I. J. Engineering and Manufacturing, **2024, 4, 37-53**
Published Online on August 8, 2024 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijem.2024.04.04

Comparative Exploration of the Contribution of Reinforcement Learning in Robotic Surgery

Cheima Bouden*
University of Constantine 2 – Abdelhamid Mehri/department of Fundamental Computer Science and its Applications, Constantine, Algeria
E-mail: cheima.bouden@univ-constantine2.dz
ORCID ID: https://orcid.org/0009-0009-2074-9733
*Corresponding Author

Received: 12 April, 2024; Revised: 14 May, 2024; Accepted: 20 June, 2024; Published: 08 August, 2024

Abstract: Problem: The precision, efficiency, and safety of surgical procedures need significant improvements. Traditional methods are limited by human capabilities, and existing robotic systems lack the advanced adaptability required for complex surgical tasks. The integration of reinforcement learning (RL) into robotic surgery represents a potential revolution in the medical field.

Methods: This comparative review synthesizes recent progress in RL applications for robotic surgery. It highlights innovative methodologies and successful applications of RL, focusing on advanced simulations to train RL agents and the importance of human demonstrations in the learning process.

Results: Emerging trends such as the effective use of simulations and human demonstrations to enhance RL in robotic surgery are identified. The review also discusses challenges associated with RL applications, emphasizing the need for clinical validation and ensuring patient safety.

Conclusion: The transformative potential of RL in robotic surgery is evident, though challenges remain. Future work should prioritize clinical validation, patient safety, and interdisciplinary collaboration.

Index Terms: Artificial Intelligence, Reinforcement Learning, Surgical Robot, Autonomy, Simulation.

1. Introduction

1.1. Structuring the comparative review

This review is structured around key sections: context and issues of robotic surgery, theoretical foundations of reinforcement learning (RL) and robotic surgery, approaches and algorithms in RL applied to robotic surgery, a comparative analysis, and concluding recommendations

1.2. Background and impact of reinforcement learning in robotic surgery

It should be noted that a crucial phase always precedes this operational phase. This is the preoperative phase, where artificial intelligence has honored its promises through several works [1-3], including the aim of better assessing the patient's conditions and mitigating their potential risks. In robotic surgery, RL improves efficiency, precision, and safety by enabling systems to learn and adapt from experience. This technology automates complex tasks, reduces surgeons' workload, and enhances surgical precision, minimizing patient risks.

1.3. Objectives of the comparative study

This study synthesizes and evaluates recent advances in RL integration in robotic surgery, identifying promising approaches, discussing results and limitations, and highlighting current and future trends. Strategic recommendations for researchers and practitioners are provided.

1.4. Reinforcement learning and robotic surgery

Reinforcement learning is a machine learning approach where an agent learns to maximize cumulative rewards through trial and error. Robotic surgery uses robots to assist in procedures, offering greater precision and control, leading to less invasive surgeries and faster recovery.

2. Foundations of Reinforcement Learning

Reinforcement learning is an essential area of artificial intelligence (AI) that focuses on how software or physical agents can learn to make optimal decisions by interacting with an environment. Before exploring through this review the most important works that have advanced RL as an innovative solution in Robotic Surgery, it is essential to reveal the related key concepts.

2.1. Definition of Reinforcement Learning

Reinforcement learning is a machine learning process in which an agent learns to make decisions by performing actions and receiving feedback (rewards or punishments) from the environment. The agent aims to learn a behavioral policy (an action strategy) that maximizes the cumulative sum of rewards received over time. Unlike supervised learning, where the agent learns from data containing the correct answers, reinforcement learning learns through experience without knowing the correct actions to take [4].

2.2. Basic concepts of reinforcement learning

This section includes the definition of key concepts relating to this domain [5]:

- Agent: The entity learns and makes decisions in an environment. It can be a robot, software, or any other system capable of interacting with its environment.
- Environment: Context in which the agent operates and with which it interacts. It can be confirmed, like a robot navigating the real world, or virtual, like an agent playing a video game.
- State: It represents the environment's current situation at a given time. It contains all the relevant information that the agent must consider to make decisions. The state can be observable or unobservable, partial or complete, depending on what the agent can perceive from its environment.
- Actions: Actions are the agent's decisions to influence its environment. Actions can be discrete (e.g., choosing from a finite set of actions) or continuous (e.g., adjusting a parameter continuously).
- Rewards: They are digital signals that the agent receives from the environment to evaluate its actions. Rewards indicate whether a given action is beneficial or detrimental to the goal of long-term reward maximization. The agent seeks to maximize the cumulative sum of rewards over time.
- Policy: A strategy defining the agent's choice of action in each state. The policy can be deterministic (a specific action for each state) or stochastic (a probability distribution over actions for each state).
- Value function: It evaluates the quality of a state (state value function) or a state-action pair (action value function) regarding expected future rewards.
- Episode: A sequence of actions, states, and rewards that ends in a terminal state or after a number of steps.

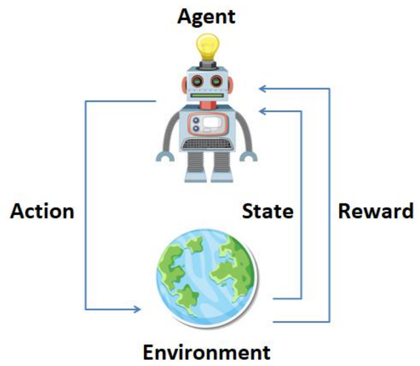


Fig.1. Principle of reinforcement learning

2.3. Main approaches and algorithms in reinforcement learning

In reinforcement learning, several approaches and algorithms are used to teach the agent how to make decisions in an environment, quite often complex to maximize a cumulative reward.

- Q-Learning: Model-free RL method. Learns the quality Q of actions: the expected value of future rewards for an action in a given state. Q-learning finds the optimal policy by learning the Q function that maximizes the cumulative reward. This method does not require knowledge of the environment and can learn only from experience [6].

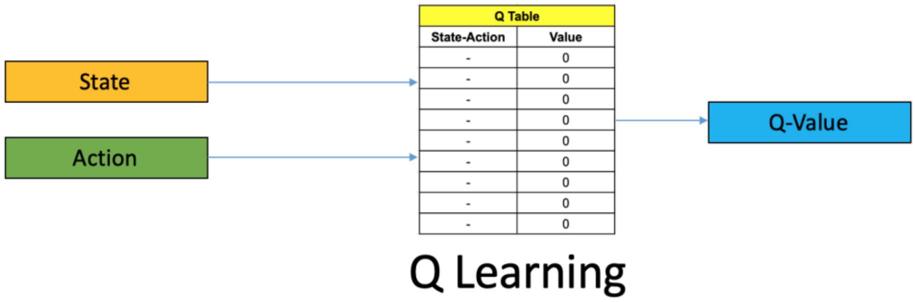


Fig.2. Principle of Q-learning method

- Deep Reinforcement Learning: It combines the techniques of RL with deep learning. This approach uses deep neural networks to approximate the agent's policy (the agent's strategy to choose actions), the value function (which evaluates how good a situation or action is), or both. This makes it possible to deal with large-scale problems with complex state and action spaces [7].

