were not found in the literature.

Table 4 presents selected examples of papers focusing on robotics in construction and using RL or IL approaches. The intention is to provide an overview of the type of research that has been carried out at the intersection of these three fields and to characterise its evolution. The table presents a brief description of the paper; the construction use-case addressed; the type of RL or IL methods used; the type of robot; the task type, i.e., robot control or route planning; and the level of development of the robotic system. Note that relevant papers employing RL approaches published before 2015 have been included in this analysis as well.

The early work at the intersection of these areas was characterised by research that developed the "building blocks" to enable robot automation. An early example was presented by Bull et al. [24], which compared the use of Q-learning, a value-based RL approach, and Genetic Algorithms (GAs) to determine the optimal gait of a wall-climbing quadruped robot. The proposed robotic system used four vacuum feet to climb building walls to support inspection and maintenance tasks. In another example, Chu [29]developed a state-space model for a mobile robotic arm that would be more useful for construction tasks than a static one. Because it would be able to move to various working areas and would not need a structured environment. The authors developed a 9 degrees of freedom (DOF) kinematic and dynamic model, 6 DOF to account for the robotic arm movements, and 3 DOF to account for translations and rotations of the mobile platform. This state-space model is a prerequisite to learn optimal arm movements using RL.

Then, research focused on enabling UAVs and UGVs to handle and transport building components among several locations. For instance, [12,11] presented a type of RL approach to determine a set of predefined actions to enable a UAV to pick and transport components to build truss-like structures. Experiments were carried out in a small-scale restricted environment. The authors claim that the approach also considers assembly, but the components are joined using magnets, which is very different from a real-life scenario. Thus, in Table 4, these studies are regarded as handling of construction components only and not for assembly. The approach was extended to account for multiple robots as well, but in this case, only simulations were presented Barros dos Santos, Nascimento and Givigi [13]. Barros dos Santos et al. [10] also presented an approach to use a mobile robot to assemble magnetic block-like structures. The approach used Particle Swarm Optimisation (PSO) to find an optimal path and an RL model-free policy-based approach to select from pre-learned subtasks to assemble four different types of structures.

More recently, neural networks have been included in the RL approaches. For example, Hu and Wang [47] presented a DRL approach to enable a UGV to navigate and avoid obstacles without precise location data. The approach is a 2D simplification that uses two sources of data as input, i.e., top-view images of the site and navigation data from the UGV sensors. Combining both types of input proved useful in a very simple environment. Note that in these examples, the action-state spaces and the environments' complexity have been simplified significantly, so most of them cannot be deployed directly in real-world scenarios.

Research has also focused on manipulating components that are less generic and that correspond better to actual building components. Liu et al. [72] proposed a value-based DRL approach to build dry stacked walls using a robotic arm. In this case, the task was also simplified by focusing only on 2D irregular objects. The DRL approach was evaluated in a small-scale restricted environment using a static robotic arm. Another example is presented by D. Liu et al. [69], in which a value-based DRL approach was employed to install stone slabs using a dual-arm robot, only results of the simulated environment were presented. Lastly, simulated implementations of DRL approaches have been also tested to enable the collaboration of agents to achieve a single task, for instance for simplified relocation of objects in discrete environments [116].

IL has been leveraged as well. Liang et al. [67] presented an IL approach for installing ceiling tiles using a robotic arm. The approach used a set of real-life videos, depicting a human installing the tiles, and computer-generated videos to demonstrate the optimal pose of the tile for installation. Autoencoders were used to encode the pose information from the demonstration videos into a representation that the agent could learn more easily; and then, behavioural cloning was used to learn the arm movements required to achieve the demonstrated tile's pose. The approach was evaluated using an industrial robotic arm emulator. Apolinarska et al. [4] presented an IL approach for assembling timber joints that combines human demonstrations in a virtual simulation environment. In this case the human demonstration is recorded in a simulation environment, in which the human uses a game controller to perform the act in the simulation environment. Then, all the training is carried out in the simulation environment, and then it is transferred into the physical robot. The authors note that their approach can generalise to real-world scenarios that have not been learned during training, which is a benefit for the intrinsic variability of construction processes.

Regarding robotic systems designed explicitly for construction (Table 4), Arai et al. [5] presented a prototype of a 6-legged climbing robot to inspect the underside of steel bridges. The feet were equipped with electromagnets, thus enabling walking on vertical surfaces and upside-down walking. Small-scale experiments were conducted in restricted laboratory conditions. Pham et al. [90] also presented a steel-structure climbing robot, but in this case, the robot uses wheels with permanent magnets for moving on steel surfaces. The purpose is to inspect the condition of steel elements in bridges. A prototype was tested in a small-scale restricted environment. In a similar example, Kitahara et al. [58] presented a vacuum-based wall-climbing robot. In this case, the robot would climb concrete surfaces and paint markings to indicate installation positions of equipment, which is usually done manually using surveying instruments. The robot was tested on a real scale restricted scenario.

Bruckmann et al. [20] presented simulations of a robot suspended by cables that could assembly block structures. In this case, four poles demarcate the robot's rectangular area of action. Cables attached to the poles suspend the robot over the area of action, and rotors move the cables enabling movement in three dimensions. The intention was that the robot could pick blocks from one location, within the demarcated area of action, and assembly structures, e.g., a wall. Iturralde and Bock [51] presented a robotic system to automate the installation process of prefabricated façade modules. The robotic system is based on a vertical stacker crane, which would be located along a side of a building. The crane would transport the modules from the ground to their correct position on the building. The system was tested using a static robotic arm in lab conditions. Note that these last two robotic solutions require a semi-structured environment, as deployment would require the installation of additional components to restrict the area of action of the robots.

6. Challenges for deploying RL in robotics applications

RL and IL have the potential to widen the use of robotics into more complex and less restricted scenarios and increase the complexity of tasks that robots can perform; for example, enabling human-like robot hand manipulation [3]). Nevertheless, there are still many and very significant challenges. In practice, RL and IL for robotics are only feasible for relatively simple tasks in very restricted scenarios due to a myriad of limiting factors [59,102], which can be categorised loosely into real-world and algorithmic challenges.

6.1. Real-world challenges

(i) *High-dimensional spaces*. Robotics demands high-dimensional state-action spaces that require extremely large amounts of data and computation to process. Thus, a vast number of sample action-state interactions are required to achieve reasonable learning levels. (ii) *Expensive real-world learning*. Robot exploration in the real-world is very expensive compared with simulated environments. It requires human supervision and careful maintenance as well as higher operational expenses. (iii) *Physical-virtual disparities*. Suboptimal policies can be learned due to differences between the simulated environments and the real-world situation, such as: time discretisation variations between simulations and robots sampling frequencies, and real-world delays in sensing and actuation not accounted for in simulations. (iv) *Dynamic environments*. Real-life environments might change constantly. Even in restricted environments, different light conditions and temperatures could change through the day and seasonally affecting robot performance. Also, robots wear, and their performance degrade. Both changes could potentially hinder the performance of the learned policies. (v) *Onerous real-time requirements*. Most of the current RL and IL approaches use large deep neural networks in various ways. When the systems are deployed, these networks will require special high-processing hardware to control the robots in real-time. (vi) *Higher risks*. Compared with ML approaches in simulated environments, the errors made by robots while interacting in the real-world environment have dramatically larger consequences. A single false output might lead to serious accidents. These higher risks increase the development and implementation costs considerably.

