Proceedings of the 4th International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/2026.39250

Reinforcement Learning in Robotic Surgery: Applications, Challenges, and Future Directions

Jinkai Zhang

Department of Science and Technology Studies, University College London, London, UK
zcrhzhb@ucl.ac.uk

Abstract. Robotic surgery improves accuracy, visualization and motion control. However, most systems still rely on surgeons to make decisions and operate instruments. This article focuses on the improved learning of how to support task level autonomy in robotic surgery. The existing research is classified according to clinical tasks, including tissue operation and cutting, needle guidance and suturing, instrument operation and intraoperative assistance, and postoperative rehabilitation. The methods in the review show that reinforcement learning can learn continuous control strategies, respond to tissue deformation, and adjust auxiliary functions according to specific tasks. Simulation platforms, expert demonstrations, secure reinforcement learning, and migration technology from simulation to reality help reduce risks and data requirements in the training process. However, at present, most of the evidence is still limited to simulated environments, artificial models, in vitro tissues or controlled animal experiments. Clinical transformation is limited by unsafe exploration, changes in the surgical environment, insufficient interpretability and insufficient multi center verification. So intensive learning is more likely to provide short term benefits through clearly defined, security restricted and supervised tasks by surgeons, rather than achieving fully autonomous surgery.

Keywords: Reinforcement learning, robotic surgery, surgical robotics, clinical translation

1. Introduction

Robot assisted minimally invasive surgeries enhance surgical accuracy by magnifying movements, reducing vibrations caused by the human body which is hard to control, and providing enhanced visualization. However, most systems still adopt a remote operation mode, like the surgeon is responsible for interpreting the intraoperative situation, planning each step of the operation, and controlling the movement of the instruments. The autonomy of surgical robots ranges from no autonomy to complete autonomy. The current applications mainly focus on assistance and task execution aspects. So, the actual goal of surgical robots is not to replace surgeons, but to automate the completion of certain clear, repetitive or highly technical secondary tasks, while maintaining human supervision.

Reinforcement Learning is suited for achieving this goal as it enables the learning of control strategies through interaction and rewards [1]. For surgical robots, the observation information can include endoscopic images, instrument positions, the actions correspond to the control of the

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.39250

instruments or cameras; the rewards can reflect accuracy, task completion degree, collision avoidance and tissue safety. This framework enables the robot to make continuous decisions in a deformable and uncertain environment, where fixed trajectories or precise analytical models are often unreliable.

Studies have shown that reinforcement learning can assist robots in performing specific surgical operations, such as enabling the surgical needle to reach the target position more accurately, and automatically adjusting the force during tissue cutting to minimize tissue damage [2, 3]. Now researchers have robots repeatedly practice various tasks such as needle handling, tissue movement, cutting, and suturing control in a computer environment [4, 5]. Existing experiments have shown that under the supervision of doctors, the robots can complete more complex operations such as suturing wounds [6, 7].

However, there is still a significant gap between laboratory performance and clinical application. Reinforcement learning requires a lot of interaction, but in patient treatment, unsafe exploration must not be allowed. If the real life conditions are different from the training conditions, the strategies trained in the computer environment may fail. A practical system must be able to return control to the surgeon before errors cause harm. Inadequately designed reward functions may only encourage task completion without fully reflecting tissue damage or abnormal recovery behaviors. The current cross-patient and cross-hospital validation is still very limited, and the improvement in autonomy also brings issues such as transparency, manual takeover, regulation, and responsibility attribution [8, 9]. Therefore, safe reinforcement learning, expert demonstration, and continuous monitoring by surgeons are necessary conditions for clinical translation [10].

The article will focus on the following aspects. First, in the second part, the article will introduce the methods. This section reviews the application of reinforcement learning in various medical robot tasks. The third part is the discussion section. This part analyzes the main limitations of current research, particularly in terms of safety and data requirements, as well as the poor adaptability to environmental changes. It explores the corresponding directions for the future development of safer, more adaptable, and clinically reliable systems.

2. Method

The reinforcement learning methods in medical robots can be classified according to the tasks performed by the robot. The main differences between various studies lie not only in the algorithms adopted, but also in how the researchers define the robot's observation information, actions, and rewards for specific clinical tasks. Since it is not allowed for the robot to conduct unrestricted trial and error on patients, most systems will first be trained in computer simulation environments, artificial tissue models, or isolated tissues. The following content introduces the existing methods according to four representative task scenarios.