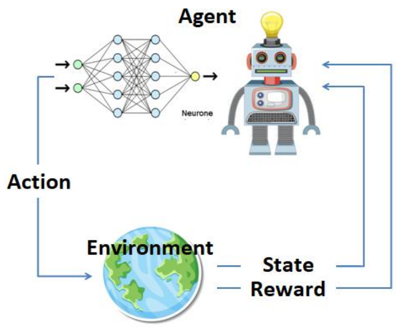


Fig.3. Principle of deep reinforcement learning

- Deep Q-Networks (DQN): It is a deep neural network extension of Q-learning, introduced by Deep Mind that can play video games competitively by observing screen pixels, using techniques like replay experience and target networks [8].

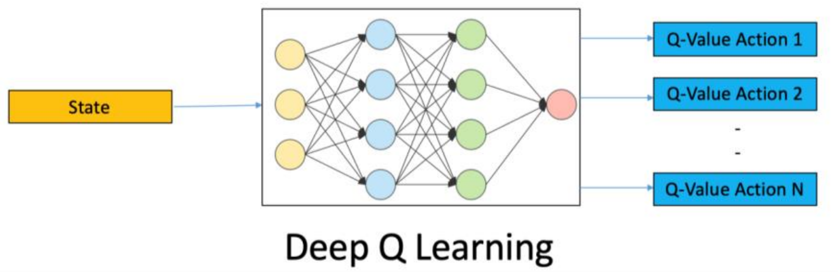


Fig.4. Principle of Deep Q-network method

- Policy Gradients: These methods directly learn the agent's policy by adjusting the parameters to maximize the expected reward. Unlike Q-learning, which learns a value function, Policy gradient methods optimize directly on the policy space. These methods benefit continuous action spaces or when the policy is complex [9].
- Actor-Critic Methods: This approach combines ideas from policy gradient methods and value methods like Q-learning. It uses two models: an "actor", which proposes actions based on the current state, and a "critic", which evaluates the action proposed by the actor using the value function. This combination allows the agent to learn more efficiently by adjusting both the policy (actor) and the value estimates (critic) [10].

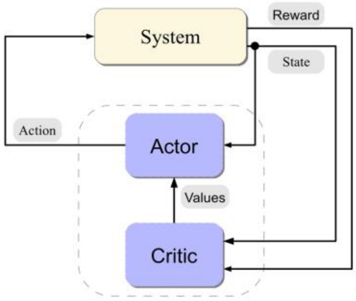


Fig.5. Principle of Actor-Critic method

3. Robotic Surgery

3.1. History and evolution of robotic surgery

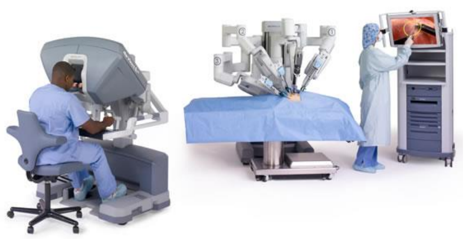
Robotic surgery began in the 1980s, when the first robotic system, the PUMA 560, was used to perform CT-guided neurosurgery in 1985 [11].

However, the da-Vinci system, developed by Intuitive Surgical in the late 1990s, marked a significant turning point in the field. This system was the first surgical robot approved by the FDA (Food and Drug Administration) for general surgery in 2000. Since then, the technology has evolved rapidly, improving existing systems and developing new robots, offering greater precision, flexibility, and control [12].

3.2. Overview of modern surgical robots and their features

Modern surgical robots represent a significant technological advancement in the medical field, providing surgeons with precise tools to perform less invasive procedures with greater precision.

Table 1. Modern surgical robots

Name	Manufactured by	Features	
Da Vinci system [12]	Intuitive Surgical	• High-definition 3D vision: Provides a magnified, three-dimensional view of the surgical site. • Articulated instruments: Robotic arms have great dexterity, allowing more precise and flexible movements than the human hand. • Teleoperation: The surgeon controls the robot from a console, reducing fatigue and improving the precision of surgical procedures.	 Fig.6. Da Vinci system [13]
Mako system [14]	Stryker	• Personalized Preoperative Planning: Uses imaging data to create a personalized surgical plan. • Robotic assistance: Helps the surgeon execute the surgical plan accurately, especially useful for knee and hip replacements.	 Fig.7. Mako system [15]
Rosa system [16]	Zimmer Biomet	• Precise guidance: Allows precise localization of surgical instruments in the brain, useful for neurosurgeries such as biopsies, deep brain stimulation electrodes and epilepsy interventions. Surgical planning. Integrates imaging data to plan intervention trajectories.	 Fig.8. Rosa system [17]

Versius system [18]	CMR Surgical	• Portability and modularity: Designed to be easily moved and configured for a variety of surgical procedures. • Articulated instruments: Offers high dexterity, allowing complex movements in tight spaces.	 Fig.9. Versius system [19]
Senhance System [20]	TransEnterix	• Haptic feedback: Provides sensory feedback to the surgeon, improving control and sensation during operations. • Eye-controlled camera: The surgeon can move the camera view by simply moving the eyes, facilitating better concentration and ergonomics.	 Fig.10. Senhance system [21]

3.3. Current areas of application of robotic surgery

Robotic surgery is currently used in a multitude of surgical specialties, including.

- General surgery, appendectomy, bariatric surgery, etc. [22].
- Urology, prostatectomy, cystectomy, etc [23].
- Gynecology, hysterectomy, myomectomy, etc [24].
- Heart surgery, valve repair, coronary bypass, etc [25].
- Orthopedic surgery, knee replacement, hip replacement, etc [26].
- Neurosurgery, biopsies, treatment of epilepsy, etc [27].

4. Integration of Reinforcement Learning in Robotic Surgery

The Integration of Reinforcement Learning in Robotic Surgery aims to improve the performance of robotic systems and increasingly relieve human presence, where humans will be called upon most for critical situations or supervision.

4.1. Role and potential of reinforcement learning in robotic surgery

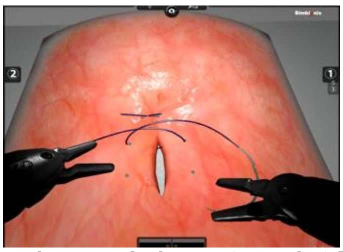
In robotic surgery, the "robot" is the agent that learns to perform surgical tasks by optimizing its actions to achieve a given objective, such as minimizing operation time or maximizing the precision of surgical gestures.

