

On the Role of Artificial Intelligence Methods in Modern Force-Controlled Manufacturing Robotic Tasks

Vincenzo Petrone^a, Enrico Ferrentino^b and Pasquale Chiacchio^c

Department of Information Engineering, Electrical Engineering and Applied Mathematics (DIEM),
University of Salerno, 84084 Fisciano, Italy
{vipetrone, eferrentino, pchiacchio}@unisa.it

Keywords: Physical Robot-Environment Interaction, Artificial Intelligence, Impedance Control, Reinforcement Learning, Peg-in-Hole.

Abstract: This position paper explores the integration of Artificial Intelligence (AI) into force-controlled robotic tasks within the scope of advanced manufacturing, a cornerstone of Industry 4.0. AI's role in enhancing robotic manipulators – key drivers in the Fourth Industrial Revolution – is rapidly leading to significant innovations in smart manufacturing. The objective of this article is to frame these innovations in practical force-controlled applications – e.g. deburring, polishing, and assembly tasks like peg-in-hole (PiH) – highlighting their necessity for maintaining high-quality production standards. By reporting on recent AI-based methodologies, this article contrasts them and identifies current challenges to be addressed in future research. The analysis concludes with a perspective on future research directions, emphasizing the need for common performance metrics to validate AI techniques, integration of various enhancements for performance optimization, and the importance of validating them in relevant scenarios. These future directions aim to provide consistency with already adopted approaches, so as to be compatible with manufacturing standards, increasing the relevance of AI-driven methods in both academic and industrial contexts.

1 INTRODUCTION

Manufacturing processes are nowadays experiencing the peak of their technological evolution, spurred by rapid advances in industrialization methods currently developing in the ongoing Fourth Industrial Revolution (Xu et al., 2018). These emerging innovations are defining the next generation of industries, leading to the so-called *Industry 4.0* (Yang and Gu, 2021).

One of the pillars of the revolution is Artificial Intelligence (AI), whose application has seen tremendous advancements and increasing popularity in recent years. Robotic manipulators, which were already one of the core drivers of the Third Industrial Revolution, are now amongst the technologies that are benefiting the most from AI (Bai et al., 2020). Merging these two powerful technologies is leading to the rise of *advanced manufacturing* (also termed *smart manufacturing*), which constitutes the foundation of *Industry 4.0* (Yang and Gu, 2021).

This position paper presents a discussion on a pe-

culiar sector of robotic tasks, namely *force-controlled* applications. Such tasks are of fundamental practical importance in manufacturing, since they deal with manipulators exerting forces on the working environment, with the objective of, e.g., refining a workpiece or manipulating and assembling objects.

Popular force-controlled tasks encompass, for instance, deburring (Lloyd et al., 2024), polishing (Iskandar et al., 2023), and assembly (Luo et al., 2019). On the one hand, the first two require exerting a specific normal force on the working surface, in order to remove burrs and production inaccuracies of a workpiece, or to smooth and finish the surface itself. On the other hand, the latter consists in the assembly of two or more objects together: the most renowned example in this context is *peg-in-hole* (PiH) (Sørensen et al., 2016). Although different in terms of requirements, all of the aforementioned tasks necessitate to control, directly or indirectly, the forces exchanged between the manipulator and the working environment.

This article discusses some modern techniques proposed by recent literature, focusing on their AI-based counterparts, elaborating on how they tackle

^a <https://orcid.org/0000-0003-4777-1761>

^b <https://orcid.org/0000-0003-0768-8541>

^c <https://orcid.org/0000-0003-3385-8866>

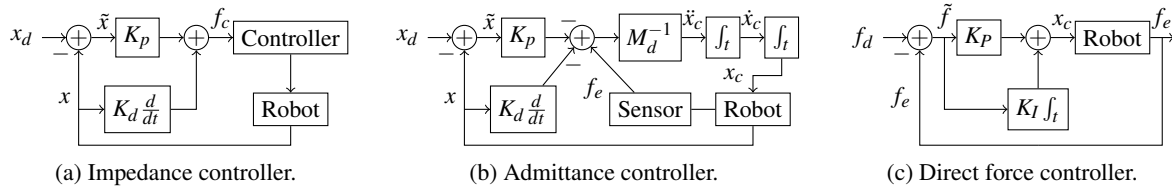


Figure 1: Controllers.

the practical issues arising from the aforementioned force-controlled applications, and highlighting differences among them. To better frame the scope of this article, we stress that we are not proposing a formal and extensive literature review, but we are discussing recent relevant research to claim our position on what the current challenges are in the aforementioned contexts, and propose possible future research directions.

The rest of this paper is organized as follows. Section 2 motivates the need of AI in force-controlled tasks, and presents in detail typical applications in such contexts. Section 3 elaborates on the challenges AI-based methods are currently facing, comparing them with state-of-the-art baselines. Section 4 summarizes the discussion and poses questions to guide future research. Section 5 concludes the paper, stating the main findings of the conducted analysis.

