Can Reinforcement Learning Be Applied to Surgery?

Masakazu Sato, Kaori Koga, Tomoyuki Fujii and Yutaka Osuga

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.76146

Abstract

Background: Remarkable progress has recently been made in the field of artificial intelligence (AI).

Objective: We sought to investigate whether reinforcement learning could be used in surgery in the future.

Methods: We created simple 2D tasks (Tasks 1–3) that mimicked surgery. We used a neural network library, Keras, for reinforcement learning. In Task 1, a Mac OS X with an 8 GB memory (MacBook Pro, Apple, USA) was used. In Tasks 2 and 3, a Ubuntu 14.04LTS with a 26 GB memory (Google Compute Engine, Google, USA) was used.

Results: In the task with a relatively small task area (Task 1), the simulated knife finally passed through all the target areas, and thus, the expected task was learned by AI. In contrast, in the task with a large task area (Task 2), a drastically increased amount of time was required, suggesting that learning was not achieved. Some improvement was observed when the CPU memory was expanded and inhibitory task areas were added (Task 3).

Conclusions: We propose the combination of reinforcement learning and surgery. Application of reinforcement learning to surgery may become possible by setting rules, such as appropriate rewards and playable (operable) areas, in simulated tasks.

Keywords: hysterectomy, robotic surgical procedures, deep learning, artificial intelligence, deep Q network (DQN), reinforcement learning

1. Introduction

Remarkable progress has recently been made in the field of artificial intelligence (AI), and new advancements are made almost every day [1–3]. Although it is not possible to discuss
all developments in AI, such as image recognition, automatic driving or trials of the game of Go, a particularly promising AI task is game playing [4–6]. AI can be trained to play games, such as the game Breakout, and can perform as well as an experienced human subject. In the context of game play, reinforcement learning is a common strategy [6, 7]. For instance, the objective of the game Breakout is to break blocks to score points. While the performance of the player at the instance of breaking a block should not necessarily be evaluated, the performance immediately prior to successfully breaking a block, such as the player’s ability to direct the ball, should be evaluated. With this aim, the AI calculates the expected values of the scores (reward) that could be obtained by various actions and learns how to achieve the maximum reward via reinforcement learning [6]. Thus, the AI succeeds in breaking more blocks and obtaining higher scores.

The concept of reinforcement learning has a commonality with surgical procedures. A hysterectomy, which is one of the most basic gynecologic procedures [8], can be considered as an example. The aim of a hysterectomy is to enucleate the uterus; however, there are no specific guidelines regarding precisely how the knife should be used at the moment that the uterus is being removed. Rather, resecting the uterine arteries in advance, for instance, may reduce total blood loss, and this possibility should be evaluated. In addition, the resection of ligaments should also be evaluated, as this step precedes extraction of the uterus.

Altogether, we simulated and investigated whether reinforcement learning could be applied to surgery and whether AI could be possibly used to perform surgeries in the future.

In this study, we created simple 2D tasks that mimicked surgery (such as a hysterectomy simulation) and investigated whether the surgery could be performed as expected via reinforcement learning. During the task, the player (an imaginary knife) moved around the task areas. The player scored when passing target areas, such as imaginary ligaments and arteries, and lost points for other actions. The task was over when the player reached the uterus or moved outside the task area. We established rules and observed the process during which movement of the imaginary knife was learned and improved. In the task with a relatively small task area (Task 1), the knife finally passed all target areas, and the expected learning was achieved. In contrast, in the task with a large task area (Task 2), significantly more time was required to complete the task, suggesting that learning was not achieved. We addressed this problem by expanding the CPU memory and adding inhibitory task areas (Task 3), and some improvements were observed.

In this study, we applied the concept of reinforcement learning to surgical procedures and identified some commonalities between reinforcement learning and the way surgeons approach an operation. We found that various aspects of the techniques of efficient learning should be developed for applying reinforcement learning to surgery. These aspects include choosing a model that closely reproduces the surgical scene using a high-performance computer for deep learning and tuning of neural networks. Surgeons will need to understand the rules, such as when a reward can be earned and where the agent is allowed to move. Efficient learning would be possible by integrating these rules into the AI by engineers.
2. Materials and methods

2.1. 2D tasks

To create the 2D task, a publicly available game code for the game “Snake” was referenced and tuned (https://github.com/farizrahman4u/qlearning4k). The code for the operation task is provided in the S7 Text. The snake game was used for this investigation because it could be easily modified to the task we needed. Therefore, in particular, the original code should not necessarily be the snake game if one could create the task which mimics surgery.