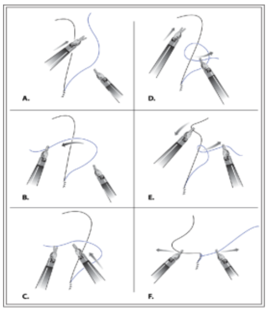
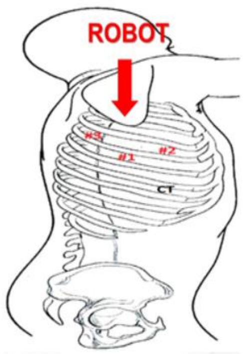
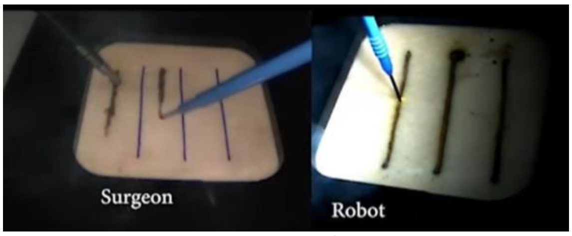
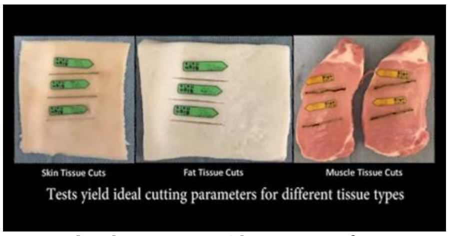
The potential of reinforcement learning in robotic surgery is immense. It would allow robotic systems to adapt and optimize their performance according to the specificities of each intervention and patient. This could lead to safer, faster, and less invasive operations, providing tailored assistance to surgeons [28].

4.2. Practical illustrations of the use of reinforcement learning in robotic surgery

Reinforcement learning in robotic surgery enables robots to master various surgical techniques and maneuvers through an algorithm dedicated to this learning process.

Table 2. Skills that robots can master

Skill		
Performing sutures and ligatures	Surgical robots can learn to perform high-precision sutures and ligatures, ensuring effective incision closure [29,30].	 Fig.11. Suture with a surgical robot – B (screenshot) [31]

		 Fig.12. Suture with surgical robot – A [32]
Organ separation	They can learn to carefully detach and manipulate organs or tissues to prepare and facilitate access to the operating area.	 Fig.13. Organ separation with the surgical robot [33]
Fabric cutting	These robots can make clean incisions in tissues, whether to remove tumors or to resection organs [34].	 Fig.14. Tissue cutting with surgeon (left) and robot (right)  Fig.15. Surgical robot cutting (skin tissue, fat tissue, muscle tissue) [35]
Handling surgical tools	They can also learn to use various surgical tools, such as forceps, scissors, or suction devices, to perform specific procedures [36] [37].	 Fig.16. Handling scalpel with surgical robot

Navigation along complex trajectories	These robotic systems can master complex routes to reach difficult-to-access areas inside the human body while preserving surrounding vital structures [38] [39].	 Fig.17. Navigation along complex trajectories with a surgical robot

4.3. Expected benefits of integrating these two areas

- Improved precision and dexterity: Robots equipped with reinforcement learning can perform surgical tasks with precision and finesse, surpassing human capabilities and reducing the risk of errors and side effects.
- Adaptation and personalization: This integration allows robotic systems to learn from each procedure and adapt to the specific characteristics of each patient, providing a more personalized surgical approach.
- Reduced surgeon fatigue: By automating some of the most complex and repetitive tasks, RL relieves surgeons of physical and mental fatigue, allowing them to focus on the critical aspects of surgery.
- Access to less invasive interventions: Combining these technologies promotes the development of less invasive surgical techniques, reducing patient recovery times and minimizing post-operative complications.
- Improved surgical efficiency: Reinforcement learning can optimize surgical trajectories and techniques, potentially reducing operation time and increasing overall procedural efficiency [40].

5. Methodology of the Comparative Review

Our choice of selection of the works addressed in this review is justified by the importance of their impact on the surgical field; careful and rigorous criteria make the selection.

5.1. Selection of the best ten research works included in the study

The bibliographic search was done on leading scientific databases, such as Google Scholar, IEEE Xplore, Springer, Science Direct, and Pub Med, using the keyword combination " reinforcement learning AND surgical AND robot". This search strategy aimed to capture the broadest possible scope of recent work applying reinforcement learning in surgical robotics. The research period extends from 2022 to 2024 to ensure the inclusion of the most recent and relevant developments in this rapidly evolving field.

5.2. Work selection criteria

The works included were filtered according to several criteria to ensure their relevance and quality.

- Research Sources: The studies were identified through a systematic search in well-known scientific databases, including Pub Med, IEEE Xplore, Google Scholar, Springer, and Science Direct. These sources were selected for their comprehensive coverage of publications in the fields of biomedical engineering, AI, and medicine.
- Access to Publications: Full article access was facilitated by institutional subscriptions at University of Constantine 2, enabling a thorough review of the available literature.
- Search Terms: The search strategy included specific terms in English: "reinforcement learning," "surgical," and "robot." This combination of keywords was specifically chosen to target studies focusing on the application of reinforcement learning in robotic surgery.
- Publication Period: The selected works were recently published between 2022 and 2024 to ensure the relevance and timeliness of the technologies and methods studied. Publications in English were preferred to facilitate analysis and comparison due to the international dissemination of scientific research in this area.
- Selection Procedure: The initial search yielded approximately 80 articles. After reviewing abstracts to assess their relevance to reinforcement learning and robotic surgery, 25 articles were deemed potentially relevant. Full texts of these articles were then reviewed to ensure they met inclusion criteria, focusing particularly on methodological rigor and the direct application of reinforcement learning in surgical contexts.

5.3. Assessment of the Quality of Included Studies

The ten studies evaluated for methodological quality and relevance in reinforcement learning-assisted robotic surgery. While most demonstrate solid methodologies and promising results, limitations include simplification of simulation models and lack of validation in real clinical settings. Despite these, these works contribute significantly to understanding and improving reinforcement learning techniques in robotic surgery.

6. Data Collection and Analysis Methods for Comparison

Data were extracted from each study, focusing on the research objectives, methodologies employed, targeted surgical applications, experimental environments used, results obtained, limitations, and future perspectives identified by the authors. Evaluation measures are also used in the different studies, such as Distance to Goal, Reward, Training Time, Learning Curve, Success Rate, and others, provide a comprehensive basis for comparing and contrasting the various approaches. This methodical approach allows the comparison to be structured around specific criteria, such as learning efficiency, task accuracy, potential clinical application, and challenges to overcome for integration into natural surgical systems.

A multidimensional evaluation framework was established for the comparative analysis to systematize the synthesis of reinforcement learning approaches in surgical robot empowerment. This includes examining the learning techniques (DRL, Q-learning, etc.), and the reported performance and limitations. This review aims to draw out current trends, persistent challenges, and opportunities for future research in reinforcement learning-assisted robotic surgery.

7. Presentation of Selected Research Works (The Top Ten)

This section will present the most relevant works integrating RL into surgical intervention. A detailed sketch will be devoted to each work to identify the essential points for a comparative analysis.

7.1. Work 1

- Title: Concept for a Reinforcement Learning Approach to Navigate Catheters Through Blood Vessels [41]
- Authors: Annika Kienzlen, Florian Jaensch, Alexander Verl, and Leo Cheng, (2022).
- Objective: Automate catheter navigation in the cardiovascular system using reinforcement learning (RL).