2 MOTIVATION

2.1 Preliminaries

Force control is an essential requirement in a vast number of applications of utmost importance in a broad spectrum of real-world contexts, ranging from industrial (Lloyd et al., 2024) to medical (Tang et al., 2023b) scenarios. For this reason, force control has been one of the major interests in robotics research in the last decades.

With the fundamental objectives of ensuring safety and preserving environment integrity, indirect force control methods, e.g. impedance (Hogan, 1985) and admittance (Newman, 1992) controllers, have been proposed to regulate the manipulator behavior, generalizing pure motion control to interaction scenarios and providing robots with compliant characteristics with respect to the environment they are in contact with.

Usually, an impedance control law (Figure 1a) can be formulated as

$$f_c = K_p \tilde{x} + K_d \dot{\tilde{x}}, \quad (1)$$

where K_p and K_d are stiffness and damping parameters, $\tilde{x} \triangleq x_d - x$ is the task-space error between the setpoint x_d and the actual end-effector (EE) position

x , and f_c is a Cartesian wrench to control via a torque-based controller, which also compensates for the manipulator's internal dynamics, in order to impose a compliant interaction at the EE. On the other hand, admittance control (Figure 1b) takes the form

$$\ddot{x}_c = M_d^{-1} (f_e - K_d \dot{x}_c - K_p \tilde{x}), \quad (2)$$

where M_d is the mass matrix of virtual mass-spring-damper dynamics EE wrenches f_e (usually measured with a force/torque sensor, mounted at the manipulator's flange) are input to. The resulting task-space acceleration \ddot{x}_c is usually commanded to a low-level motion controller.

As an additional performance requirement, industrial tasks usually demand a desired force to be accurately exerted on the working surface: in this case, direct force control (DFC) strategies can be employed to track a reference force (Khatib, 1987). Typical examples of force-related tasks are illustrated in Figure 2, i.e. workpiece deburring (Figure 2a) and surface polishing (Figure 2b). Typically, this is implemented with a PI loop as

$$x_c = K_P \tilde{f} + K_I \int_t \tilde{f} dt, \quad (3)$$

where K_P and K_I are proportional and integral control parameters, respectively, and $\tilde{f} \triangleq f_d - f_e$ is the wrench error between the desired wrench f_d and the actual one (Figure 1c).

2.2 AI-Based Methods

One of the most challenging issues force control algorithms have to face is the inaccuracy and unpredictability of the environment geometry and dynamics. Most of advanced techniques are based on the assumption that the environment force follows a simple linear spring model in the form

$$f_e = K_e (x - x_r), \quad (4)$$

with K_e and x_r being the (unknown) environment stiffness and rest position, respectively. However, this assumption does not always hold, especially at high speeds (Iskandar et al., 2023), and it represents, in general, an approximation, since it is not expected for the environment to behave linearly (Matschek et al.,

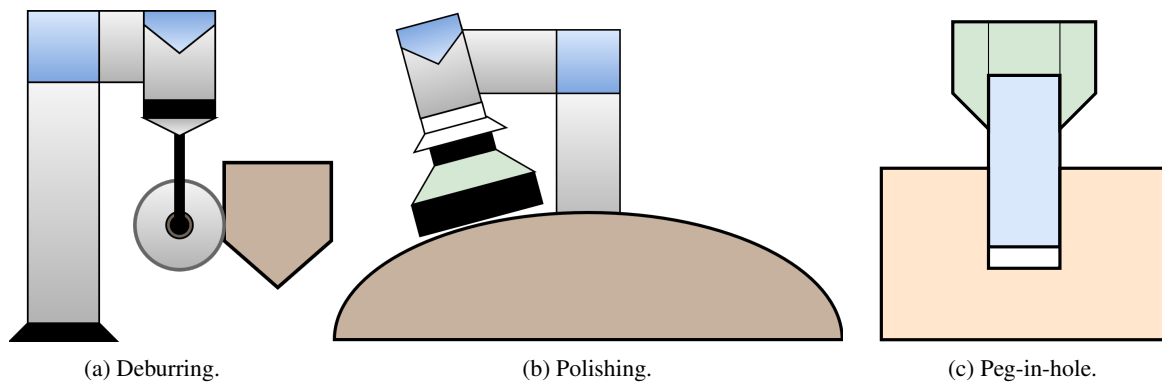


Figure 2: Force-controlled tasks.

2023). This aspect motivates the necessity of employing data-driven strategies, exploiting AI and Machine Learning (ML) to compensate for the inherent difficulties risen from the aforementioned complex scenarios.