6.2. Algorithmic challenges

(i) *Under-modelling and model uncertainty*. Developing a sufficiently accurate model of the robot and its environment is a very challenging task. It is also difficult to find the most effective balance between a detailed but slow model and a rough but fast one. Moreover, model uncertainties are difficult to identify and thus to include in the models. (ii) *Defining effective rewards*. While defining a reward function is significantly easier than defining a task explicitly, it is still very complicated to define a reward function that leads to a desired robot behaviour. RL systems are infamous for exploiting reward functions in unanticipated manners. IL approaches that reconstruct the reward function from expert demonstrations do not require to specify the reward function manually; however, the optimal or intended rewards are not always easily achieved either. (iii) *Algorithms' low stability and robustness*. Compared with other ML methods, RL and IL are relatively unstable and sensitive to minor deviations in configurations and parameter tuning [46]. (iv) *Low generalisation*. Most of the RL and IL approaches perform satisfactorily for the tasks that were trained for, but it is difficult to leverage those learned behaviours for other tasks.

7. Characterising the construction site requirements for deploying RL-based robotics

This section presents a characterisation of the typical circumstances in construction sites to facilitate understanding of the level of technical challenges facing the implementation of robotics for construction. The major technical challenge in robotics for construction resides in the nature of construction sites and construction tasks. Table 5 presents a list of the main attributes characterising construction sites and tasks; and it compares them with manufacturing shop floors and tasks, a sector in which robotics have been implemented widely.

Construction sites are highly-unstructured environments, in which various crews work on different activities sharing the same space and at the same time. Construction sites represent very large action-state spaces, in which very many different actions can happen that could change the environment in a multitude of unexpected manners.

Table 5
Environment and tasks characteristics for manufacturing and construction.

Environment characteristics ***Manufacturing shop floor***	***Construction site***
- Structured environment	- Highly unstructured environment
- Large action-state spaces	- Very large action-state spaces
- Constant or low changing environments	- Fluctuating (constantly changing environments)
- Easy to constrain	- Difficult to constrain
- Low level of human interaction	- High level of human interaction
- Mid difficulty testing in real-world scenarios	- High difficulty testing in real-world scenarios
- Low variations among different shop floors	- Large variations among different sites
Tasks characteristics	
Manufacturing tasks	***Construction tasks***
- Mid-complexity	- High-complexity
- Low diversity	- Large diversity
- Low interdependency	- Large interdependency
- High compartmentalisation	- Low compartmentalisation
- Difficult to capture expert knowledge	- Difficult to capture expert knowledge
- Existing environment and agent models	- Lack of existing environment and agent models
- Existing simulation environments	- Lack of existing simulation environments

Moreover, construction sites are in constant change affecting the way actions can and should be carried out. Compared with shop floors, construction sites are difficult to constrain or to designate specific areas for robotic work. Also, in construction sites there is a higher level of human interaction than in shop floors, which robotics systems will need to address effectively because safety is a high priority. Testing robotic systems in floor shops and construction sites is expensive in both cases; however, it is more difficult in construction sites, and the disruptions will be higher due to their highly-unstructured nature. Lastly, there are larger topological variations among construction sites than among manufacturing shop floors, which makes more difficult the implementation of generic robotic systems.

Although some construction tasks are somewhat similar, they are very different from the repetitive tasks usually carried out on shop floors. Many construction tasks are repetitive, but the cycles are more complex than in manufacturing and are not identical [36]). Even in the most homogeneous and repetitive tasks, there are subtle differences (component size variations, large tolerances) that affect the action significantly, especially when considered from a robot's perspective [67]. Furthermore, construction tasks are usually complex, there is a broad diversity among tasks, and they are mutually dependent. It is difficult to break down tasks into subtasks and compartmentalise them. For both cases, it is difficult to capture demonstrations from experts; however, for manufacturing activities, there are existing environment models, agent models, and simulation environments that facilitate RL-based robotics development, which are not available for construction sites.

Table 6 lists the RL challenges discussed in the previous subsection and indicates the level of impact that those challenges have for three use-cases, i.e. (a) computer-based solutions, (b) robotics for manufacturing, and (c) robotics for construction. Computer-based solutions refer to solutions that do not require hardware interfaces, such as optimal scheduling, e.g. [8]) and automated gameplay, e.g. [96], for which RL approaches have proven very useful. For each challenge, a 3-level rating has been assigned for the respective use-case. Computer-based solutions are the least affected by the challenges as the real-world challenges do not apply or have limited impact. The algorithmic challenges are also less impactful since the environments are usually smaller, can be modelled fully, and rewards are easier to define. Whereas, robotics for manufacturing and for construction are significantly more affected. Robotics for construction can be considered as a very hard problem for RL and IL approaches, particularly concerning the real-world challenges due to the highly unstructured and complex nature

Table 6
Comparison of the significance levels of the RL challenges for three different use cases.

RL and IL Challenges	Computer-based Solutions	Robotics in Manufacturing	Robotics in Construction
Real-world challenges			
- High-dimensional spaces	★	★★	★★★
- Expensive real-world learning	NA	★★	★★
- Physical-virtual disparities	NA	★	★★★
- Dynamic environments	NA	★	★★★
- Onerous real-time requirements	★	★★★	★★★
- Higher risks	★	★★★	★★★
Algorithmic Challenges			
- Under-modelling and model uncertainty	★	★★	★★★
- Defining effective rewards	★	★★	★★
- Algorithms' low stability and robustness	★★	★★	★★
- Low generalisation	★★	★★	★★

of construction sites. For algorithmic challenges, robotics in construction is on par with robotics for manufacturing; nevertheless, the lack of existing environments and agent models represents still a substantial obstacle.

7.1. Limits on problem size and dimensionality

The most critical challenge of RL and IL approaches for construction robotics is the strong limitations of the problem complexity and dimensionality that these approaches can handle, hindering its application for real-world tasks. Table 7 presents a comparison between a selection of model-based and model-free RL and IL approaches. It presents three metrics: (1) the type and size of environment, this could be discrete or continuous; (2) the size of the state space, i.e., the number of different states in which an agent can be; and (3) the size of the action space, i.e., the number of actions that an agent can take. Given the construction site and activity requirements discussed above and the presented review, all of the examples construction robotics that leverage RL or IL are very simple and highly restricted in terms of environment, state and action spaces. Actions are restricted to single digits, while state spaces and discrete environments a couple of dozens maximum. Note as well that not all authors report these figures in a clear manner, which hinders comparisons between different approaches. Another aspect that should be presented clearly is the processing requirements and training times as the size of the problem increases, as many of these approaches have limitations on processing scalability [7]. Another important aspect is that there are no clear differences between RL and IL approaches in terms of the size and complexity of the problems, for both cases there are strong size limitations. It is the same for model-based and model-free approaches both a strongly limited by the size of the problem; while mode-based approaches enable to reduce the search space this only affects the training time but not the problem size. In sum, all RL and IL approaches are restricted to be applied to simplified versions of real-world tasks, particularly in construction applications in which unstructured and dynamic environments increase the problem size considerably.

7.2. New directions for RL and IL

This section presents new directions in developing further the RL and IL approaches that are relevant to the improvement of robotics for construction.

7.2.1. Multi agent distributed reinforcement learning and collaboration

One way to reduce the searching time for a successful policy is to leverage multiple agents to learn a homogeneous distributed policy. In this case, multiple agents collaborate to learn a common policy without interacting among them. For example, [92] proposed an actor-critic approach in which the policy is learn in a centralised manner, but the policy execution is decentralised. Another important aspect is collaboration. This approach also enables the transfer of the sum of experience of all agents to entire groups of agents or swarms in a way that all together work toward the same goal. The approach is demonstrated in a multi-robot construction problem in which agents collaborate to arrange block elements according to a specified structure. The approach enables to use swarms of different swarm sizes without requiring additional training.