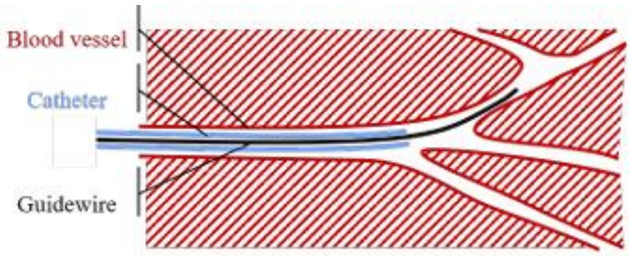


Fig.18. Catheters navigation through blood vessels [41]

- Methodology: Simulated environment with RL elements:
  Agent: Catheter model / Environment: Vascular system simulation / States: Catheter position / Actions: Push, pull, rotate / Rewards: Proximity to target
- Surgical application: Minimally invasive procedures for cardiovascular diseases.
- Experimental environment: SOFA-based simulation, OpenAI Gym integration.
- Metrics Used: Distance to Goal, Reward, Training Time.
- Results: RL agent successfully navigated to the target point in the simulation.
- Limitations: Simplified modeling with rigid vessel walls, no blood flow.
- Future perspectives: Model flexible vessel walls, variable target points, and adjust observation and reward functions for realism.

7.2. Work 2

- Title: GESRsim: Gastrointestinal Endoscopic Surgical Robot Simulator [42]
- Authors: Huxin Gao, Zedong Zhang, Changsheng Li, Xiao Xiao, Liang Qiu, Xiaoxiao Yang, Ruoyi Hao, Xiuli Zuo, Yanqing Li, and Hongliang Ren, (2022).
- Objective: Enhance development and autonomy of gastrointestinal endoscopic surgical robots (GESR) using CoppeliaSim to create the first robotic simulator for the GESR system.

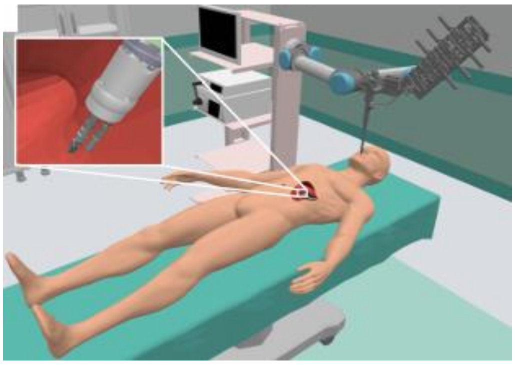


Fig.19. Gastrointestinal endoscopic surgical robot simulator [42]

- Methodology: RL-based training scenarios for autonomous robot training with increasing complexity:
  Agents: Robotic components / Environment: Anatomically detailed simulations / States: Spatial configurations of instruments / Actions: Precise surgical maneuvers / Rewards: Task accuracy and error minimization
- Surgical Application: Simulator for robot-assisted gastrointestinal endoscopic surgery (GES), a type of NOTES.
- Experimental Environment: CoppeliaSim-based with integration for hardware and AI frameworks, providing training scenes and verifying control algorithms.
- Metrics: Learning curve, reward.
- Results: GESRsim developed as a versatile tool, implementing VS and RL control algorithms, and creating ESD training scenarios.
- Limitations: Need for integrating more physics engines for realistic soft material simulation.
- Future Perspectives: Incorporate physics engines like Vortex and SOFA for improved realism and integrate advanced algorithms for robotic autonomy.

7.3. Work 3

- Title: Autonomous Tissue Manipulation via Surgical Robot Using Deep Reinforcement Learning and Evolutionary Algorithms [43]
- Authors: Amin Abbasi Shahkoo and Ahmad Ali Abin, (2022).
- Objective: Optimize tensioning policy and action sequence for precise soft tissue cutting in robotic surgery.

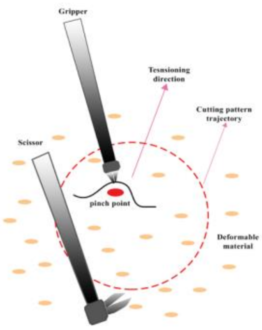


Fig.20. Autonomous tissue manipulation via surgical robot [43]

- Methodology: Combines Deep Reinforcement Learning (DRL) and evolutionary algorithms:
  Agent: Surgical robot with scissors and pincers / Environment: Simulated soft tissue / States: Tension in tissues and instrument position / Actions: Tension sequences / Rewards: Accuracy of cuts and reduction of tissue damage, with policies maximizing these rewards.
- Surgical Application: Precise soft tissue cutting using surgical scissors and forceps in robotic surgery.
- Experimental Environment: Modified Thananjeyan simulator to evaluate method accuracy and reliability.
- Metrics: Cutting accuracy, reward.
- Results: Significant improvement in cutting accuracy and tissue damage reduction compared to existing methods, with a 72.2% average improvement in symmetrical difference quality over tension-free methods.
- Limitations: Limited to non-intersecting patterns; discrete space movement of robot's tensioner arm.

- Future Perspectives: Explore tension policies with continuous space movement for more precise cutting on complex patterns, and use multiple grippers.

7.4. Work 4

- Title: Deep Reinforcement Learning in Continuous Action Space for Autonomous Robotic Surgery [44]
- Authors: Amin Abbasi Shahkoo, Ahmad Ali Abin, (2023).
- Objective: To provide a deep reinforcement learning-based approach to control the gripping arm during soft tissue cutting in a continuous action space to increase surgical precision and reduce stress and pressure on the surgeon.

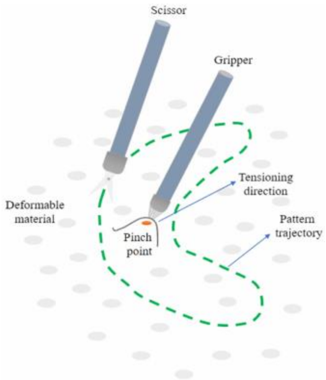


Fig.21. Autonomous Tissue manipulation via surgical robot [44]

- Methodology:
  Agent: Robotic arm performing surgical tasks / Environment: Simulated soft tissue cutting in minimally invasive surgery / Actions: Precise arm movements, such as cutting angles and depths, continuously merged for increased precision / Rewards: Encouraging precise cuts and minimizing tissue damage, with evolving policies from exploratory to exploitative
- Surgical Application: Soft tissue cutting in minimally invasive surgery (MIS).
- Experimental Environment: Modified simulator supporting continuous action space with high-resolution simulated soft tissue.
- Metrics: Cutting accuracy, symmetric difference.
- Results: Demonstrated better performance than discrete action space methods, with a 70% average improvement in cutting precision over baselines.
- Limitations: Limited in handling crossed contour patterns, lack of tactile sensation, and separate training phases for each pattern.
- Future Perspectives: Develop optimal tension policies for multiple pinch points, improve pinch point sampling algorithms, and enable the tension arm to move to different points during cutting.

7.5. Work 5

- Title: Human-in-the-Loop Embodied Intelligence With Interactive Simulation Environment for Surgical Robot Learning [45]
- Authors: Yonghao Long, Wang Wei, Tao Huang, Yuehao Wang, and Qi Dou, (2023).
- Objective: Integrate human embodied intelligence in an interactive simulation platform to enhance surgical robot learning through human demonstrations and reinforcement learning (RL).
- Methodology:
  Agent: Surgical robot learning tasks through an interactive simulator with haptic inputs
  Environment: Simulator replicating complex surgical procedures / Actions: Continuous surgical maneuvers for precision / Rewards: Promote precision and efficiency, aligned with human expert demonstrations
  Policy: Refined through human demonstrations and continuous feedback
- Surgical Application: Training and automating surgical tasks, improving robotic skills through RL and human demonstrations.
- Experimental Environment: SurRoL simulator, supporting high-quality human interactions via a haptic input device, with realistic scene rendering and physical interactions.
- Metrics: Success rate, steps to complete, economy of motion.
- Results: Improved learning efficiency, stability, and performance of policies learned from human expert demonstrations compared to non-experts or scripts.