Lack of accurate and reliable environment modeling becomes even more problematic in contact-rich assembly tasks, e.g., PiH (Figure 2c) or similar dual setups, such as gear assembly (Luo et al., 2019). These challenging scenarios suffer from additional problems, namely unknown or inaccurate locations of the objects to manipulate (for instance, the hole location or the held peg orientation in PiH), and the complexity of the forces due to frequent contacts, occurring because of the limited tolerance between the parts to assemble. Powerful tools to cope with these inescapable uncertainties and complexities encompass vision-based approaches (Zhang et al., 2023), to estimate hole locations, and Reinforcement Learning (RL), through which control policies are devised from experience and computed according to an objective function to maximize, usually modeling the task goal (Elguea-Aguinaco et al., 2023).

A comprehensive review on PiH control strategies is available in (Jiang et al., 2020). However, this survey does not consider some recent popular methods: indeed, the considered AI- and ML-based methods mostly rely on *learning from demonstrations*, requiring the physical presence of a human operator in the loop to collect data and devise policies to deploy on the robot (Zhang et al., 2021).

The next section will present some of the recent advancements in AI for force-related and contact-rich tasks, and will highlight the practical challenges they are demanded to tackle.

3 CHALLENGES

3.1 Problem Definition

In literature, PiH is considered as a benchmark for force control strategies applied to contact-rich assembly tasks (Jiang et al., 2020). As displayed in Figure 3, it consists of different phases: first, the robot searches for the hole location where the held peg has to be inserted, typically with translational movements (Figure 3a). When the peg engages the hole, the EE rotates to align the former against the latter's walls (Figure 3b). Lastly, the insertion phase actually places the peg into the hole slot (Figure 3c).

For this task, inherent difficulties arise, caused by sub-millimetric tolerance between peg's and hole's dimensions, inaccuracies in estimating the hole's exact location, and complex forces and torques to be managed at the EE. In recent years, RL has been the most popular approach with which these challenges have been faced (Ji et al., 2024), as it usually does not rely on a specific model, instead trying to optimize motion policies according to a tailored reward function, which is constantly updated as data are collected during the task execution.

3.2 Stability and Safety

The first practical objective RL usually struggles to accomplish is *guaranteed asymptotic stability*, i.e. it does not usually provide a theoretical proof formally guaranteeing the RL policy to actually converge towards the objective. In (Khader et al., 2021), this paramount problem is analyzed, thus devising RL policies with guaranteed stability. To achieve this feature, a *variable impedance controller*¹ – orig-

¹A variable impedance controller is similar to the one in (1), where K_p and/or K_d are updated online in an outer optimization loop.

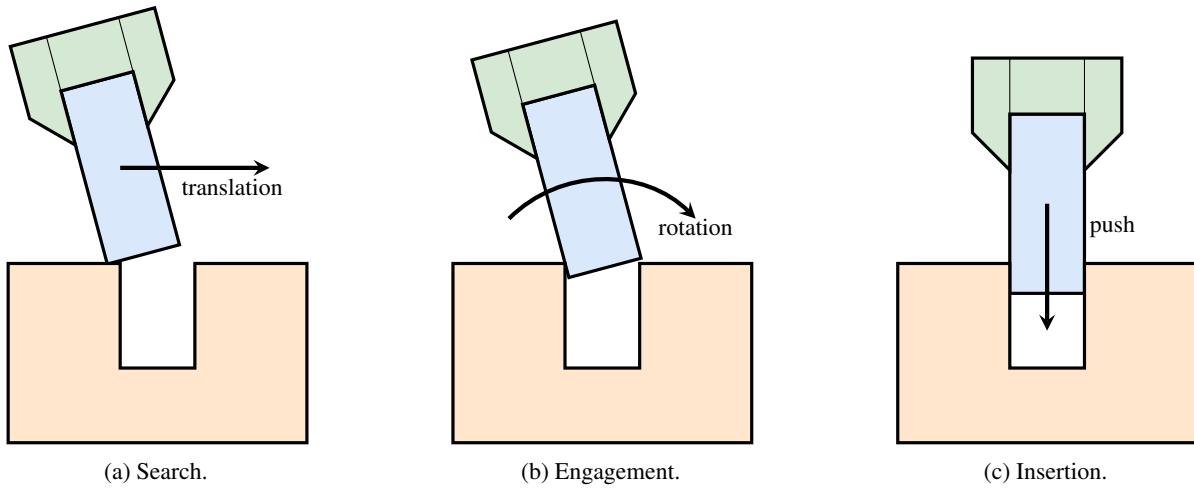


Figure 3: Peg-in-hole task.

inally defined in (Khansari et al., 2014) – is proposed, which is globally asymptotically stable if its matrices are symmetric positive definite (SPD): so, (Khader et al., 2021) proposes to generate them from Wishart distributions. Interestingly, in the PiH experiment, (Khader et al., 2021) clarifies that the stiffness and damping matrices of the variable impedance controller defined in (Khansari et al., 2014) are initialized as, *but not constrained to*, diagonal matrices, hence non-diagonal matrices are computed by the RL policy when solving the task. This is in contrast with classical methods, as K_p and K_d in (1) are usually chosen as diagonal matrices.