7.2.2. Human-centred collaborative robots

Another critical aspect is human-robot collaboration. In an ideal collaboration scenario, the support that the robot provides to the human has to be proactive as in traditional human-to-human collaborations in which changes in the environment and task requirement are addressed in a seamless manner. In this regard, the robot must contribute to the task with incomplete data about the state space including the human and the environment. The agent must be able to select an appropriate action that will contribute to accomplish the task given variations in the environment and task requirements. A major challenge is how to deal with uncertainties given changing environments and incomplete knowledge of the state space. Ghadirzadeh et al. [39] proposed a supervised-learning approach that can address uncertainties and find an optimal balance between quick and effective actions while minimising potential mistakes. The authors claim that this approach allows for more fluent collaboration avoiding delays when changes in the environment arise.

In general, IL seems to provide larger benefits for human-robot

Table 7
RL and IL approaches comparison of problem complexity and dimensionality.

Task	Approach	environment	state space	action space	Reference
Inspection	RL > Model-Free > Value-based	discrete (20)	4	3	[24]
Assembly	RL > Model-Free > Policy-based	continuous	2	8	[10]
Assembly	DRL > Model-Free > Value-based	discrete (10–16)	3	2	[72]
Assembly	DRL > Model-Free > Value-based	continuous	12	6	[69]
Navigation	DRL > Model-Free > Value-based	continuous	4	4	[47]
Assembly	IL > Behavioural Cloning > Model-free	continuous	13	6	[67]
Object placement	DRL > Model-Free > Value-based	discrete (36)	4	2	[116]
Assembly	IL > Behavioural Cloning > Model-free	continuous	13	6	[4]

interaction as could generalise in better manner given variations and uncertainties common of complex human interactions. For instance, Sasagawa et al., [93] presented an IL method that could execute force adjustments given variability in the human manipulation. And Wang et al. [107] presented an IL approach for coordinated human–robot collaboration that leverages hidden state-space models that enables an agent to select between three different tasks given changing states in the human actions.

7.2.3. Long-horizon planning

Another critical challenge, highly relevant for construction robotics, is how to deal with tasks that require multiple steps and inputs at different and varying times, commonly known as long-horizon planning. A major problem with long-horizon planning tasks is the generation and collection of the massive data, across the entire search space, required so that RL an IL approaches can generalise effectively. In this case, a notable approach is to collect data on demonstration and visited states and the ability to revisit those states and relabelling as demonstrations for policy learning [75].

7.2.4. Cloud federated learning and meta learning

Federated learning enables multiple agents to learn concurrently and then improve their policy learning by acquiring knowledge in the cloud compiled by all the robot's experimentation. For example, Liu et al. [68] proposed an IL approach for cloud federated learning in which multiple sensor data is shared and distributed among agents. The approach improves the efficiency of learning and fuses the different learnt policies locally in each agent.

Meta-learning seeks to improve the task parameters of the learning itself using the data collected in several learning episodes. Thus, meta learning algorithms learn about two aspects of the models, i.e. (a) a policy to complete a task successfully and (b) the change in task parameters when given examples of a new task. These types of approaches enable to learn policies that are adaptable to changing environments and varying tasks in complex and dynamic environments. For example, Song et al. [99] proposed a *meta*-learning approach that enables agent to adapt to changes in environments and states with large noise in state data. Kaushik et al. [56] proposed an approach that defines multiple initial parameters for learning the policy and enables pre-trained agent to select the most effective initial parameters adapting the model to the current scenario to minimise the task completion steps. In a different manner, particle swarm optimisation algorithms have been used to self-tune and optimise agents when affected by external environment changes, parameter variations, and random noise [74,71,73,114].

7.2.5. Gaps between simulation and reality

Similar to other machine learning approaches, RL and IL approaches also suffer from the immense obstacles of collecting relevant real-world data due to the extremely costly and laborious task of gathering sufficient data from a wide enough sample. Thus, approaches to use simulation environments to generate data and to train agents those simulations environments are being developed, e.g., [115]. These types of approaches provide a potential infinite source of data and enables safe training of agents at initial stages of development. Then, the trained agents are transferred to real robots and in some cases additional training is carried out in real environments. However, a major drawback exists, the simulated environments are only an abstraction of the real-world, thus a gap between the simulated and real environment limits the performance of the learned policies, which could degrade over time if the real-world environment change trough time. This gap between simulation and reality is a major challenge for the development of construction robotics as the environments in construction are complex and dynamic, which require of high-fidelity virtual environments for training agents.

8. Discussion

8.1. Key findings

Here is a summary of the most relevant findings presented in this paper grouped into three categories i.e., high-level findings, mid-level findings, and research characteristics.

8.1.1. High-level findings

There are strong indications that research on robotics for construction has not increased considerably since the 1980 s; and, it could have even decreased if measured by the number of publications in the literature.

Research that leverages RL for robotics has been significantly more prominent than IL. This could be explained because it is relatively easier to implement.

DRL has surged in recent years, most probably driven by the massive interest in deep learning approaches to computer vision.

RL and IL main advantages are that there is no need for large labelled datasets and a relatively easy implementation. The disadvantages are that it is very computationally expensive and rewards are difficult to define.

8.1.2. Mid-level findings

The thematic analysis indicates that robotic systems for construction require three essential properties (referred here as 3Ms): i.e.: (i) mobile, they should be able to move to different work areas in the site, (ii) multi-purpose, they should be able to carry out different tasks, and (iv) multi-robot, they should be able to collaborate with other robots to carry out tasks.

Also, essential themes for research on robotics for construction are human-robot interaction, navigation capabilities and computer vision, and AR and VR capabilities.

ML has been used for research on robotics for construction only limitedly, e.g., less than 4 % of all the papers published at ISARC concerning robotics address ML. RL and IL have been employed even far less than that.

Publications addressing the 3Ms has been minimal. Mobility is the feature most addressed primarily through research on UAVs and UGVs.

Research on HRI in the construction context is very limited as well, and the number of publications on this subject has decreased in recent years.

8.1.3. Characteristics of the reviewed research outputs

Inspection is the most employed use-case in research on robotics for construction accounting for almost 40 % of all publications, which is probably the easier use-case to tackle. Thirty-four percent is focused on research concerning the handling and assembly of construction components.

RL and IL approaches have not been widely adopted, as traditional control approaches account for ~ 72 % of the publications.

Research is concentrated in robotic systems that are relatively easy to acquire and to test, namely UAVs, UGVs, and small-scale static robotic arms, accounting for around 62 % of the publications analysed, while only 7 % is evaluated in real-world albeit restricted environments.

Almost 80 % of the publications employ off-the-shelf robotic systems or adapted systems designed initially for other purposes. Moreover, the newly developed robotic systems specific for construction have not been tested on real site conditions.

The complex and dynamic conditions at construction sites coupled with the complex cycles and variable construction tasks represent a very hard problem for RL and IL approaches.

8.2. Outlook: The prospects of deploying RL for construction robotics

8.2.1. Construction is a very hard problem for RL-based robotics

RL approaches have been very successful for specific problems with simple state-action spaces and clear reward functions, e.g., videogames [82]) or boardgames [95]. However, the translation to the real-world is very difficult, and the same level of success has not been achieved in robotics. The computational power required for RL and IL approaches increases massively with low increments in the state-action spaces, which in most cases makes them impossible to implement for robotics in real-life situations [102]. In the case of robotics for construction, this problem is even more significant as the state-action spaces are even larger due to the unstructured nature of the construction sites. The tasks are also more complex; while they are also repetitive, their cycles are more varied. In addition, the cost of training a robotic system using RL is exceptionally high because it will require training in both the simulated environment and the real environment. For example, a recent effort, reported in literature, to train a static robotic arm to grasp different objects required seven arms grasping continuously and in parallel for four months (~800 robot-hours), generating 600,000 sample grasps, and a deep neural network with 1.2 million parameters [54]. These numbers could be significantly higher for training a robotic system with a larger and more complex state-action space required for construction use-cases.