- Limitations: Focused on simulation; requires validation in real-world surgical scenarios.
- Future Perspectives: Explore surgical safety in policy learning from human demonstrations, investigate learning from human feedback and human-robot collaboration, and develop an open-source embedded AI platform for surgical robots.

7.6. Work 6

- Title: Learning Needle Pick-and-Place Without Expert Demonstrations [46]
- Authors: Rokas Bendikas, Valerio Modugno, Dimitrios Kanoulas, Francisco Vasconcelos, and Danail Stoyanov, (2023).
- Objective: Develop a novel approach for learning complex needle manipulation tasks in surgical applications using reinforcement learning without expert demonstrations.

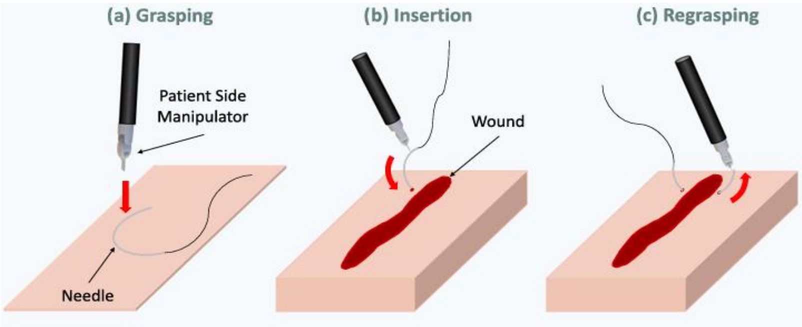


Fig.22. Grasping – insertion – regrasping [46]

- Methodology: Task Decomposition: Break down tasks into simpler subtasks
  Algorithm: Actor-critic with deterministic policy / Training Steps: NeedleReach (Learning to reach and grasp the needle) and NeedlePickAndPlace (Completing the needle picking and placement task) / Agent: Robotic system / Environment: Modified SurRoL simulation / Actions: Precise movements for needle manipulation / Rewards: Promoting efficient and precise task completion / Policy: Guided by actor-critic algorithm
- Surgical Application: Enhance efficiency and precision in surgical needle handling tasks.
- Experimental Environment: Modified SurRoL simulation with Da Vinci Research Kit (dVRK).
- Metrics: Success rate, episode length.
- Results: Achieved a 91% success rate, outperforming methods using expert demonstrations.
- Limitations: Manual task decomposition, lower sampling efficiency, untested in real-world scenarios.
- Future Perspectives: Automate task decomposition, improve sample efficiency, bridge simulation-to-real-world gap, extend to other robotic surgery tasks.

7.7. Work 7

- Title: Surgical Continuum Manipulator Control Using Multi-agent Team Deep Q Learning [47]
- Authors: Guanglin Ji, Qian Gao, Minyi Sun, Guanyu Mi, Xinyao Hu, and Zhenglong Sun, (2023).
- Objective: Develop and test a Team Deep Q (TDQN) learning framework for precise control of a 2 DoF surgical continuum manipulator driven by four cables, where each cable pair forms a unique agent.
- Methodology:
  RL Algorithm: Team Deep Q Network (TDQN) / Agents: Each cable pair acts as a separate agent / Environment: Simulated, manipulating a 2 DoF continuum manipulator / Rewards: Incentivize precision and stability / Coordination: TDQN optimizes agents' actions for control
- Surgical Application: Used in neurosurgery, vascular surgery, and abdominal surgery for accessing confined surgical sites.
- Experimental Environment: Simulations with a 3D printed continuum manipulator, controlled by motors via an EPOS4 controller, tracked in real-time with a camera.
- Metrics: Root Mean Square Error (RMSE), reward.
- Results: TDQN showed better target accuracy and robustness to disturbances, with reduced RMSE compared to the Multi-Agent Deep Q (MADQN) algorithm.
- Limitations: Focus on 2 DoF manipulator may limit generalizability to more complex configurations.
- Future Perspectives: Extend TDQN to multi-segment continuous robots, perform ex vivo experiments with multi-DoF robots to validate effectiveness in real surgical scenarios.

7.8. Work 8

- Title: Towards Safe and Efficient Reinforcement Learning for Surgical Robots Using Real-Time Human Supervision and Demonstration [48]
- Authors: Yafei Ou and Mahdi Tavakoli, (2023).
- Objective: Present a deep reinforcement learning (DRL) framework for surgical robot automation that integrates real-time human supervision during training to prevent failures and accelerate learning.

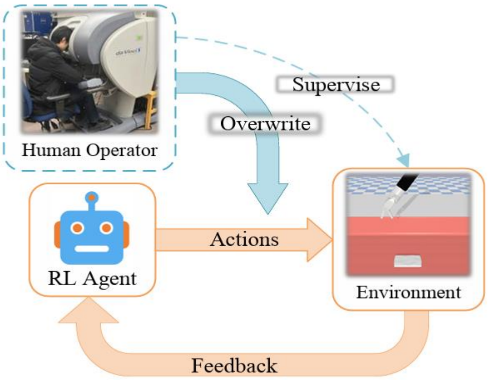


Fig.23. Human-guided surgical robots [48]

- Methodology:
  RL Approach: Combines RL and Generative Adversarial Imitation Learning (GAIL) / Environment: Simulated, with human interventions via teleoperation / Rewards: Minimize errors and failures / Policy: Continuous, enhanced by human interventions / Optimization: Refine reward function, optimize human-machine interactions
- Surgical Application: Increase surgical robot autonomy to reduce surgeons' workload in tasks like needle re-entry and tissue retraction.
- Experimental Environment: Two simulations with human intervention via joystick controller and da Vinci Research Kit (dVRK) MTM.
- Metrics: Success rate, learning curve.
- Results: Outperformed baseline algorithms, demonstrating greater learning efficiency and failure avoidance.
- Limitations: Tested in simple simulated environments, limiting generalizability to real-world scenarios.
- Future Perspectives: Extend to real-world surgical scenarios and specific tasks, explore using experienced physicians as supervisors.

7.9. Work 9

- Title: Trajectory Planning and Obstacle Avoidance Behavior of Manipulator Based on QL Algorithm [49]
- Authors: Guo Xinlan, Li Tao, Zhang Jian, (2023).
- Objective: Correct trajectory deviation in the movement and obstacle avoidance of a robotic arm using a control algorithm to ensure the actual trajectory is close to the ideal trajectory.
- Methodology:
  Agent: Surgical robotic arm navigating in a MATLAB simulator / Environment: MATLAB simulation / Rewards: Minimize deviation and reduce travel cost for efficient navigation / Policy: Guided by a backpropagation neural network, optimizing performance objectives / Algorithm: Continuously approximates Q values, adjusting state modeling for improved convergence speed and migration cost
- Surgical Application: While not directly addressed, the principles of path planning and obstacle avoidance are relevant for advanced surgical robotic arms.
- Experimental Environment: MATLAB simulation testing the Q-learning algorithm's control over a manipulator with three joints and two connecting rods.
- Metrics: Convergence speed, path planning effectiveness, migration cost.
- Results: The control algorithm showed fast convergence, better path planning than traditional schemes, and the lowest migration cost.
- Limitations: Focuses on simulation environments without real-world validation or specific surgical applications.
- Future Perspectives: Apply the improved Q-learning algorithm to real-world surgical scenarios to enhance the precision and reliability of surgical robotic arms.