The same issue is considered with another approach in (Zhang et al., 2024), which proposes a variable-stiffness impedance controller, where the stiffness matrix K_p and the task-space reference x_d in (1) are computed with two RL policies. This control scheme is applied to the contact-rich cable routing manipulation task, where the robot has to explore a path constrained by walls, whose location is unknown, hence the robot can only rely on contact

forces to have a feedback on the environment. The two policies are called “task policy” and “recovery policy”: the former is used to accomplish the task (i.e., reaching the goal at the end of the maze), and the latter is used to recover from an *unsafe* state-action pair (see Figure 4). It is noteworthy to specify that stability is not formally guaranteed, but safety is promoted with the concept of “risk learning”: indeed, a “safety critic” network is trained to compute the “degree of risk” ϵ of a given task policy action. If ϵ is above a safety threshold α , then the recovery policy action is applied. The “risk learning” phase is performed in a preliminary offline training, where the risk (i.e., the output of the “safety critic” network) depends on the satisfaction of a safety constraint requiring the measured force to stay below a predefined threshold.

In order to stress how the aspect of safety is a strict requirement in these tasks, it is worth pointing out that, whenever it is not explicitly solved with tailored strategies, it is common for researchers to clip the RL policy action in a limited bound (Pozzi et al., 2023). For instance, this guideline is followed by (Hou et al., 2022), which performs the multiple PiH task updating stiffness and damping parameters with a fuzzy-logic controller, and selecting the optimal EE control action making use of DQN (Mnih et al., 2013) and DDPG (Lillicrap et al., 2016) RL frameworks, limiting the action of the latter in a bounded range.

3.3 Optimization Strategies

Although PiH is one of the most common assembly tasks, there is still no consolidated methodology to approach it. Indeed, even considering the spectrum of RL-based techniques, various diverse strategies have

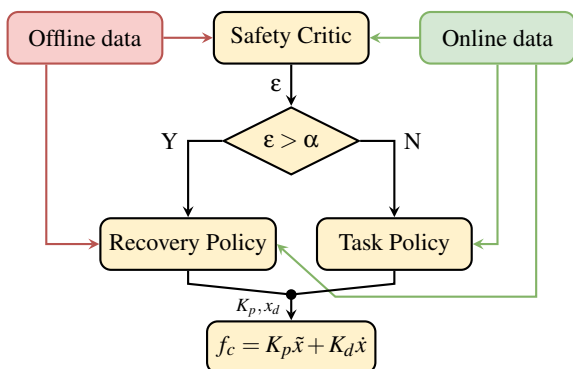


Figure 4: Safe Reinforcement Learning with Variable Impedance Control (SRL-VIC) used in (Zhang et al., 2024).

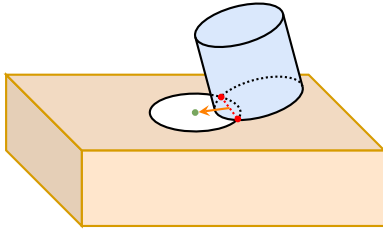


Figure 5: PiH approach used in (Unten et al., 2023): the contact points (in red) are estimated, and a translational motion (in orange) is planned towards the goal (in green).

been proposed, both in terms of *actions* (i.e., the output of the RL policy) and optimization algorithms used to devise the policy itself.

Indeed, the already referenced works choose the action semantics to be either K_p (Khader et al., 2021) or x_d (Hou et al., 2022) or both (Zhang et al., 2024). However, other approaches are possible: for instance, (Ji et al., 2024) solves the task by separating the phases of hole searching and hole insertion (see Figure 3) and optimizing two different configurations of *non-diagonal* (similarly to (Khader et al., 2021)) proportional matrices K_p in (3) through DDPG.

All the aforementioned works rely on model-free frameworks. Instead, (Luo et al., 2019) tackles an assembly task with a model-based RL approach, i.e., iLQG (Todorov and Li, 2005), with which it is possible to compute an optimal policy in closed form, outputting the impedance Cartesian wrench f_c directly.

Lastly, it is worth mentioning a recent novel approach that actually falls outside the RL realm. In fact, (Unten et al., 2023) exploits force-related information to devise a “motion planning” approach to effectively solve the PiH problem. In particular, given force/torque measurements and known peg geometry, the contact points at which the peg enters in contact with the hole walls are estimated. Then, from the two contact points the direction along which the peg must be moved to precisely match the hole is computed as the line connecting the midpoint and the hole center (see Figure 5). Currently, it has never been assessed whether this solution is advantageous compared to RL: as will be discussed in Section 4, we believe this aspect is a prospect to be analyzed in future research.