8.2.2. Transcending the low-hanging-fruit approach.

Robotics for construction suffers from the lack of dedicated systems as most of the research and development efforts are adaptations from other industries. As presented in this study, robotic systems designed explicitly for construction use-cases account for only ~ 17 % of all the research efforts reported in literature. Using of-the-shelve robots is the easier and less expensive way to start developing robotic systems for construction; and in many cases, it is the best option to start research on the subject. This is the low-hanging-fruit approach. Nevertheless, this approach has intrinsic limitations that hinder the development of genuinely effective robotic systems for construction. For instance, approaches to robotics used in manufacturing are not optimal for construction, as the structured nature of the environments typical in manufacturing is entirely different in the construction industry. Also, there are large differences in scale and accuracy and precision requirements. For a truly step-change in robotics for construction, dedicated robotic systems must be developed that address the specific requirements of construction; otherwise, advances in robotics for construction will remain limited and at a basic level.

8.2.3. Structuring construction sites vs smarter robots

There are two major ways to advance robotics for construction, both with advantages and disadvantages. One way is to organise and structure construction sites to make them similar to manufacturing shop floors and manufacture building components off-site. In this way, prefabricated building components could be assembled by robots in semi-structured construction sites. The advantages of this way are that robotics approaches already developed for manufacturing could be translated into construction in a relatively easy manner. However, it requires that the design, manufacturing, and assembly of building components become highly integrated. The large investments required to accomplish this integration are only justifiable for very large and repetitive construction projects. The other way entails developing smarter and more flexible robotic systems that can operate in unstructured construction sites. The advantage of this approach is that it will enable the use of robotics for a wide variety of constructions projects from small one-offs to large repetitive ones. The disadvantage is that developing sufficiently smart robotic systems is a very tall order. The authors believe that both approaches are relevant, should be pursued, and will provide substantial benefits to the construction industry as a whole.

8.3. Directions for future research in construction robotics

Concerning further work, the authors suggest four research areas that would aid fostering research in this subject. (i) The formalisation of the construction problem for RL. It is necessary to analyse in detail the construction tasks characteristics, e.g., by structuring construction tasks into different complexity levels [36], thus facilitating the development of workflows appropriate for robotic systems. The identification of opportunities for robotics requires analysis of construction work at the most basic levels and the analysis of its cycles and relationships with higher-level workflows. (ii) Identification of limiting and driving factors. Limiting and driving factors for robotics have been reported in literature, but in a disjointed manner, e.g., [22,31,89]. Integrating knowledge about limiting and driving factors will help to devise effective strategies to advance research in the area, e.g., [33]. (iii) The development of a research roadmap that outlines the principal research avenues alongside outstanding challenges and technology limitations. This roadmap would help researchers to focus their efforts more effectively, and it would facilitate the identification of bottlenecks limiting research progress, e.g. [18],Davila Delgado et al. [32]. For instance, [44] presented framework for the implementation of robotics in construction that define three major aspects to consider i.e., technology, organisation and people. Each of the aspects contains a set of recommended actions that could facilitate the implementation of robotics in construction. This study represents a first step in the development of a roadmap for the adoption of robotics in the construction industry. In addition, (iv) specific evaluation criteria and benchmarks for different types of robots in construction need to be defined in a systematic manner. In this regard, Ma and Hartmann [76] presented seven criteria to evaluate wall-climbing robots and assessed three adhesion techniques and three locomotion techniques most commonly used in wall climbing robots. These types of studies are essential to guide a further research into robotics ensuring correct applicability for construction specific requirements. In this sense, clear declarations on the size of environments, action and state spaces, and processing requirements must be present in all published research efforts. Lastly, (v) it is necessary the development of a maturity model that guides the evaluation of the robotic systems being developed, thus facilitating recording and keeping track of the achieved research progress.

8.3.1. Addressing the high-level challenges for robotics in construction through research

Identifying the high-level factors that limit the adoption of robotics for the construction industry is an essential first step to address the low adoption issue. A few research efforts have identified a number of limiting factors, e.g., [78,22]; while, [31] presented an in-depth study into the limiting factors, outlining four distinct categories, i.e. contractor-side economic factors, client-side economic factors, technical and work-culture factors, and weak business case factors. However, these studies have focused on high-level challenges for industry adoption such as high capital and maintenance costs, low maturity of the technologies, and the fragmented nature of the construction industry. While these studies provide a relevant overview of the problem and enable a rough understanding of the obstacles, more granular studies are required that shed light on how these limitations can be overcome. For example, the limiting factors identified are very different in nature and different strategies are necessary to tackle them.

Here, a brief examination is presented of how some of the limiting factors referred above can be addressed directly through research while taking into consideration the findings of this study. The two limiting factors identified by all the three studies above [78,22,31], are the “fragmented nature of the construction industry” and the “the unskilled workforce and lack of experts in the field”. From these limiting factors, only the latter can be addressed through research. Lack of experts in the field can be addressed by training construction robotics experts instructing them on relevant areas including ML, RL, cyber-physical systems, and robotics in conjunction with construction-related areas such as civil engineering. Three limiting factors were identified by more than one study, i.e., "low R&D budgets in construction", "unproved effectiveness and unclear business case", and "immature technology". In this case, only the latter two can be addressed by research directly. The perception that robotics for construction has not proved its effectiveness and that it is not a mature technology can be tackled by developing solutions for the construction industry from the ground up and not by only adapting what is being developed for other sectors. Most of the current solutions reported in literature give the impression of being literal translations of existing solutions that do not tackle the essential characteristics of the problem. Robotic solutions developed specifically for the construction industry will facilitate demonstrating their effectiveness and maturing the developed technologies. More importantly, note that the highest costs of deploying robotic solutions are the software and the required adequations to the environment, rather than the robot itself [98]. Both of these can be addressed by research through developing low-maintenance software and smarter agents that require fewer physical adequations.

8.4. Limitations of the study

There are two main limitations in the study presented here, i.e.: (1) some publications that are not explicitly related to construction might have been included in the high-level and mid-level analyses. This is due to filtering limitations and ambiguities in search terms, which can potentially distort the actual numbers of publications and ratios presented here. However, the authors believe that this potential deviation is not significant, and that the trends and overall landscape of the state-of-research presented in this study are representative of the actual situation. Moreover, this limitation was mitigated by including publications from the ISARC database, which focuses specifically on construction. (2) The detailed analysis might not have considered some relevant publications on the subject as only a sample of publications were selected. In this case, the authors believe that an exhaustive analysis that assesses all the existing publications on the subject is not indispensable and that a significant sample can provide most of the relevant information required for this study. All in all, note that the statistics, graphs, and analyses presented here are intended to provide only an overview of the state-of-research at the intersection of robotics, reinforcement learning, and construction. Lastly, other studies have recently focused on surveying RL and IL approaches in robotic applications in general e.g., [48,97], which arrive to very similar categorisations and insights as this study in terms of limitations and future research directions.

9. Conclusions

This paper has presented a series of analyses that provide an overview of the state-of-research at the intersection of robotics, reinforcement learning, and construction. Overall, it can be concluded that the amount of research on robotics for construction has not increased significantly as in other fields, and it probably has remained constant since the mid-1980s. RL and IL approaches have not been used widely in robotics for construction, and traditional control methods are still the most used. The intrinsic characteristics of construction, namely the unstructured and dynamic nature of construction sites and the complex task cycles, make construction a tough problem for RL-based robotics. In order to enable a step-change in robotics for construction, it is imperative to develop dedicated hardware and software systems that address the specific requirements of the construction industry. Special attention is essential at the software level because it is the bottleneck that limits the development of smarter and more flexible robotic systems needed for unstructured and dynamic construction sites. In this sense, it is important to consider that the highest cost of implementing a robotic solution is the cost of software at ~ 45 % of the total cost; while ~ 30 % are costs for constraining the environment [98]. Thus, RL and IL are promising approaches to reduce these costs because they represent a reduction in software development costs and a reduction in the need for constrained environments.