7.10. Work 10

- Title: Deep Reinforcement Learning for Concentric Tube Robot Path Following [50]
- Authors: Keshav Iyengar, Sarah Spurgeon, and Danail Stoyanov, (2024).
- Objective: Propose a deep reinforcement learning approach to control concentric tube robots (CTRs) in surgery, enhancing control, precision, and handling of rotational constraints.

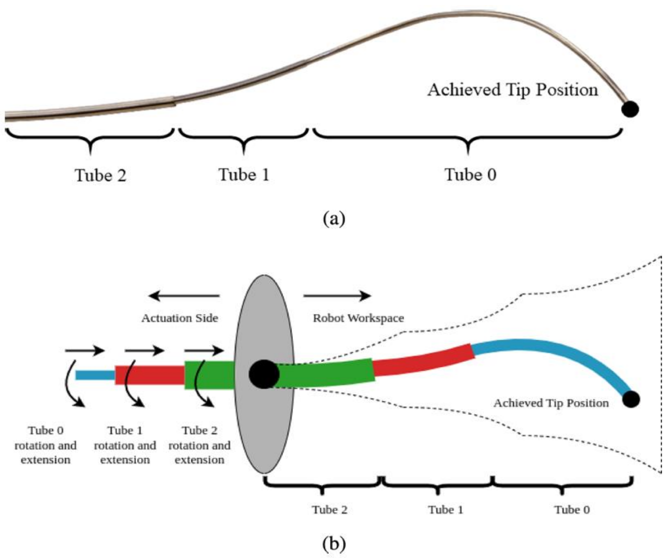


Fig.24. Concentric tube robot path following [50]

- Methodology:
  Agent: Robot controller / Environment: Simulation of CTR kinematics / Actions: Adjust tube positions to follow a trajectory / Rewards: Minimize trajectory tracking and inverse kinematic errors / Algorithm: Deep Deterministic Policy Gradient (DDPG) with Hindsight Experience Replay (HER)
- Surgical Application: CTRs for minimally invasive procedures in brain, eye, fetoscopic, pulmonary, cardiac, and prostate surgeries.
- Experimental Environment: OpenAI Gym-based simulation to train and evaluate control policies, study inverse kinematics, and path following.
- Metrics: Mean path error, inverse kinematic error.
- Results: The approach generalizes control across multiple CTR systems, improves inverse kinematics and path tracking accuracy, and reduces the impact of rotational constraints.
- Limitations: Simulation-only study without hardware validation; slamming problem and domain transfer methods need exploration.
- Future Perspectives: Hardware validation, exploring domain transfer methods, and developing dynamic simulation frameworks to avoid slamming.

8. Comparative Analysis of Works

This part serves as an analytical interface of previous works in order to derive the most recommended for the field of robotic surgery.

8.1. Comparison and Discussion

The selected works highlight several innovative applications of reinforcement learning (RL) in robotic surgery, notably for catheter navigation, tissue manipulation, trajectory planning, and manipulator control. Several studies have successfully demonstrated a significant improvement in surgical robots' precision, autonomy, and efficiency by applying advanced RL techniques, such as Deep Reinforcement Learning (DRL), Q-Learning, and Actor-Critic approaches.

To further understand the impact of these techniques, this paper compares the learning efficiency across the selected works. The effectiveness of each approach varies based on the specific task and environment.

Table 3. Learning Efficiency of each work

Algorithm	Task	Learning Efficiency
Work 1: DQN	Catheter Navigation	40% of training
Work 2: PPO	Endoscopic Manipulator Guidance	500k steps
Work 3: DRL + EA	Tissue Manipulation and Cutting	300 generations
Work 4: TRPO	Pattern Cutting on Soft Tissue	25 iterations with batch size 500
Work 5: DDPG with Human Demonstrations	Needle Manipulation with Human Guidance	50 epochs
Work 5: TD3 with HER	Pick-and-Place Needle Task	2 million steps
Work 7: TDQN and MADQN	Continuum Manipulator Control	50 episodes
Work 8: SAC with GAIL	Surgical Tasks with Human Supervision	10,000 to 40,000 steps
Work 9: Q-learning Improved	Trajectory Planning and Obstacle Avoidance	500 iterations
Work 10: DRL with PPO	Path Following for Concentric Tube Robots	2 million steps

8.2. *Identifying trends and similarities*

A notable trend is the integration of advanced simulations for training RL agents before their application in real-world environments, highlighting the importance of simulation for the safe and efficient development of robotic surgical technologies. Several works have also explored hybrid approaches, combining RL with other machine learning techniques, to overcome specific challenges, such as adaptation to complex surgical environments and precise manipulation of tissues or instruments.

8.3. *The most effective approaches and the most promising results*

Among the various approaches, some stand out due to their effectiveness and promising results:

- Work 1: DQN for catheter navigation achieved 40% training efficiency, showing precision improvement.
- Work 2: PPO for endoscopic manipulator guidance reached efficiency after 500k steps, aided by anatomically detailed simulations.
- Work 3: DRL combined with evolutionary algorithms (EA) for tissue manipulation and cutting demonstrated a 72.2% improvement in cutting accuracy.
- Work 6: TD3 with HER for needle pick-and-place tasks achieved high precision after 2 million steps.
- Work 8: SAC with GAIL integrating human supervision showed improved learning efficiency and failure avoidance.

9. Advantages and Limitations

To help with the validity and reliability of previous work, conducting a critical analysis allows researchers to suggest improvements and better guide future work.

9.1. *Discussion of the advantages retained in the work*

The works offer several advantages that enhance the field of robotic surgery:

- Improved Precision and Autonomy: Many studies demonstrated substantial improvements in the precision and autonomy of surgical robots, reducing the need for constant human intervention.
- Enhanced Training Efficiency: Advanced RL algorithms, especially when combined with human demonstrations or supervision, have shown to accelerate the training process and improve learning outcomes.
- Versatility in Application: The works cover a range of surgical tasks, from catheter navigation to tissue manipulation and needle handling, showcasing the broad applicability of RL in surgical robotics.

9.2. *Discussion of limitations and challenges mentioned in the work*

- Realism of Simulations: While simulations are essential for training RL agents, there is a significant gap between simulated and real-world environments. This gap can lead to discrepancies in the robot's performance when deployed in actual surgical settings.
- Transfer of Skills to Real Environments: Ensuring that skills acquired in simulated environments are effectively transferred to real-world scenarios is a critical challenge. The differences in physical properties and dynamics between simulations and real environments can impact the robot's performance.