2.1. Direct organization manipulation and cutting

Soft tissues will deform when being grasped, pulled or cut, making organization manipulation a very challenging task. The robot needs to constantly adjust its movements according to the real time changes of the tissues, rather than relying solely on preset fixed motion routes.

Nguyen et al. proposed a method using deep reinforcement learning to maintain appropriate tension during tissue cutting [11]. In their designed system, one instrument performs the cutting along a predetermined route, while the other grasping instrument needs to decide which part of the simulated tissue to hold and in which direction to pull. The reward of this method takes into account

DOI: 10.54254/2755-2721/2026.39250

both the cutting accuracy and the tissue tension, enabling the robot to learn a series of pulling actions. The simulation results show that compared with methods using fixed rules to control tissue tension, this method has better stability.

Abbasi Shahkoo and Abin studied a similar organization manipulation problem and combined continuous action deep reinforcement learning with evolutionary algorithms [3]. The robot learns to determine the position and movement mode of the instrument, while adjusting according to the deformation of the tissue. Continuous actions are very important because the movement of surgical instruments usually cannot be simply divided into several fixed instructions. This study shows that reinforcement learning can control the robot to complete continuous operations in environments prone to deformation.

D'Ettorre et al. proposed a reinforcement learning method based on contextual information for organ manipulation in surgery [12]. The robot controls a flexible surgical guide, moving the organ to make the target area for the surgery more clearly visible. This strategy not only learns a fixed operation situation but also receives contextual information such as the position of the guide and the target position. Therefore, the same method can learn multiple basic operations and adapt to different task conditions.

Iyama et al. implemented automatic counter-traction using the soft actor-critic algorithm [13]. The system reads the three-dimensional point cloud information of the tissue surface and then continuously adjusts the pulling direction and distance of the instrument. While keeping the surgical area clearly visible, the system also needs to avoid excessive deformation of the tissue. These studies show that when the robot needs to control both the shape of the tissue and the surgical field, reinforcement learning has great application value.

2.2. Needle insertion, suturing and fine manipulation

The needle insertion task requires high positional accuracy, and it must also take into account the bending of the needle and the deformation of the tissue. Tan et al. defined flexible needle puncture as a reinforcement learning problem [2]. The robot continuously selects the next adjustment action based on the needle and the target position. This method enables the needle to continuously correct its route during the puncture process, rather than relying solely on the calculated in advance motion trajectory before the surgery. This real-time correction ability is particularly important when the needle bends or the tissue moves.

Keller et al. proposed a robot corneal needle puncture method based on expert demonstration reinforcement learning [14]. They combined the optical coherence tomography technology that can provide depth information with the demonstration data of human operators, and then further improved the control strategy of the robot using the deep deterministic policy gradient method. In experiments using isolated human corneas, this robot reached the target depth more stably than the participating surgeons in the controlled environment set by the study. The significance of this research lies in that the robot had already obtained correct operation demonstrations before starting reinforcement learning, thus reducing random attempts.

In suturing and other fine operations, Xu et al. developed the SurRoL platform [4], and Scheikl et al. developed the LapGym platform [5]. These platforms provide simulation tasks such as needle approach to the target, needle grasping, object transfer, tissue manipulation, and cutting. Researchers can set rewards for requirements such as reaching the target, controlling the direction of the instrument, avoiding collisions, and completing the task, and compare different reinforcement learning algorithms in a reusable unified environment.

DOI: 10.54254/2755-2721/2026.39250

Huang et al. developed DEX, a reinforcement learning method guided by expert demonstrations, using a similar environment [15]. DEX gives higher evaluations to actions similar to expert demonstrations and uses non-parametric regression methods to guide the robot to handle states not present in the demonstration data. This method was tested on ten surgical operation tasks in SurRoL and was transferred to the Da Vinci research platform. The results show that expert demonstrations can improve the efficiency of robot learning and increase the success rate of completing complex tasks.

Shademan et al. and Saeidi et al. respectively achieved autonomous robot suturing and intestinal anastomosis under supervision [6, 7]. These systems are not entirely based on reinforcement learning, but can provide important references for reinforcement learning research: they divide complex surgeries into multiple steps with clear goals, continuously update the surgical plan using sensor information, and retain the supervision of human doctors throughout the process. Compared to controlling the entire surgery with a single reinforcement learning strategy, this task level automation mode is more suitable as a recent development direction.