This paper provides a very relevant contribution to knowledge. This paper contributes by (i) consolidating, structuring, and summarising research knowledge at the intersection of robotics, reinforcement learning, and construction. (ii) Identifying strengths and weaknesses of RL-based robotics approaches and translating and putting them in context according to the intrinsic characteristics of construction. And, (iii) hinting new possible research avenues and research gaps. The contribution of this paper will help researchers kick-start new research efforts on robotics for construction and boost existing ones. It facilitates the understanding of existing limitations of RL-based robotics and provides high-level information on how to employ them effectively. It also aids to identify what approaches are more useful for different situations and offers a rough idea of the essential requirements to start developing robotic systems for construction.

Declaration of Competing Interest

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

Acknowledgements

The authors would like to gratefully acknowledge the Engineering and Physical Sciences Research Council EPSRC for supporting this study through a research grant ID EP/S031480/1.

References

[1] P. Abbeel, A.Y. Ng, Apprenticeship learning via inverse reinforcement learning, in: Proceedings, Twenty-First International Conference on Machine Learning, ICML 2004, Association for Computing Machinery (ACM), New York, New York, USA, 2004, pp. 1–8, https://doi.org/10.1145/1015330.1015430.

[2] R. Akrour, M. Schoenauer, M. Sebag, Preference-based policy learning, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial intelligence and Lecture Notes in Bioinformatics), Springer, Berlin, Heidelberg, 2011, pp. 12–27, https://doi.org/10.1007/978-3-642-23780-5_11.

[3] O.A.M. Andrychowicz, et al., Learning dexterous in-hand manipulation, Int. J. Rob. Res. 39 (1) (2020) 3–20, https://doi.org/10.1177/0278364919887447.

[4] A.A. Apolinarska, et al., Robotic assembly of timber joints using reinforcement learning, Autom. Constr. 125 (2021), 103569, https://doi.org/10.1016/J.AUTCON.2021.103569.

[5] Arai, T. et al., 2015. Inspection robot in complicated 3D environments. In: 32nd International Symposium on Automation and Robotics in Construction and Mining: Connected to the Future, Proceedings. International Association for Automation and Robotics in Construction I.A.A.R.C). doi:10.22260/isarc2015/0082.

[6] B.D. Argall, et al., A survey of robot learning from demonstration, Rob. Auton. Syst. 57 (5) (2009) 469–483, https://doi.org/10.1016/j.robot.2008.10.024.

[7] Arora, S., Doshi, P., 2018. A survey of inverse reinforcement learning: challenges, methods and progress. Available at: http://arxiv.org/abs/1806.06877 (Accessed: September 22, 2020).

[8] R. Atallah, C. Assi, M. Khabbaz, Deep reinforcement learning-based scheduling for roadside communication networks, in: 2017 15th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks, WiOpt 2017, Institute of Electrical and Electronics Engineers Inc., 2017, https://doi.org/10.23919/WIOPT.2017.7959912.

[9] D. Bai, et al., Improved single shot multibox detector target detection method based on deep feature fusion, Concurr. Computat.: Pract. Exp. 34 (4) (2022) e6614.

[10] S.R. Barros dos Santos, et al., A stochastic learning approach for construction of brick structures with a ground robot, IFAC-PapersOnLine 50 (1) (2017) 5654–5659, https://doi.org/10.1016/j.ifacol.2017.08.1114.

[11] Barros dos Santos, S.R., Givigi, S.N., Nascimento, C.L., 2013. Autonomous construction of structures in a dynamic environment using Reinforcement Learning. In: SysCon 2013 - 7th Annual IEEE International Systems Conference, Proceedings, pp. 452–459. doi: 10.1109/SysCon.2013.6549922.

[12] S.R. Barros dos Santos, S.N. Givigi, C.L. Nascimento, Autonomous construction of multiple structures using learning automata: description and experimental validation, IEEE Syst. J. 9 (4) (2015) 1376–1387, https://doi.org/10.1109/JSYST.2014.2374334.

[13] S.R. Barros dos Santos, C.L. Nascimento, S.N. Givigi, Planning and learning for cooperative construction task with quadrotors, in: 8th Annual IEEE International