- Complex Action Spaces: Managing complex action spaces in RL, particularly in the context of surgical procedures, remains a significant challenge. The need for precise and continuous control in a highly dynamic environment adds to the complexity.
- Safety and Reliability: Ensuring the safety and reliability of autonomous surgical procedures is paramount. The lack of haptic feedback and other sensory inputs in current robotic systems increases the risk of complications during surgery.

10. Conclusion and Perspectives

This manuscript has compared advances in artificial intelligence within surgical interventions, highlighting how the integration of new AI models is revolutionizing robotic surgery. These innovations are making procedures more precise, autonomous, and efficient. The application of reinforcement learning (RL) in robotic surgery shows immense potential for improving surgical outcomes. The analysis of ten works demonstrates significant advancements in precision, autonomy, and training efficiency driven by advanced RL algorithms.

However, further research is needed to address the limitations of simulated environments and simplified models, and to validate findings in real-world scenarios. Future directions include integrating more realistic physics engines, enhancing human-robot collaboration, and exploring new RL techniques to optimize performance. Ongoing research in these areas will expand the possibilities in robotic surgery, leading to safer and more effective interventions.

It is recommended to continue the development of interactive simulations and precise modeling for robotic training. Using human demonstrations and imitation learning techniques can enhance the learning process. Implementing multi-agent learning and collaborative methods for more complex procedures is also crucial. Ensuring clinical validation and rigorous safety testing is necessary to verify the effectiveness and safety of these technologies. Finally, interdisciplinary collaboration between engineers, AI researchers, surgeons, and healthcare professionals will drive innovation and practical application. Such efforts are essential for overcoming current challenges and maximizing the potential of AI-driven surgical robotics.

Declaration of AI and AI-assisted Technologies in the Writing Process

During the preparation of this manuscript, AI-assisted tools such as ChatGPT, Grammarly and Turnitin were used for drafting and revising the content. All contributions from these tools were thoroughly reviewed and revised by the authors to ensure accuracy and originality of the final manuscript.

Acknowledgments

Thanks to the University of Constantine 2 - Abdelhamid Mehri for providing the resources and support necessary to conduct this research.

References

[1] Bouden, C., & Mezioud, C. (2023, August). An Intelligent Decision Support System for the Surgical Preoperative Phase: An Approach Based on Machine Learning. In International conference on WorldS4 (pp. 483-494). Singapore: Springer Nature Singapore.
[2] Bouden, C., & Mezioud, C. (2023). Towards an Intelligent Surgical Preoperative Decision Support System: An Approach Based on Decision Trees and the Random Forest. Intelligent Methods In Engineering Sciences, 2(3), 67-74.
[3] Bouden, C., & Mezioud, C. (2023, April). Towards an Intelligent Preoperative Surgical Decision Support System: A machine learning-based approach. In Proceedings of International Conference on Intelligent Systems and New Applications (Vol. 1, pp. 19-26).
[4] L. P. Kaelbling, M. L. Littman, & A. W. Moore, (1996). Reinforcement learning: A survey. Journal of artificial intelligence research, 4, 237-285.
[5] R. S. Sutton, & A. G. Barto (1999). Reinforcement learning. Journal of Cognitive Neuroscience, 11(1), 126-134.
[6] P. Dayan, & CJCH Watkins, (1992). Q-learning. Machine Learning, 8(3), 279-292.
[7] Li, Y. (2017). Deep reinforcement learning: An overview. arXiv preprint arXiv:1701.07274.
[8] M. Roderick, J. MacGlashan, & S. Tellex (2017). Implementing the deep q-network. arXiv preprint arXiv:1711.07478.
[9] J. Peters, & S. Schaal (2008). RL of motor skills with policy gradients. Neural networks, 21(4), 682-697.
[10] V. Konda, & J. Tsitsiklis (1999). Actor-critic algorithms. Advances in neural information processing systems, 12.
[11] B. Armstrong, O. Khatib, & J. Burdick (1986, April). The explicit dynamic model and inertial parameters of the PUMA 560 arm. In Proceedings. 1986 IEEE International Conference on Robotics and Automation (Vol. 3, pp. 510-518).
[12] S. Tsuda, D. Oleynikov, J. Gould, D. Azagury, B. Sandler, M. Hutter & R. Satava (2015). SAGES TAVAC safety and effectiveness analysis: da Vinci® surgical system (Intuitive Surgical…). Surgical endoscopy, 29, 2873-2884.
[13] www.maxongroup.ch/maxon/view/application/Am-OP-Tisch-steht-ein-Roboterchirurg?isoCode=fr
[14] J. D. Sires, J. D. Craik, & C. J. Wilson (2019). Accuracy of bone resection in MAKO total knee robotic-assisted surgery. The journal of knee surgery, 34(07), 745-748.