3.4 Reward Formulation

As mentioned in Section 3.1, formulating a significant reward is fundamental for a RL approach to succeed. Indeed, the reward function R should coherently express the goal of the task, so as to drive the optimization algorithm to yield the optimal policy.

In RL-based PiH, the most popular rewards are

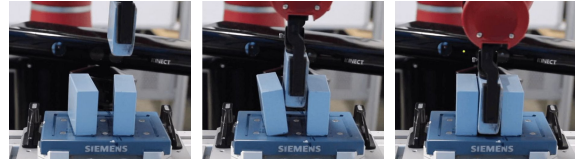


Figure 6: Residual RL approach (Johannink et al., 2019).

Euclidean distance to goal position x_g , in the form

$$R_g = -\|x_g - x\|^2, \quad (5)$$

and the time to complete the task, expressed as

$$R_T = 1 - \frac{k}{T}, \quad (6)$$

where $k \in \mathbb{N}$ is the current time step and $T \in \mathbb{N}$ a parameter denoting the maximum number of steps the task should be completed in.

Additionally, some penalties may be added for the sake of safety, e.g., a penalty on the norm of the action a :

$$R_a = -\|a\|^2, \quad (7)$$

or a safety penalty to avoid the manipulator to exert excessive forces, i.e.,

$$R_f = \begin{cases} -P, & f_e > f_{ih} \\ 0, & \text{otherwise} \end{cases}, \quad (8)$$

where $f_{ih} \in \mathbb{R}^+$ is the safety threshold and $P \in \mathbb{R}^+$ is the penalty.

A unique reward is used by (Johannink et al., 2019), in which the RL action u_r is added as a residual to that of a low-level controller: in particular, with the objective of inserting a peg in a hole between two fixed blocks (see Figure 6), the reward includes a contribution to describe how much the “left” and “right” blocks are tilting from their upright position:

$$R_h = -\alpha_\theta (|\theta_l| + |\theta_r|) - \alpha_\phi (|\phi_l| + |\phi_r|), \quad (9)$$

where θ and ϕ represent the hole blocks’ tilting angles, and $\alpha_\theta, \alpha_\phi \in \mathbb{R}^+$ are hyperparameters.

As evident, the reward does not always follow a particular formulation: researchers typically tend to design them according to certain safety, accuracy or time requirements. In the next section, we will report what the reference works choose as reward, possibly linearly combining them in the form (e.g. summing (5) and (7))

$$R = \alpha_g R_g + \alpha_a R_a, \quad (10)$$

with $\alpha_g, \alpha_a \in \mathbb{R}^+$ being two weights.

3.5 Force-Tracking Tasks

Although PiH does not require force tracking, similar strategies and learning frameworks have been applied to pursue this specific requirement as well, as

Table 1: Differences between PiH approaches. Reward functions sums are intended to be weighted sums.

Reference	Reward	Stability	RL algorithm	Actions	Unified	Simulator
(Khader et al., 2021)	$R_g + R_a$	✓	CEM-like	K_p, K_d	✓	MuJoCo
(Ji et al., 2024)	$R_f + R_T$	✗	DDPG	K_p	✗	PyBullet
(Narang et al., 2022)	R_g	✗	PPO	x_d	✓	IsaacGym
(Hou et al., 2022)	$R_f + R_T$	✗	DQN, DDPG	x_d	✓	—
(Zhang et al., 2024)	$R_g + R_f$	✓	DDPG	K_p, x_d	✓	MuJoCo
(Luo et al., 2019)	R_g	✗	iLQG	f_c	✓	—
(Tang et al., 2023a)	R_g	✗	PPO	x_d	✓	IsaacGym
(Johannink et al., 2019)	$R_g + R_h$	✗	TD3	u_r	✓	MuJoCo

mentioned in Section 1. Such methodologies are developed with the aim of increasing the performance of a standard PI DFC (3) in terms of force-tracking error \tilde{f} . For instance, (Pozzi et al., 2023) uses RL to compute the optimal setpoint x_d for an impedance controller, minimizing \tilde{f} .

Other relevant works employing AI foresee exploiting Neural Networks (NNs) in learning interaction dynamics in variable damping (Huang et al., 2021; Hamedani et al., 2021) and variable stiffness (Liu et al., 2021; Anand et al., 2023) impedance controllers, or selecting the optimal action enhancing that of a low-level DFC (Petrone et al., 2024). It is evident that, in fact, similar AI-enhanced force controllers are employable in various contexts, possibly different than that of PiH.