2.3. Surgical instrument operations and intraoperative assistance

The surgical robot can also provide assistance to the doctor without directly performing the main surgical steps. Typical tasks include moving to the location of the instrument, grasping or repositioning the needle, transferring objects, and maintaining a clear surgical field by pulling on the tissue.

SurRoL and LapGym have designed several of these operations as separate reinforcement learning tasks [4, 5]. The robot's observation information can include the position of the instrument and the camera image, the action represents the movement of the instrument, and the reward is used to evaluate whether the robot has successfully grasped the object, whether it is approaching the target, whether the movement is smooth, and whether it avoids collisions. Training strategies for a task with a clear objective separately can reduce the testing difficulty and allow the doctor to continue to be responsible for the planning and decision-making of the entire surgery.

The contextual reinforcement learning method proposed by D'Ettorre et al. shows that when the target position changes [12], an auxiliary strategy can continue to be used without the need to retrain for each surgical scenario.

Iyama et al.'s anti-pulling system provides another example [13]. This robot does not determine the surgical target or perform tissue separation itself, but by pulling the tissue, it keeps the surgical area clearly visible, allowing the main operator to complete the surgery more effectively.

Huang et al.'s research also shows that when simple rewards cannot fully explain how the robot should approach or grasp an object [15], expert demonstrations can be used to guide the robot. In summary, these methods apply reinforcement learning to automatically complete specific auxiliary tasks with a clear scope, while the overall planning and supervision of the surgery remain the responsibility of the doctor.

2.4. Postoperative rehabilitation and personalized assistance

Reinforcement learning has also been applied to rehabilitation robots. In this case, the goal of the robots is to adjust the level of assistance based on the actual situation of each patient during multiple rehabilitation training sessions. Although these devices do not belong to intraoperative surgical robots, they represent an important application scenario related to postoperative rehabilitation, as they also require learning how to have safe and continuous physical contact with the human body.

DOI: 10.54254/2755-2721/2026.39250

Pareek et al. proposed AR3n, an on-demand assistance control method for robot-assisted writing rehabilitation training [16]. This system uses the soft actor-critic algorithm to adjust the assistance provided by the robot in real time based on the user's motion trajectory errors. Rewards encourage users to accurately complete the actions and penalize unnecessary robot assistance. Therefore, the robot can provide support for the user, but it will not completely replace the user in completing the task. The researchers first trained the system using a virtual patient model and then tested it through simulation experiments and real-person participation experiments.

Zhang et al. used a reinforcement learning method that does not rely on precise human models to establish personalized impedance fields for gait training [17]. The controller can learn at which stage of the walking cycle the power ankle foot device should provide assistance and how much assistance should be provided. So the mechanical assistance provided by the device can change according to the user's walking stage.

Guo et al. also combined reinforcement learning with adaptive admittance controllers for lower limb rehabilitation robots [18]. This method adjusts the interaction parameters between the human and the robot based on the patient's movement performance, thereby providing more personalized and gentle assistance.

These studies indicate that the main value of reinforcement learning in postoperative rehabilitation may be to provide automatically adjustable assistance. When the patient's ability improves, the robot can reduce the assistance; when the patient has difficulty completing the action, the robot can increase the assistance. However, the clinical effectiveness of these methods still needs to be verified through real patient studies, and conclusions cannot be drawn solely based on simulation experiments or small-scale laboratory studies.

3. Discussion

Reinforcement learning has enabled robots to autonomously perform goal-oriented operations. However, there is still a significant gap between the success in the laboratory and its safe application in patient treatment. There are three main challenges, the safety and data efficiency of the learning process, the stability of the robot in the face of environmental changes, and the interpretability and clinical validation of the system.

3.1. Challenges and limitations

3.1.1. Safety and data

Reinforcement learning requires a large number of repeated trials to improve performance and inevitably includes failed actions. This learning approach is acceptable in computer simulation environments, because people need this kind of failed to make a improvement. But it cannot be directly used to actual surgeries because a mistake or incorrect cutting could cause harm to the patient.