[15] www.docmartinortho.vegas/mako-smartrobotics/
[16] C. Batailler, D. Hannouche, F. Benazzo, & S. Parratte (2021). Concepts and techniques of a new robotically-assisted technique for total knee arthroplasty: the ROSA knee system. Archives of Orthopedic and Trauma Surgery, 141(12), 2049-2058.
[17] www.zimmerbiomet.com/en/products-and-solutions/zb-edge/robotics/rosa-brain.html
[18] B. C. Thomas, M. Slack, M. Hussain, N. B, A. P, E. D, & G.D.S (2021). Preclinical evaluation of the versus surgical system, a new robot-assisted surgical device for use in minimal access renal and prostate surgery. European urology focus, 7(2), 444-452.
[19] www.maxhealthcare.in/technology/versius-next-generation-robotic-surgery
[20] T. deBeche -Adams, W. S. Eubanks, & SG de la Fuente, (2019). Early experience with the Senhance ®-laparoscopic/robotic platform in the US. Journal of Robotics Surgery, 13(2), 357-359.
[21] www.medicaldevice-network.com/projects/asensus-surgicals-senhance-surgical-system-usa/?cf-view&cf-closed
[22] E.B. Wilson, (2009). The evolution of robotic general surgery. Scandinavian Journal of Surgery, 98(2), 125-129.
[23] A. A. Shah, J. Bandari, D. Pelzman, B. J. Davies, & B. L. Jacobs (2021). Diffusion and adoption of the surgical robot in urology. Translational andrology and urology, 10(5), 2151.
[24] A. Gioè, G. Monterossi, S. Gueli Alletti, G. Panico, G. Campagna, B. Costantini & G. Scambia, (2024). The new robotic system HUGO RAS for gynecologic surgery: First European experience from Gemelli Hospital. International Journal of Gynecology & Obstetrics.
[25] P. Q. Thuan, PT V. Chuong, NHNam & NH Dinh (2023). Coronary Artery Bypass Surgery: Evidence-Based Practice. Cardiology in Review, 10-1097.
[26] D, K., & Koutserimpas, C. (2024). Pitfalls with the MAKO Robotic-Arm-Assisted Total Knee Arthroplasty. Medicine, 60(2), 262.
[27] O. Khanna, R.D. Beasley, F, & S. DM (2021). The path to surgical robotics in neurosurgery. Operative Neurosurgery, 20(6), 514-520.
[28] J. Chen, H. Y. Lau, W. Xu, & H. Ren (2016, February). Towards transferring skills to flexible surgical robots with demonstration and reinforcement learning programming. In 2016 Eighth International Conference on Advanced Computational Intelligence (ICACI) (pp. 378-384). IEEE.
[29] Guru, Khurshid, Mohd Raashid Sheikh, Syed Johar Raza, Andrew Stegemann and John Nyquist. “New knotting technique for robot-assisted surgery” The Canadian Journal of Urology 19 4 (2012): 6401-3.
[30] V. Varier, VM, Rajamani, DK, Goldfarb, N., Tavakkolmoghaddam, F., Munawar, A., & Fischer, GS (2020, August). Collaborative suturing: A reinforcement learning approach to automate hand-off tasks in suturing for surgical robots. In 2020, 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN) (pp. 1380-1386). IEEE.
[31] www.youtube.com/watch?v=u33fOYOjCZ8
[32] RC Jackson and MC Cavusoglu, “Needle path planning for autonomous robotic surgical suturing,” in 2013 IEEE International Conference on Robotics and Automation, Karlsruhe, Germany: IEEE, May 2013, p. 1669 1675. doi: 10.1109/ICRA.2013.6630794.
[33] F. Gharagozloo and M. Meyer, “Robotic Transthoracic First Rib Resection for Neurogenic Thoracic Outlet Syndrome,” WJCS, vol. 12, no. 01, p. 1 11, 2022, doi: 10.4236/wjcs.2022.121001.
[34] B. Thananjeyan, A. Garg, S. Krishnan, C. Chen, L. Miller, and K. Goldberg, "Multilateral surgical pattern cutting in 2D orthotropic gauze with deep reinforcement learning policies for tensioning," 2017 IEEE International Conference on Robotics and Automation (ICRA), Singapore, 2017, pp. 2371-2378, doi: 10.1109/ICRA.2017.7989275.
[35] R. Bogue, “Cutting robots: a review of technologies and applications”, Industrial Robot: An International Journal, vol. 35, no. 5, p. 390 396, August 2008, doi: 10.1108/01439910810893554.
[36] M. Jacob, Y.-T. Li, G. Akingba, and JP Wachs, “ Gestonurse: a robotic surgical nurse for handling surgical instruments in the operating room,” J Robotic Surg, vol. 6, no. 1, p. 53 63, March 2012, doi: 10.1007/s11701-011-0325-0.
[37] X. Tan, C. B. Chng, Y. Su, K. B. Lim, & C. K. Chui (2019). Robot-assisted training in laparoscopy using deep reinforcement learning. IEEE Robotics and Automation Letters, 4(2), 485-492.
[38] J. Sun, D. Zhou, Y. Liu, J. Deng, and S. Zhang, “Development of a Minimally Invasive Surgical Robot Using Self-Helix Twisted Artificial Muscles,” IEEE Trans. Ind. Electron., vol. 71, no. 2, p. 1779 1789, Feb. 2023, doi: 10.1109/TIE.2023.3257374.
[39] W. Chi, J. Liu, M. E. Abdelaziz, G. Dagnino, C. Riga, C. Bicknell, & G. Z. Yang (2018, October). Trajectory optimization of robot-assisted endovascular catheterization with reinforcement learning. In 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (pp. 3875-3881). IEEE.
[40] A. Bourdillon, A. Garg, H. Wang, Y. J. Woo, M. Pavone, & J. Boyd (2023). Integration of reinforcement learning in a virtual robotic surgical simulation. Surgical Innovation, 30(1), 94-102.
[41] A. Kienzlen, F. Jaensch, A. Verl and L. Cheng, "Concept for a Reinforcement Learning Approach to Navigate Catheters Through Blood Vessels," 2022 28th International Conference on Mechatronics and Machine Vision in Practice (M2VIP), Nanjing, China, 2022, p. 1-4, doi: 10.1109/M2VIP55626.2022.10041096.
[42] H. Gao et al., " GESRsim: Gastrointestinal Endoscopic Surgical Robot Simulator," 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Kyoto, Japan, 2022, pp. 9542-9549, doi: 10.1109/IROS47612.2022.9982138.
[43] AA Shahkoo and AA Abin, "Autonomous Tissue Manipulation via Surgical Robot Using Deep Reinforcement Learning and Evolutionary Algorithms," in IEEE Transactions on Medical Robotics and Bionics, vol. 5, no. 1, pp. 30-41, Feb. 2023, doi: 10.1109/TMRB.2023.3237772.
[44] AASHahkoo, AAAbin “Deep reinforcement learning in continuous action space for autonomous robotic surgery,” Int J CARS 18, 423–431 (2023). https://doi.org/10.1007/s11548-022-02789-8
[45] Y. Long, W. Wei, T. Huang, Y. Wang, and Q. Dou, "Human-in-the-Loop Embodied Intelligence With Interactive Simulation Environment for Surgical Robot Learning,” in IEEE Robotics and Automation Letters, vol. 8, no. 8, pp. 4441-4448, Aug. 2023, doi: 10.1109/LRA.2023.3284380.

[46] R. Bendikas, V. Modugno, D. Kanoulas, F. Vasconcelos and D. Stoyanov, “Learning Needle Pick-and-Place Without Expert Demonstrations,” in IEEE Robotics and Automation Letters, vol. 8, no. 6, pp. 3326-3333, June 2023, doi: 10.1109/LRA.2023.3266720.

[47] G. Ji, Q. Gao, M. Sun, G. Mi, X. Hu, and Z. Sun, "Surgical Continuum Manipulator Control Using Multi-agent Team Deep Q Learning," 2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Sydney, Australia, 2023, pp. 1-5, doi: 10.1109/EMBC40787.2023.10340943.

[48] Y. Ou and M. Tavakoli, "Towards Safe and Efficient Reinforcement Learning for Surgical Robots Using Real-Time Human Supervision and Demonstration," 2023 International Symposium on Medical Robotics (ISMR), Atlanta, GA, USA, 2023, pp. 1-7, doi: 10.1109/ISMR57123.2023.10130214.

[49] G. Xinlan, L. Tao and Z. Jian, "Trajectory Planning and Obstacle Avoidance Behavior of Manipulator Based on Q-learning Algorithm," 2023 IEEE International Conference on Control, Electronics and Computer Technology (ICCECT), Jilin, China, 2023, pp. 1235-1241, doi: 10.1109/ICCECT57938.2023.10140814.

[50] K. Iyengar, S. Spurgeon and D. Stoyanov, "Deep Reinforcement Learning for Concentric Tube Robot Path Following," in IEEE Transactions on Medical Robotics and Bionics, vol. 6, no. 1, pp. 18-29, Feb. 2024, doi: 10.1109/TMRB.2023.3310037.

Authors’ Profiles

Cheima Bouden is a PhD student specializing in artificial intelligence at Abdelhamid Mehri Constantine 2 University. Previously an IT engineer, now a part-time teacher of artificial intelligence courses, conducting research on AI applications in the medical and surgical fields.

How to cite this paper: Cheima Bouden, "Comparative Exploration of the Contribution of Reinforcement Learning in Robotic Surgery", International Journal of Engineering and Manufacturing (IJEM), Vol.14, No.4, pp. 37-53, 2024. DOI:10.5815/ijem.2024.04.04