Nevertheless, in industrial contexts, solely ensuring force-tracking accuracy might not be completely satisfactory: indeed, AI-based methods are not currently coping with practical problems such as tool wear (Lloyd et al., 2024) in deburring applications (Figure 2a) and high-speed motions (Iskandar et al., 2023) in polishing-like tasks (Figure 2b). We claim that both these aspects are of utmost importance in industries, since they aim at maximizing the efficiency, efficacy and quality of the delivered products, thus a future challenge for AI in force control is properly managing these problems, in conjunction with force-tracking.

3.6 Summary

In Table 1 we summarize the major differences in the PiH approaches discussed in the previous sections, in terms of (i) reward; (ii) explicit or implicit stability; (iii) RL algorithm²; (iv) RL policy action; (v) unification of the hole-searching and hole-insertion phases; (vi) simulator. Simulators are paramount in rapidly training RL policies, but their fidelity w.r.t. crucial factors such as system dynamics and real-

²More details on PPO and TD3 can be found in (Schulman et al., 2017) and (Fujimoto et al., 2018)

ism in force/torque measures may heavily influence their success when deployed on the real hardware (Sørensen et al., 2016)³. In this sense, (Narang et al., 2022) proposed a realistic dataset of simulated assets, and (Tang et al., 2023b) introduced specific learning and control strategies to bridge the gap between simulated and real worlds.

Given the evident diversities among recently proposed literature in all the major aspects of the selected approaches, we stress that facing the peculiar PiH task with RL can still be considered an open problem, as a consolidated solution does not currently exist. Our position is that future research should concentrate on limiting the gaps between these strategies, formally comparing their peculiarities or possibly merging their advancements. We will further elaborate this claim in Section 4.

4 FUTURE DIRECTIONS

In the light of the discussions done in Section 3, we now state our position on the topic of empowering manufacturing processes with AI methods. We deem that future research on this subject should concentrate on consolidating the novel technologies that are rapidly emerging in the latest years, both in force-tracking applications and in contact-rich assembly tasks.

To this aim, we suggest to devise common methods to formally compare RL-based techniques, both among them and against standard approaches, defining quantitative metrics according to which existing and novel methodologies shall be validated. For instance, possible performance metrics in PiH might be (i) success rate; (ii) amount of exerted forces on the workpiece; (iii) execution time. In this sense, it is required to define a specific reward shape (see Section 3.4), and to formally state performance require-

³Popular examples are, as also listed in Table 1, MuJoCo (Todorov et al., 2012), IsaacGym (Makoviychuk et al., 2021) and PyBullet (Coumans and Bai, 2016)

ments in terms of, e.g., peg-hole tolerance, so as to fairly compare these methods.

We consider some of the features of the referenced works to be essential in practical scenarios, e.g. guaranteed asymptotic stability in (Khader et al., 2021). Hence, we deem that, in the future, researchers shall make an effort in trying to integrate the various enhancements independently proposed in the relevant recent literature, in order to accomplish various performance requirements, as stated above.

As regards other relevant applications we discussed, namely deburring and polishing, we claim that researchers and practitioners should continue to pursue methods minimizing force-tracking error, as it clearly is the most immediate quantitative index describing the quality of such tasks. However, it is of sheer importance to ensure that novel methods deal with the subjects highlighted in Section 3.5, i.e. performance degradation in highly dynamic scenarios and impact on workpiece machining quality and tool wear. To the best of the authors' knowledge, such topics are not currently covered by AI-driven methods, thus we suggest to invest on this direction, in order to increase the relevance of future works in both academic and industrial contexts.

These considerations are in line with the final goal of providing equivalence with standard and consolidated approaches in terms of *perceived compatibility*. This aspect is indeed a paramount enabler for the use of AI in manufacturing, as demonstrated in (Merhi and Harfouche, 2023), according to which "any innovation is considered compatible with an organization only when it is perceived as consistent with existing business processes, practices, and values".

5 CONCLUSIONS

This paper reported recent advancements on AI methodologies applied to manufacturing robotic tasks. The rationale behind employing these technologies is two-fold. First, in the context of Industry 4.0, they can further optimize manufacturing processes, increasing the production quality, efficiency and throughput. Moreover, they can compensate for inherent limits of classical model-based control methods, as usually happens in challenging force-related and contact-rich applications.

We analyzed issues and objectives these methods are demanded to undertake when applied in real-world industrial scenarios, and analyzed the differences in recent relevant research works on this topic, both on methodological and implementation-related aspects. In conclusion, we claimed our position

on possible directions future research should pursue, in order to accommodate for specific performance requirements, and proposed suggestions researchers shall potentially follow to increase the relevance of future works on both academic and practical level.