Systems Conference, SysCon 2014 - Proceedings, IEEE Computer Society, 2014, pp. 57–64, https://doi.org/10.1109/SysCon.2014.6819236.
[14] Bellemare, M.G., Dabney, W., Munos, R., 2017. A Distributional Perspective on Reinforcement Learning. In: 34th International Conference on Machine Learning, ICML 2017, 1, pp. 693–711.
[15] Billard, A. et al., 2008. Robot programming by demonstration. In: Springer Handbook of Robotics. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 1371–1394. doi:10.1007/978-3-540-30301-5_60.
[16] Bock, T., Ashida, S., 2002. A concept of the robotoid manager with AR. In: Proceedings of the 19th International Symposium on Automation and Robotics in Construction (ISARC). International Association for Automation and Robotics in Construction (IAARC). doi:10.22260/isarc2002/0084.
[17] T. Bock, T. Linner, Construction robots : elementary technologies and single-task construction robots, Cambridge University Press, 2017.
[18] Boje, C. et al., 2020. Towards a semantic construction digital twin: directions for future research. Automation in Construction. Elsevier B.V., p. 103179. doi: 10.1016/j.autcon.2020.103179.
[19] Bric J.D. et al., 2016. Current state of virtual reality simulation in robotic surgery training: a review. Surgical Endoscopy. Springer New York LLC, pp. 2169–2178. doi:10.1007/s00464-015-4517-y.
[20] Bruckmann, T. et al., 2016. Automated Construction of Masonry Buildings using Cable-Driven Parallel Robots. In: doi: 10.22260/ISARC2016/0041.
[21] Bruzl, M. et al., 2016. Optimizing the trajectory of the painting robot. In: ISARC 2016 - 33rd International Symposium on Automation and Robotics in Construction. International Association for Automation and Robotics in Construction I.A.A.R.C), pp. 605–612. doi:10.22260/isarc2016/0073.
[22] J. Buchli, et al., Digital in situ fabrication - challenges and opportunities for robotic in situ fabrication in architecture, construction, and beyond, Cem. Concr. Res. 112 (2018) 66–75, https://doi.org/10.1016/J.CEMCONRES.2018.05.013.
[23] Bulgakov, A. et al., 2019. Dispersed cyber-physical coordination and path planning using unmanned aerial vehicle. In: Proceedings of the 36th International Symposium on Automation and Robotics in Construction, ISARC 2019. International Association for Automation and Robotics in Construction I.A.A.R. C), pp. 730–734. doi:10.22260/isarc2019/0098.
[24] Bull, L. et al., 1995. Adaptive gait acquisition using multi-agent learning for wall climbing robots, in: Proceedings of the 12th International Symposium on Automation and Robotics in Construction (ISARC). International Association for Automation and Robotics in Construction (IAARC). doi:10.22260/isarc1995/0040.
[25] G.C. Burdea, Invited review: the synergy between virtual reality and robotics, IEEE Trans. Robot. Automat. 15 (3) (1999) 400–410, https://doi.org/10.1109/70.768174.
[26] Camilo-Lima, G., 2019 The Growth of AI and Machine Learning in Computer Science Publications, Medium. Available at: https://medium.com/@thegcamilo/the-growth-of-ai-and-machine-learning-in-computer-science-publications-603d75467c38 (Accessed: September 24, 2020).
[27] Chen, K.-L. et al., 2019. Simulating extreme points of crane by robot arm in virtual reality, doi:10.22260/ISARC2019/0166.
[28] Cho, J., Lee, S., 2017. Force-egulated impact control of a reinforced concrete box culvert chipping robot, in: doi: 10.22260/ISARC2017/0063.
[29] Chu, B., 2012. Modeling of a Mobile Manipulator for Redundancy Resolution. In: 2012 Proceedings of the 29th International Symposium of Automation and Robotics in Construction, ISARC 2012. International Association for Automation and Robotics in Construction I.A.A.R.C. doi: 10.22260/isarc2012/0075.
[30] Clark, W.E., Sivan, M., O'Connor, R.J., 2019. Evaluating the use of robotic and virtual reality rehabilitation technologies to improve function in stroke survivors: a narrative review. J. Rehabilitation Assistive Technol. Eng., 6, 205566831986355. doi: 10.1177/2055668319863557.
[31] Davila Delgado, J.M. et al., 2019. Robotics and automated systems in construction: understanding industry-specific challenges for adoption, J. Build. Eng. 26(November). doi:doi.org/10.1016/j.jobe.2019.100868.
[32] J.M. Davila Delgado, L. Oyedele, P. Demian, et al., A research agenda for augmented and virtual reality in architecture, engineering and construction, Adv. Eng. Inf. 45 (2020), 101122, https://doi.org/10.1016/j.aei.2020.101122.
[33] J.M. Davila Delgado, L. Oyedele, T. Beach, et al., Augmented and virtual reality in construction: drivers and limitations for industry adoption, J. Constr. Eng. Manage. 146 (7) (2020) 04020079, https://doi.org/10.1061/(ASCE)CO.1943-7862.0001844.
[34] N.J. van Eck, et al., A comparison of two techniques for bibliometric mapping: Multidimensional scaling and VOS, J. Am. Soc. Inform. Sci. Technol. 61 (12) (2010) 2405–2416, https://doi.org/10.1002/asi.21421.
[35] N.J. van Eck, L. Waltman, Software survey: VOSviewer, a computer program for bibliometric mapping, Scientometrics 84 (2) (2010) 523–538, https://doi.org/10.1007/s11192-009-0146-3.
[36] J.G. Everett, A.H. Slocum, Automation and robotics opportunities: construction versus manufacturing, J. Constr. Eng. Manage. 120 (2) (1994) 443–452, https://doi.org/10.1061/(ASCE)0733-9364(1994)120:2(443).
[37] C. Finn, et al., "A Connection between Generative Adversarial Networks. Inverse Reinforcement Learning, and Energy-Based Models", 2016.
[38] Q. Gao, J. Liu, Z. Ju, Robust real-time hand detection and localization for space human–robot interaction based on deep learning, Neurocomputing 390 (2020) 198–206, https://doi.org/10.1016/J.NEUCOM.2019.02.066.
[39] A. Ghadirzadeh, et al., Human-centered collaborative robots with deep reinforcement learning, IEEE Rob. Autom. Lett. 6 (2) (2021) 566–571, https://doi.org/10.1109/LRA.2020.3047730.
[40] Gong, Z., Zhang, Y., 2018. Temporal spatial inverse semantics for robots communicating with humans. In: Proceedings - IEEE International Conference on Robotics and Automation. Institute of Electrical and Electronics Engineers Inc., pp. 4451–4458. doi:10.1109/ICRA.2018.8460754.
[41] Ha, D., Schmidhuber, J., 2018. Recurrent world models facilitate policy evolution. In: 32nd Conference on Neural Information Processing Systems (NeurIPS 2018), pp. 2450–2462. doi: 10.5281/zenodo.1207631.
[42] T. Haarnoja, et al., Soft actor-critic: off-policy maximum entropy deep reinforcement learning with a stochastic actor, NeurIPS (2018) 1–9.
[43] Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning. Second, Springer series in statistics. Second. doi: 10.1007/978-0-387-84858-7.
[44] Hatoum, M.B., Hala Nassereddine, 2020. Developing a framework for the implementation of robotics in construction enterprises, in: EG-ICE 2020 Workshop on Intelligent Computing in Engineering, pp. 453–462.
[45] R. Heikkilä, et al., Development of an earthmoving machinery autonomous excavator development platform, in: Proceedings of the 36th International Symposium on Automation and Robotics in Construction, ISARC 2019. International Association for Automation and Robotics in Construction I.A.A.R. C), 2019, pp. 1005–1010, https://doi.org/10.22260/isarc2019/0134.
[46] Henderson, P. et al., 2017. Deep Reinforcement Learning that Matters, in: 32nd AAAI Conference on Artificial Intelligence, AAAI, 2018, pp. 3207–3214.
[47] Hu, R., Wang, Z., 2019. A lunar robot obstacle avoidance planning method using deep reinforcement learning for data fusion, in: Proceedings - 2019 Chinese Automation Congress, CAC 2019. Institute of Electrical and Electronics Engineers Inc., pp. 5365–5370. doi:10.1109/CAC48633.2019.8997266.
[48] Hua, J. et al., 2021. Learning for a robot: deep reinforcement learning, imitation learning, transfer learning, Sensors 21 (2021) 1278, 21(4), p. 1278. doi:10.3390/S21041278.
[49] L. Huang, et al., Detection algorithm of safety helmet wearing based on deep learning, Concurr. Computat.: Pract. Exp. 33 (13) (2021) e6234.
[50] D. Inoue, T. Ueno, S. Nakamura, Robotic inspection tests of tunnel lining concrete with crack light-section device on variable guide frame, in: ISARC 2018–35th International Symposium on Automation and Robotics in Construction and International AEC/FM Hackathon: The Future of Building Things. International Association for Automation and Robotics in Construction I.A.A.R.C), 2018, https://doi.org/10.22260/isarc2018/0074.
[51] K. Iturralde, T. Bock, Integrated, automated and robotic process for building upgrading with prefabricated modules, in: ISARC 2018–35th International Symposium on Automation and Robotics in Construction and International AEC/FM Hackathon: The Future of Building Things. International Association for Automation and Robotics in Construction I.A.A.R.C), 2018, https://doi.org/10.22260/isarc2018/0048.
[52] K. Iturralde, T. Kinoshita, T. Bock, Grasped element position recognition and robot pose adjustment during assembly, in: Proceedings of the 36th International Symposium on Automation and Robotics in Construction, ISARC 2019. International Association for Automation and Robotics in Construction I.A.A.R. C), 2019, pp. 461–468, https://doi.org/10.22260/isarc2019/0062.
[53] C. Jiang, A. Mita, Automatic floorplan generation of living space for simulating a life of an elderly resident supported by a mobile robot, in: Proceedings of the 36th International Symposium on Automation and Robotics in Construction, ISARC 2019. International Association for Automation and Robotics in Construction I.A. A.R.C), 2019, pp. 688–695, https://doi.org/10.22260/isarc2019/0092.
[54] Kalashnikov, D. et al., 2018. QT-opt: scalable deep reinforcement learning for vision-based robotic manipulation.
[55] K. Kamiyama, et al., Automatic inspection of embankment by crawler-type mobile robot, in: ISARC 2018–35th International Symposium on Automation and Robotics in Construction and International AEC/FM Hackathon: The Future of Building Things. International Association for Automation and Robotics in Construction I.A.A.R.C), 2018, https://doi.org/10.22260/isarc2018/0097.
[56] R. Kaushik, T. Anne, J.B. Mouret, Fast online adaptation in robotics through meta-learning embeddings of simulated priors, in: IEEE International Conference on Intelligent Robots and Systems, 2020, pp. 5269–5276, https://doi.org/10.1109/IROS45743.2020.9341462.
[57] P. Kim, J. Park, Y.K. Cho, in: As-is geometric data collection and 3D visualization through the collaboration between UAV and UGV, International Association for Automation and Robotics in Construction I.A.A.R.C), 2019, pp. 544–551, https://doi.org/10.22260/isarc2019/0073.
[58] T. Kitahara, K. Satou, J. Onodera, Marking robot in cooperation with three-dimensional measuring instruments, in: ISARC 2018–35th International Symposium on Automation and Robotics in Construction and International AEC/FM Hackathon: The Future of Building Things. International Association for Automation and Robotics in Construction I.A.A.R.C), 2018, https://doi.org/10.22260/isarc2018/0042.
[59] J. Kober, J.A. Bagnell, J. Peters, Reinforcement learning in robotics: a survey, Int. J. Rob. Res. 32 (11) (2013) 1238–1274, https://doi.org/10.1177/0278364913495721.
[60] H. Kretzschmar, et al., Socially compliant mobile robot navigation via inverse reinforcement learning, Int. J. Robot. Res. 35 (11) (2016) 1289–1307, https://doi.org/10.1177/0278364915619772.
[61] Kucuksubasi, F., Sorguc, A.G., 2018. Transfer learning-based crack detection by autonomous UAVs. In: ISARC 2018 - 35th International Symposium on Automation and Robotics in Construction and International AEC/FM Hackathon: The Future of Building Things. International Association for Automation and Robotics in Construction I.A.A.R.C). doi: 10.22260/isarc2018/0081.

