# Bridging AI and Surgery: Exploring Reinforcement Learning in Surgical Contexts

**Antonio Damiano**
Alma Mater Studiorum
Msc Artificial Intelligence
antonio.damiano3@studio.unibo.it

## Abstract

This project explores the convergence of Artificial Intelligence (AI) and surgical procedures, leveraging reinforcement learning and deep neural networks. We investigate AI's application in surgery through the development of 2D surgical simulations, the integration of reinforcement learning using neural networks, and a comprehensive examination of the learning process. The results demonstrate AI's potential to enhance surgical techniques, subject to defined rules, rewards, and specialized optimizations. This work replicates an existing experiment (1), modifying the environment and model used, focusing on a prostatectomy environment compared to the original hysteroctomy setting.

## 1 Introduction

The integration of artificial intelligence (AI) and machine learning into healthcare has ushered in transformative possibilities. The paper "Can Reinforcement Learning Be Applied to Surgery?" authored by Masakazu Sato, Kaori Koga, Tomoyuki Fujii, and Yutaka Osuga (1), explores applying reinforcement learning (RL) techniques to surgical procedures, with a focus on hysteroctomy.

RL offers the potential to revolutionize surgical practices by introducing AI agents that assist surgeons. These agents optimize surgical trajectories, adapt to patient-specific variations, and enhance surgical skills in real-time. This paper investigates the feasibility and implications of this integration, aiming to unlock new frontiers in surgical precision and patient safety through AI-enhanced surgery.

The provided code implements a simulation environment for surgical procedures, focusing on prostatectomy. It employs concepts from reinforcement learning and deep neural networks to train an AI agent for surgical decision-making. The environment is represented as a grid, and the agent's objective is to navigate this grid while performing surgical tasks efficiently and safely.

### 1.1 Environment

The surgical environment class defines its dimensions, grid representation, object positions (prostate), target positions (cutting points), bladder positions, positions to avoid (nerves and bladder conduits), and scalpel positions. It also provides methods to create and display the surgical environment, perform actions, and reset the environment. This representation simplifies an actual surgical procedure but serves the project's objectives effectively.

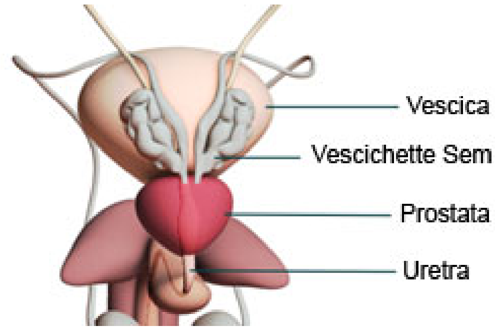


Figure 1: Actual Prostate

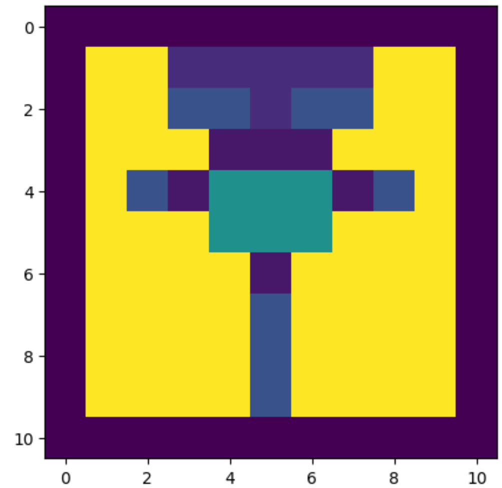

Figure 2: Environment Representation

### 1.2 Model

The code provides a simple neural network model for RL in a surgical environment, contrasting with the paper's more complex model. This model, implemented using TensorFlow, comprises key components:

- A 2D convolutional layer with 16 filters and ReLU activation.
- A Flatten layer to reshape the output.
- A fully connected Dense layer with 32 units and ReLU activation.
- A final Dense layer with units corresponding to possible actions (up, right, down, left) and linear activation.

The model is compiled with the RMSprop optimizer and mean squared error as the loss function.

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d (Conv2D)             (None, 9, 9, 16)          160