The reward function cannot fully represent the quality of clinical operations. For example, the robot may receive a high reward for quickly reaching the target, but in this process, it uses a force that may damage the tissue. The robot may also discover certain behaviors that can obtain high rewards mathematically, but these behaviors do not conform to the real intentions of doctors.

Fan et al. incorporated safety constraints into the reinforcement learning process of surgical tasks [10], Huang et al. used expert demonstrations to reduce meaningless or dangerous attempts [15]. However, these methods still require researchers to determine in advance which states and actions

DOI: 10.54254/2755-2721/2026.39250

are considered clinical dangerous behaviors. The collection of training data is also a problem, as collecting expert demonstrations, establishing real tissue models, and conducting physical robot experiments all require high time and economic costs.

3.1.2. Insufficient adaptability to varying environmental and data distribution shifts

Reinforcement learning strategies are usually trained within a limited range of image appearance, human anatomy, and tissue physical properties. But in real surgeries, lighting conditions may suddenly change, blood, smoke, or other instruments may block the target, and tissue hardness may vary depending on the patient, even changing during the surgery.

So strategies that only use clear images or a single tissue model for training may receive completely different information in actual surgeries from the training data. D'Ettorre et al. added different task conditions during training [12], and Iyama et al. set multiple camera positions and tissue states [13]. These methods can improve the robot's adaptability, but still cannot cover all possible clinical scenarios.

Ou and Tavakoli used reinforcement learning to train a robot to move internal target points in a finite element simulation environment [19], and then transferred the strategy to an actual robot. Their research reflects the current from simulation to reality problem if the tissue deformation, friction, sensor information, or robot movement characteristics in the simulation environment are inconsistent with reality, the transferred strategy may produce incorrect or even dangerous actions.

3.2. Future prospects

3.2.1. Establishing safer reinforcement learning frameworks

The future system should integrate expert demonstrations, and clear safety constraints. Robots can first learn from the recorded expert operations and then continue to improve in the simulation environment instead of directly attempting on patients.

The system should also have an independent safety protection layer. Even if the reinforcement learning strategy requires the robot to perform a certain action, the safety protection layer should be able to prevent the use of excessive force, entering protected areas, or performing actions beyond the movement range of the equipment.

In the future the uncertainty assessment functions should be added. When the robot encounters situations it has not seen before or the system lacks confidence in its own judgment, the robot should stop operating and request assistance from the doctor. The expert demonstration method proposed by Huang et al. [15] and the safety constraint method proposed by Fan et al. [10] provide a foundation for this direction. Future research should not only report the success rate of the task but also report the number of violations of safety constraints, the number of times the doctor took over.

3.2.2. Enhance the robot adaptability to environmental changes

The solution to the problems of environmental changes and data distribution shifts should start from the training stage. The simulation environment should not only present an ideal surgical condition. It should actively change factors such as lighting and camera position.

The digital model established for the patient, combined with multiple information based sensing methods, may also improve the stability of the system. For example, the robot can simultaneously use different ways to avoid complete reliance on data that may be erroneous.

DOI: 10.54254/2755-2721/2026.39250

The simulation to reality transfer method studied by Ou and Tavakoli still needs to be tested on various physical tissues and different robot platforms [19]. Making limited strategic adjustments during the surgery may have certain value, but such adjustments can only be made within a verified safe range. Allowing the robot to conduct unrestricted real time learning during the patient's treatment would bring unacceptable uncertainty.

4. Conclusion

This review indicates that reinforcement learning has meaningful potential to improve the autonomy and consistency of medical robots. Previous studies have applied it to tissue manipulation, cutting, needle guidance, instrument handling, visual-field maintenance, and personalised rehabilitation assistance. These results demonstrate that robots can learn complex control strategies in simulation and perform selected tasks on physical research platforms. Nevertheless, current evidence is mainly based on controlled laboratory conditions and does not yet establish clear benefits for patients. Safety risks, differences between simulated and real environments, limited interpretability, and insufficient clinical validation remain important barriers. Future research should combine expert demonstrations, explicit safety constraints, uncertainty assessment, varied simulation environments, and staged clinical evaluation. Reinforcement learning should therefore be developed as a method for automating well-defined surgical subtasks under continuous human supervision. Its realistic purpose is not to replace surgeons, but to reduce repetitive workload, improve operational consistency, and provide reliable assistance while preserving professional judgment and human responsibility.