[62] Kyjanek, O. et al., 2019. Implementation of an augmented reality AR workflow for human robot collaboration in timber prefabrication. In: Proceedings of the 36th International Symposium on Automation and Robotics in Construction, ISARC 2019. International Association for Automation and Robotics in Construction I.A.A.R.C), pp. 1223–1230. doi:10.22260/isarc2019/0164.
[63] Lachekhab, F., Tadjine, M., 2016. Goal seeking of mobile robot using fuzzy actor critic learning algorithm. In: Proceedings of 2015 7th International Conference on Modelling, Identification and Control, ICMIC 2015. Institute of Electrical and Electronics Engineers Inc. doi: 10.1109/ICMIC.2015.7409370.
[64] Lee, S. et al., 2009. Development of a heuristics-based task planning system for intelligent excavating system, in: 2009 26th International Symposium on Automation and Robotics in Construction, ISARC 2009. International Association for Automation and Robotics in Construction I.A.A.R.C), pp. 307–316. doi: 10.22260/isarc2009/0037.
[65] Lee, Y.S. et al., 2015. Proposal of workspace mapping method for conversionless unmanned excavation system to improve operative performance. In: 32nd International Symposium on Automation and Robotics in Construction and Mining: Connected to the Future, Proceedings. International Association for Automation and Robotics in Construction I.A.A.R.C). doi: 10.22260/isarc2015/0085.
[66] Li, R.Y.M., Leung, T.H., 2017. Leading safety indicators and automated tools in the construction industry. In: ISARC 2017 - Proceedings of the 34th International Symposium on Automation and Robotics in Construction. International Association for Automation and Robotics in Construction I.A.A.R.C), pp. 758–765. doi: 10.22260/isarc2017/0106.
[67] C.J. Liang, V.R. Kamat, C.C. Menassa, Teaching robots to perform quasi-repetitive construction tasks through human demonstration, Autom. Constr. 120 (2020), 103370, https://doi.org/10.1016/j.autcon.2020.103370.
[68] B. Liu, et al., Federated imitation learning: a novel framework for cloud robotic systems with heterogeneous sensor data, IEEE Rob. Autom. Lett. 5 (2) (2020) 3509–3516, https://doi.org/10.1109/LRA.2020.2976321.
[69] D. Liu, J. Cao, X. Lei, Slabstone installation skill acquisition for dual-arm robot based on reinforcement learning, in: IEEE International Conference on Robotics and Biomimetics, ROBIO 2019, Institute of Electrical and Electronics Engineers Inc., 2019, pp. 1298–1305, https://doi.org/10.1109/ROBIO49542.2019.8961805.
[70] Liu, T.Y. et al., 2019. Measuring and positioning system design of robotic floor-tiling. In: Proceedings of the 36th International Symposium on Automation and Robotics in Construction, ISARC 2019. International Association for Automation and Robotics in Construction I.A.A.R.C), pp. 716–721. doi:10.22260/isarc2019/0096.
[71] X. Liu, et al., Genetic algorithm-based trajectory optimization for digital twin robots, Front. Bioeng. Biotechnol. 9 (2022) 1433, https://doi.org/10.3389/FBIOE.2021.793782/BIBTEX.
[72] Y. Liu, et al., Deep Q-Learning for Dry Stacking Irregular Objects, in: IEEE International Conference on Intelligent Robots and Systems, Institute of Electrical and Electronics Engineers Inc., 2018, pp. 1569–1576, https://doi.org/10.1109/IROS.2018.8593619.
[73] Y. Liu, et al., Self-tuning control of manipulator positioning based on fuzzy PID and PSO algorithm, Front. Bioeng. Biotechnol. 9 (2022) 1443, https://doi.org/10.3389/FBIOE.2021.817723/BIBTEX.
[74] Y. Liu, et al., Target localization in local dense mapping using RGBD SLAM and object detection, Concur. Computat.: Pract. Exp. 34 (4) (2022) e6655.
[75] Luo, S., Kasaei, H., Schomaker, L., 2021. Self-imitation learning by planning. In: Proceedings - IEEE International Conference on Robotics and Automation, 2021-May, pp. 4823–4829. doi:10.1109/ICRA48506.2021.9561411.
[76] Ma, L., Hartmann, T., 2020. A review of wall-climbing robots: technical analysis and potential for indoor building inspection. In: Ungureanu, L.-C., Hartmann, T. (eds) EG-ICE 2020 Workshop on Intelligent Computing in Engineering, pp. 463–471. doi:10.14279/depositonce-9977.
[77] G.J. Maeda, et al., Probabilistic movement primitives for coordination of multiple human–robot collaborative tasks, Autonomous Robots 41 (3) (2017) 593–612, https://doi.org/10.1007/s10514-016-9556-2.
[78] R. Mahbub, An investigation into the barriers to the implementation of automation and robotics technologies in the construction industry, Queensland University of Technology, 2008.
[79] Mahmood, A.R. et al., 2018. "Benchmarking Reinforcement Learning Algorithms on Real-World Robots.
[80] Mantha, B.R.K., Garcia de Soto, B., 2019. Designing a reliable fiducial marker network for autonomous indoor robot navigation. In: Proceedings of the 36th International Symposium on Automation and Robotics in Construction, ISARC 2019. International Association for Automation and Robotics in Construction I.A.A.R.C), pp. 74–81. doi:10.22260/isarc2019/0011.
[81] McLaughlin, E., Charron, N. and Narasimhan, S., 2019. Combining deep learning and robotics for automated concrete delamination assessment. In: Proceedings of the 36th International Symposium on Automation and Robotics in Construction, ISARC 2019. International Association for Automation and Robotics in Construction I.A.A.R.C), pp. 485–492. doi:10.22260/isarc2019/0065.
[82] Mnih, V. et al., 2013. Playing Atari with Deep Reinforcement Learning.
[83] Moerland, T.M., Broekens, J., Jonker, C.M., 2020 Model-based reinforcement learning: a survey. In: Proceedings of the International Conference on Electronic Business (ICEB), 2018-December, pp. 421–429.
[84] Navon, R., Retik, A., 1996. A new approach to construction robot programming. In: Proceedings of the 13th International Symposium on Automation and Robotics in Construction. International Association for Automation and Robotics in Construction (IAARC). doi:10.22260/isarc1996/0013.
[85] Ng, A., Rusell, S., 2000. Algorithms for inverse reinforcement learning. In: Proceedings of the Seventeenth International Conference on Machine Learning, pp. 663–670.
[86] O'Brien, J., 2017. A humanoid-robotics method for predicting construction work performance. In: Proceedings of the 16th IAARC/IFAC/IEEE International Symposium on Automation and Robotics in Construction. International Association for Automation and Robotics in Construction (IAARC). doi:10.22260/isarc1999/0036.
[87] Osa, T. et al., 2018. An Algorithmic Perspective on Imitation Learning, Foundations and Trends in Robotics. Now Publishers. doi:10.1561/2300000053.
[88] T. Osa, N. Sugita, M. Mitsuishi, Online trajectory planning and force control for automation of surgical tasks, IEEE Trans. Autom. Sci. Eng. 15 (2) (2018) 675–691, https://doi.org/10.1109/TASE.2017.2676018.
[89] M. Pan, et al., Influencing factors of the future utilisation of construction robots for buildings: a Hong Kong perspective, J. Build. Eng. 30 (2020), 101220, https://doi.org/10.1016/j.jobe.2020.101220.
[90] Pham, N.H. et al., 2016. Visual and 3D mapping for steel bridge inspection using a climbing robot. In: ISARC 2016 - 33rd International Symposium on Automation and Robotics in Construction. International Association for Automation and Robotics in Construction I.A.A.R.C), pp. 141–149. doi:10.22260/isarc2016/0018.
[91] Plaat, A., Kosters, W., Preuss, M., 2020. Model-based deep reinforcement learning for high-dimensional problems, a survey.
[92] G. Sartoretti, et al., Distributed reinforcement learning for multi-robot decentralized collective construction, Springer Proc. Adv. Robot. 9 (2019) 35–49, https://doi.org/10.1007/978-3-030-05816-6_3.
[93] A. Sasagawa, et al., Imitation learning based on bilateral control for human-robot cooperation, IEEE Rob. Autom. Lett. 5 (4) (2020) 6169–6176, https://doi.org/10.1109/LRA.2020.3011353.
[94] Schulman, J. et al., 2015. Trust region policy optimization. In: 32nd International Conference on Machine Learning, ICML 2015. International Machine Learning Society (IMLS), pp. 1889–1897.
[95] Silver, D., Hubert, T., et al., 2017. Mastering chess and shogi by self-play with a general reinforcement learning algorithm.
[96] D. Silver, J. Schrittwieser, et al., Mastering the game of Go without human knowledge, Nature 550 (7676) (2017) 354–359, https://doi.org/10.1038/nature24270.
[97] B. Singh, R. Kumar, V.P. Singh, Reinforcement learning in robotic applications: a comprehensive survey, Artif. Intell. Rev. 55 (2) (2022) 945–990, https://doi.org/10.1007/S10462-021-09997-9/FIGURES/13.
[98] Sirkin, H.L., Zinser, M., Rose, J.R., 2015. The Robotics Revolution: the next great leap in manufacturing.
[99] X. Song, et al., Rapidly adaptable legged robots via evolutionary meta-learning, in: IEEE International Conference on Intelligent Robots and Systems, 2020, pp. 3769–3776, https://doi.org/10.1109/IROS45743.2020.9341571.
[100] D.K. Sukumar, et al., Augmented reality-based tele-robotic system architecture for on-site construction, in: 32nd International Symposium on Automation and Robotics in Construction and Mining: Connected to the Future, Proceedings. International Association for Automation and Robotics in Construction I.A.A.R.C), 2015, https://doi.org/10.22260/isarc2015/0075.
[101] R.S. Sutton, A.G. Barto, Reinforcement Learning: An Introduction, The MIT Press, Second, 2020.
[102] Tai, L. et al., 2016. A survey of deep network solutions for learning control in robotics: from reinforcement to imitation.
[103] T. Tsuruta, K. Miura, M. Miyaguchi, Development of automated mobile marking robot system for free access floor, in: ISARC 2018–35th International Symposium on Automation and Robotics in Construction and International AEC/FM Hackathon: The Future of Building Things. International Association for Automation and Robotics in Construction I.A.A.R.C), 2018, https://doi.org/10.22260/isarc2018/0085.
[104] T. Tsuruta, K. Miura, M. Miyaguchi, Improvement of automated mobile marking robot system using reflectorless three-dimensional measuring instrument, in: Proceedings of the 36th International Symposium on Automation and Robotics in Construction, ISARC 2019. International Association for Automation and Robotics in Construction I.A.A.R.C), 2019, pp. 756–763, https://doi.org/10.22260/isarc2019/0102.
[105] L. Waltman, N.J. van Eck, E.C.M. Noyons, A unified approach to mapping and clustering of bibliometric networks, J. Informet. 4 (4) (2010) 629–635, https://doi.org/10.1016/j.joi.2010.07.002.
[106] K. Wang, J.C.P. Cheng, in: Integrating Hardware-in-the-loop simulation and BIM for planning UAV-based as-built MEP inspection with deep learning techniques, International Association for Automation and Robotics in Construction I.A.A.R.C), 2019, pp. 310–316, https://doi.org/10.22260/isarc2019/0042.
[107] L. Wang, et al., Imitation learning for coordinated human–robot collaboration based on hidden state-space models, Rob. Comput. Integr. Manuf. 76 (2022), 102310, https://doi.org/10.1016/J.RCIM.2021.102310.
[108] Wang, T. et al., 2019. Benchmarking model-based reinforcement learning.
[109] X. Wang, N. Kwok, Inspecting interior built environments using Augmented Reality-based autonomous robot, in: Automation and Robotics in Construction - Proceedings of the 24th International Symposium on Automation and Robotics in Construction. International Association for Automation and Robotics in Construction I.A.A.R.C), 2007, pp. 105–109, https://doi.org/10.22260/isarc2007/0021.
[110] Weber, T. et al., 2017. Imagination-augmented agents for deep reinforcement learning. Adv. Neural Inform. Process. Syst. 5691–5702.