REFERENCES

- Anand, A. S., Gravdahl, J. T., and Abu-Dakka, F. J. (2023). Model-based variable impedance learning control for robotic manipulation. *Robot. Auton. Syst.*, 170. Art. no. 104531.
- Bai, C., Dallasega, P., Orzes, G., and Sarkis, J. (2020). Industry 4.0 technologies assessment: A sustainability perspective. *Int. J. Prod. Econ.*, 229. Art. no. 107776.
- Coumans, E. and Bai, Y. (2016). PyBullet, a Python module for physics simulation for games, robotics and machine learning.
- Elguea-Aguinaco, I., Serrano-Muñoz, A., Chrysostomou, D., Inziarte-Hidalgo, I., Bøgh, S., and Arana-Arexolaleiba, N. (2023). A review on reinforcement learning for contact-rich robotic manipulation tasks. *Robot. Computer-Integr. Manufact.*, 81. Art. no. 102517.
- Fujimoto, S., van Hoof, H., and Meger, D. (2018). Addressing Function Approximation Error in Actor-Critic Methods. In *Proc. Int. Conf. Mach. Learn.*, volume 4, pages 2587–2601.
- Hamedani, M. H., Sadeghian, H., Zekri, M., Sheikholeslam, F., and Keshmiri, M. (2021). Intelligent Impedance Control using Wavelet Neural Network for dynamic contact force tracking in unknown varying environments. *Contr. Eng. Pract.*, 113. Art. no. 104840.
- Hogan, N. (1985). Impedance Control: An Approach to Manipulation: Part I - Theory. *J. Dyn. Syst. Meas. Contr.*, 107(1):1–7.
- Hou, Z., Li, Z., Hsu, C., Zhang, K., and Xu, J. (2022). Fuzzy Logic-Driven Variable Time-Scale Prediction-Based Reinforcement Learning for Robotic Multiple Peg-in-Hole Assembly. *IEEE Trans. Automat. Sci. Eng.*, 19(1):218–229.
- Huang, H., Yang, C., and Philip Chen, C. L. (2021). Optimal Robot-Environment Interaction Under Broad Fuzzy Neural Adaptive Control. *IEEE Trans. Cybern.*, 51(7):3824–3835.
- Iskandar, M., Ott, C., Albu-Schäffer, A., Siciliano, B., and Dietrich, A. (2023). Hybrid Force-Impedance Control for Fast End-Effector Motions. *IEEE Robot. Automat. Lett.*, 8(7):3931–3938.
- Ji, Z., Liu, G., Xu, W., Yao, B., Liu, X., and Zhou, Z. (2024). Deep reinforcement learning on variable stiffness compliant control for programming-free robotic assembly in smart manufacturing. *Int. J. Prod. Res.*, 62(19):7073–7095.
- Jiang, J., Huang, Z., Bi, Z., Ma, X., and Yu, G. (2020). State-of-the-Art control strategies for robotic PiH assembly. *Robot. Computer-Integr. Manufact.*, 65. Art. no. 101894.
- Johannink, T., Bahl, S., Nair, A., Luo, J., Kumar, A., Loskyll, M., Ojea, J. A., Solowjow, E., and Levine, S.