 flatten (Flatten)           (None, 1296)              0

 dense (Dense)               (None, 32)                41504

 dense_1 (Dense)             (None, 4)                 132

=================================================================
Total params: 41,796
Trainable params: 41,796
Non-trainable params: 0
_________________________________________________________________
```

Figure 3: Actual Prostate

This simpler model in the code is suitable for limited hardware resources, such as CPUs or GPUs with less computational power. It serves educational and experimental purposes, while the paper's model may be tailored for real-world surgical applications.

## 2 Agent

The agent in the code is designed for RL in a simulated surgical environment, consisting of:

### 2.1 Surgical Environment

The agent interacts with a grid-based simulated surgical environment, which includes objects, targets, bladder, and areas to avoid. The agent's goal is to navigate this environment efficiently while performing surgical tasks.

### 2.2 Rewards

The rewards after each action take into account the cut tissue follow this table:

Table 1: Agent Rewards

| Episode | Reward |
|---|---|
| Border | -5 |
| Bladder or Prostate | -5 |
| Target Points | +10 |
| Points to avoid | -3 |
| Peritoneum | -1 |

### 2.3 Neural Network Model

The agent employs a neural network to approximate the action-value function (Q-function) for RL. The model takes the current environment state as input and outputs Q-values for possible actions.

### 2.4 Reinforcement Learning Algorithm

The agent uses a Q-learning variant for training, including experience replay and updates based on the Bellman equation. An epsilon-greedy strategy balances exploration and exploitation.

### 2.5 Training and Evaluation

The agent undergoes training episodes, exploring the environment, taking actions, and updating Q-values. Hyperparameters include episode count, minibatch size, memory buffer size, exploration rate, epsilon decay, and discount factor.

After training, the agent is saved as a trained model and evaluated in a testing phase.

### 2.6 Hardware Resource Considerations

The model's simplicity ensures compatibility with hardware limitations. It is accessible for educational and experimental purposes and suitable for CPUs or GPUs with lower processing power.

### 2.7 Issues

#### 2.7.1 Time Constraints

Due to time limitations, the grid search for hyperparameters was prematurely halted. Hyperparameters may not be thoroughly fine-tuned.

#### 2.7.2 Lack of Exploration

The final trained agent exhibited limited exploration, often repeating the same actions. This issue requires further investigation and adjustments.

These issues highlight areas for improvement in future work.

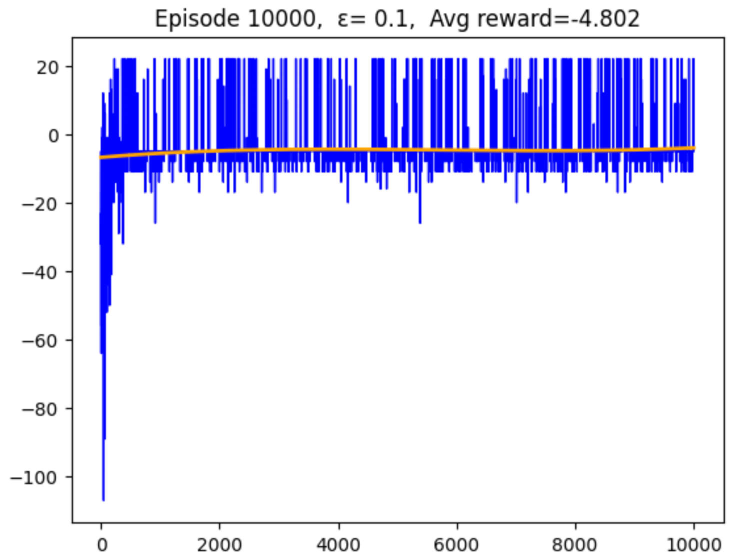

Figure 4: Training Rewards and Trend Line

## 3 Results

Results from the current implementation may not meet expectations initially. However, refinement through advanced techniques, hyperparameter optimization, and extended training iterations holds promise for substantial improvement. These refinements aim to enhance the agent's proficiency in surgical environments and task performance. As you can see in the treining reward trend line results does not converge to the optimal solution.

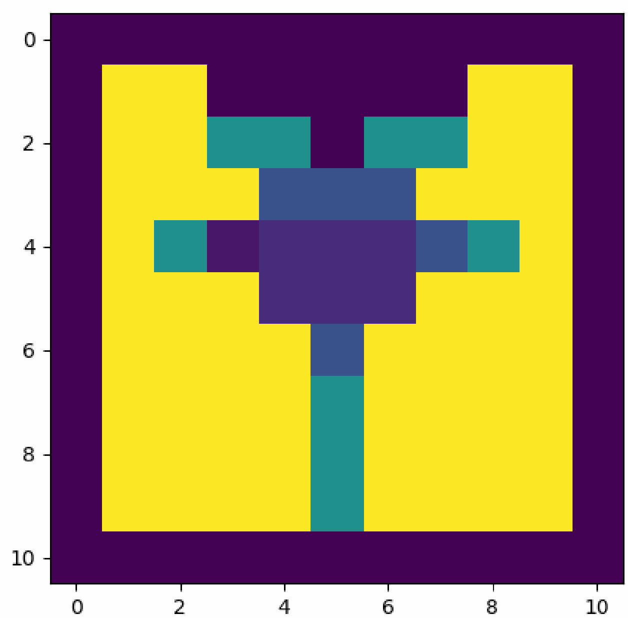

Figure 5: Trained Agent

## References

## References

[1] Masakazu Sato, Kaori Koga, Tomoyuki Fujii, and Yutaka Osuga. *Can Reinforcement Learning Be Applied to Surgery?*