[111] Y. Weng, et al., Enhancement of real-time grasp detection by cascaded deep convolutional neural networks, Concur. Computat.: Pract. Exp. 33 (5) (2021) e5976.
[112] S. Xiang, R. Wang, C. Feng, Towards mobile projective AR for construction co-robots, in: Proceedings of the 36th International Symposium on Automation and Robotics in Construction, ISARC 2019. International Association for Automation and Robotics in Construction I.A.A.R.C), 2019, pp. 1106–1113, https://doi.org/10.22260/isarc2019/0147.
[113] T. Zedin, et al., A method based on C-K Theory for fast STCR development: The case of a drilling robot design, in: ISARC 2017 - Proceedings of the 34th International Symposium on Automation and Robotics in Construction. International Association for Automation and Robotics in Construction I.A.A.R.C), 2017, pp. 464–471, https://doi.org/10.22260/isarc2017/0064.
[114] X. Zhang, et al., Time optimal trajectory planing based on improved sparrow search algorithm, Front. Bioeng. Biotechnol. 10 (2022) 292, https://doi.org/10.3389/FBIOE.2022.852408/BIBTEX.
[115] Zhao, W., Queralta, J.P., Westerlund, T., 2020. Sim-to-real transfer in deep reinforcement learning for robotics: a survey. 2020 IEEE Symposium Series on Computational Intelligence, SSCI 2020 737–744. doi:10.1109/SSCI47803.2020.9308468.
[116] Zhu, A., Pauwels, P., de Vries, B., 2020. Robot construction simulation using deep reinforcement learning. EG-ICE 2020 Workshop on Intelligent Computing in Engineering, Proceedings, 472–480.