- (2019). Residual Reinforcement Learning for Robot Control. In *Proc. IEEE Int. Conf. Robot. Automat.*, pages 6023–6029.
- Khader, S. A., Yin, H., Falco, P., and Kragic, D. (2021). Stability-Guaranteed Reinforcement Learning for Contact-Rich Manipulation. *IEEE Robot. Automat. Lett.*, 6(1):1–8.
- Khansari, M., Kronander, K., and Billard, A. (2014). Modeling robot discrete movements with state-varying stiffness and damping: A framework for integrated motion generation and impedance control. In *Proc. Robot. Sci. Syst.*
- Khatib, O. (1987). A unified approach for motion and force control of robot manipulators: The operational space formulation. *IEEE J. Robot. Automat.*, 3(1):43–53.
- Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D., and Wierstra, D. (2016). Continuous control with deep reinforcement learning. In *Proc. Int. Conf. Learn. Represent.* Art. no. 149803.
- Liu, X., Ge, S. S., Zhao, F., and Mei, X. (2021). Optimized Interaction Control for Robot Manipulator Interacting With Flexible Environment. *IEEE/ASME Trans. Mechatron.*, 26(6):2888–2898.
- Lloyd, S., Irani, R. A., and Ahmadi, M. (2024). Precision robotic deburring with Simultaneous Registration and Machining for improved accuracy, quality, and efficiency. *Robot. Computer-Integr. Manuf.*, 88. Art. no. 102733.
- Luo, J., Solowjow, E., Wen, C., Ojea, J. A., Agogino, A. M., Tamar, A., and Abbeel, P. (2019). Reinforcement Learning on Variable Impedance Controller for High-Precision Robotic Assembly. In *Proc. IEEE Int. Conf. Robot. Automat.*, pages 3080–3087.
- Makoviychuk, V., Wawrzyniak, L., Guo, Y., Lu, M., Storey, K., Macklin, M., Hoeller, D., Rudin, N., Allshire, A., Handa, A., and State, G. (2021). Isaac Gym: High Performance GPU-Based Physics Simulation For Robot Learning. In *Proc. Adv. Neural Inform. Process. Syst.*, volume 1.
- Matschek, J., Bethge, J., and Findeisen, R. (2023). Safe Machine-Learning-Supported Model Predictive Force and Motion Control in Robotics. *IEEE Trans. Contr. Syst. Technol.*, 31(6):2380–2392.
- Merhi, M. I. and Harfouche, A. (2023). Enablers of artificial intelligence adoption and implementation in production systems. *Int. J. Prod. Res.*, 62(15):5457–5471.
- Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., and Riedmiller, M. (2013). Playing Atari with Deep Reinforcement Learning. *arXiv preprint: 1312.5602*.
- Narang, Y., Storey, K., Akinola, I., Macklin, M., Reist, P., Wawrzyniak, L., Guo, Y., Moravanszky, A., State, G., Lu, M., Handa, A., and Fox, D. (2022). Factory: Fast Contact for Robotic Assembly. In *Proc. Robot. Sci. Syst.*
- Newman, W. S. (1992). Stability and Performance Limits of Interaction Controllers. *J. Dyn. Syst. Meas. Contr.*, 114(4):563–570.
- Petrone, V., Puricelli, L., Pozzi, A., Ferrentino, E., Chiacchio, P., Braghin, F., and Roveda, L. (2024). Optimized Residual Action for Interaction Control with Learned Environments. *TechRxiv Preprint: 21905433.v2*.
- Pozzi, A., Puricelli, L., Petrone, V., Ferrentino, E., Chiacchio, P., Braghin, F., and Roveda, L. (2023). Experimental Validation of an Actor-Critic Model Predictive Force Controller for Robot-Environment Interaction Tasks. In *Proc. Int. Conf. Inform. Contr. Automat. Robot.*, volume 1, pages 394–404.
- Schulman, J., Wolski, F., Dhariwal, P., Radford, A., and Klimov, O. (2017). Proximal Policy Optimization Algorithms. *arXiv preprint: 1707.06347*.
- Sørensen, L. C., Buch, J. P., Petersen, H. G., and Kraft, D. (2016). Online Action Learning using Kernel Density Estimation for Quick Discovery of Good Parameters for Peg-in-Hole Insertion. In *Proc. Int. Conf. Inform. Contr. Automat. Robot.*, volume 2, pages 166–177.
- Tang, B., Lin, M. A., Akinola, I. A., Handa, A., Sukhatme, G. S., Ramos, F., Fox, D., and Narang, Y. S. (2023a). IndustReal: Transferring Contact-Rich Assembly Tasks from Simulation to Reality. In *Proc. Robot. Sci. Syst.*
- Tang, Z., Wang, P., Xin, W., Xie, Z., Kan, L., Mohanakrishnan, M., and Laschi, C. (2023b). Meta-Learning-Based Optimal Control for Soft Robotic Manipulators to Interact with Unknown Environments. In *Proc. IEEE Int. Conf. Robot. Automat.*, pages 982–988.
- Todorov, E., Erez, T., and Tassa, Y. (2012). MuJoCo: A physics engine for model-based control. In *Proc. IEEE Int. Conf. Intell. Robots Syst.*, pages 5026–5033.
- Todorov, E. and Li, W. (2005). A generalized iterative LQG method for locally-optimal feedback control of constrained nonlinear stochastic systems. In *Proc. Am. Contr. Conf.*, volume 1, pages 300–306.
- Unten, H., Sakaino, S., and Tsuji, T. (2023). Peg-in-Hole Using Transient Information of Force Response. *IEEE/ASME Trans. Mechatron.*, 28(3):1674–1682.
- Xu, L. D., Xu, E. L., and Li, L. (2018). Industry 4.0: state of the art and future trends. *Int. J. Prod. Res.*, 56(8):2941–2962.
- Yang, F. and Gu, S. (2021). Industry 4.0, a revolution that requires technology and national strategies. *Compl. Intell. Syst.*, 7(3):1311–1325.
- Zhang, H., Solak, G., Lahr, G. J. G., and Ajoudani, A. (2024). SRL-VIC: A Variable Stiffness-Based Safe Reinforcement Learning for Contact-Rich Robotic Tasks. *IEEE Robot. Automat. Lett.*, 9(6):5631–5638.
- Zhang, K., Wang, C., Chen, H., Pan, J., Wang, M. Y., and Zhang, W. (2023). Vision-based Six-Dimensional Peg-in-Hole for Practical Connector Insertion. In *Proc. IEEE Int. Conf. Robot. Automat.*, pages 1771–1777.
- Zhang, X., Sun, L., Kuang, Z., and Tomizuka, M. (2021). Learning Variable Impedance Control via Inverse Reinforcement Learning for Force-Related Tasks. *IEEE Robot. Automat. Lett.*, 6(2):2225–2232.