References

[1] Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). MIT Press. Cambridge,MA.
[2] Tan, X., Lee, Y., Chng, C. B., et al. (2020). Robot-assisted flexible needle insertion using universal distributional deep reinforcement learning. *International Journal of Computer Assisted Radiology and Surgery, 15*, 341–349.
[3] Abbasi Shahkoo, A., & Abin, A. A. (2023). Deep reinforcement learning in continuous action space for autonomous robotic surgery. *International Journal of Computer Assisted Radiology and Surgery, 18*, 423–431.
[4] Xu, J., Li, B., Lu, B., et al. (2021). SurRoL: An open-source reinforcement learning centered and dVRK compatible platform for surgical robot learning. *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 1821–1828.
[5] Scheikl, P. M., Gyenes, B., Younis, R., et al. (2023). LapGym: An open source framework for reinforcement learning in robot-assisted laparoscopic surgery. *Journal of Machine Learning Research, 24*, 1–42.
[6] Shademan, A., Decker, R. S., Opfermann, J. D., et al. (2016). Supervised autonomous robotic soft tissue surgery. *Science Translational Medicine, 8*, 337ra64.
[7] Saeidi, H., Opfermann, J. D., Kam, M., et al. (2022). Autonomous robotic laparoscopic surgery for intestinal anastomosis. *Science Robotics, 7*, eabj2908.
[8] Yang, G. Z., Cambias, J., Cleary, K., et al. (2017). Medical robotics: Regulatory, ethical, and legal considerations for increasing levels of autonomy. *Science Robotics, 2*.
[9] Attanasio, A., Scaglioni, B., De Momi, E., et al. (2021). Autonomy in surgical robotics. *Annual Review of Control, Robotics, and Autonomous Systems, 4*, 651–679.
[10] Fan, K., Chen, Z., Ferrigno, G., et al. (2024). Learn from safe experience: Safe reinforcement learning for task automation of surgical robot. *IEEE Transactions on Artificial Intelligence, 5*, 3374–3383.
[11] Nguyen, T. T., Nguyen, N. D., Bello, F., & Nahavandi, S. (2019). A new tensioning method using deep reinforcement learning for surgical pattern cutting. *2019 IEEE International Conference on Industrial Technology*, 1339-1344.
[12] D'Ettorre, C., Zirino, S., Dei, N. N., Stilli, A., De Momi, E., & Stoyanov, D. (2022). Learning intraoperative organ manipulation with context-based reinforcement learning. *International Journal of Computer Assisted Radiology*

DOI: 10.54254/2755-2721/2026.39250

and Surgery, 17, 1419-1427.
[13] Iyama, Y., Takahashi, Y., Chen, J., Noda, T., Hara, K., Kobayashi, E., Sakuma, I., & Tomii, N. (2025). Autonomous countertraction for secure field of view in laparoscopic surgery using deep reinforcement learning. *International Journal of Computer Assisted Radiology and Surgery, 20*, 625-633.
[14] Keller, B., Draelos, M., Zhou, K., Qian, R., Kuo, A. N., Konidaris, G., Hauser, K., & Izatt, J. A. (2020). Optical coherence tomography-guided robotic ophthalmic microsurgery via reinforcement learning from demonstration. *IEEE Transactions on Robotics, 36*, 1207-1218.
[15] Huang, T., Chen, K., Li, B., Liu, Y. H., & Dou, Q. (2023). Demonstration-guided reinforcement learning with efficient exploration for task automation of surgical robot. *2023 IEEE International Conference on Robotics and Automation*, 4640-4647.
[16] Pareek, S., Nisar, H. J., & Kesavadas, T. (2024). AR3n: A reinforcement learning-based assist-as-needed controller for robotic rehabilitation. *IEEE Robotics & Automation Magazine, 31*, 74-82.
[17] Zhang, Y., Li, S., Nolan, K. J., & Zanotto, D. (2021). Shaping individualized impedance landscapes for gait training via reinforcement learning.
[18] Guo, B., Han, J., Li, X., & Yan, L. (2019). Human–robot interactive control based on reinforcement learning for gait rehabilitation training robot. *International Journal of Advanced Robotic Systems, 16*, 1-16.
[19] Ou, Y., & Tavakoli, M. (2023). Sim-to-real surgical robot learning and autonomous planning for internal tissue points manipulation using reinforcement learning.