2.2. Task 1

We created 9 × 9 squares of a simple 2D task representing a surgery (a hysterectomy) as shown in Figure 1A. The goal of the task is to remove the object (uterus, yellow); indirect goals, such as cutting ligaments and arteries, were also established. In the context of surgery, there are preferred cutting points that consider the densities of the arteries and the distance from the object, and these points were set as target points (green). Thus, the objective was for the tip of the knife (orange) to pass (cut) all the target points. In other words, if surgeons see the mass (tumors or uterus what so ever) to remove like yellow areas in Figure 1A, then they would cut the ligament (green areas) by using knife as drawing a circle, without approaching the area where the bleeding is expected to occur (red). Rewards, such as +1 for passing green areas, −1 for passing red areas, and 0 for passing blue areas (peritoneum), were defined. In addition, −1

![Figure 1. Creation of task. (A) Rules of the task. The name and reward of each area were as follows: orange, tip of knife; dark blue, path where tip of knife had passed; red, ligament or artery, −1; green, target area, +1; yellow, uterus, −1 (task over); blue, peritoneum, 0; brown, out of area, −1 (task over). Starting point was located in the upper middle of area. (B) Results of learning. The results after 500 epochs of learning (upper) and 1500 epochs of learning (lower) are shown. The knife passed all the green points in the shortest amount of time after learning.](image-url)
was given and the task was defined as over if the knife passed over the yellow area (uterus) or if the knife passed outside the task area. Finally, −1 was given for passing areas were previously passed.

2.3. Task 2

We extended the task to $48 \times 48$ squares. This imposed significantly heavier burdens on the computer compared to Task 1, and we extended the CPU memory as described below.

2.4. Task 3

We added inhibitory areas (which would trigger the task to be over) to Task 2, and named it Task 3.

2.5. Deep Q network

We used a neural network library, Keras, for the reinforcement learning (https://keras.io) and TensorFlow as its backend (https://www.tensorflow.org) [9]. The codes for the agent can be found at https://github.com/farizrahman4u/qlearning4k.

The code simply obey the rules’ $Q(S, a) = r + \text{gamma} \cdot Q(S', a')$.

$Q(S, a)$ means the maximum score the agent will get by the end of the game, if it does action $a$, when the game is in state $S$. On performing action $a$, the game will jump to a new state $S'$, giving the agent an immediate reward $r$.

In short, $Q(S, a) = \text{Immediate reward from this state} + \text{Max-score from the next state onwards}$.

Gamma is a discounting factor to give more weight to the immediate reward.

We are gynecologists and understand the very minimum of deep q learning, and we just tuned nb_frames (the number of frames should the agent remeber), batch_size (batch size for training), gamma (discounting factor) and nb_epoch (number of epochs to train). The original neural networks contained two convolutional layers and two dense layers. We slightly tuned the network by adding convolutional layers or changing the size of frames, but we did not see much improvement. Thus, the readers might find a more efficient way of learning. However, the objective of the present study is to propose the combination of reinforcement learning and surgery, and we did not investigate the tuning any further after obtaining conclusions.

2.6. Development environment

The development environments used in this study were Python 2.7.12, Keras 1.1.0, TensorFlow 0.8.0, and Matplotlib1.5.3. In Task 1, a Mac OS X with an Intel Core i5 processor and 8 GB memory (MacBook Pro, Apple, USA) was used. In Tasks 2 and 3, a Ubuntu 14. 04LTS with an Intel Xenon E5 v2 processor and 26 GB memory (Google Compute Engine, Google, USA) was used.
3. Results

3.1. Learning process and trials for appropriate learning

We created 9 × 9 squares of a simple 2D task representing the surgical scene (Figure 1A). We provide representative movies after 10 epochs, 500 epochs and 1500 epochs of learning (Figure 1B and Movies 1–3). The knife passed all the green points in the shortest amount of time after learning.

We next extended the task to 48 × 48 squares (Movies 4 and 5). This imposed significantly heavier burdens on the computer than in Task 1, and therefore, we expanded the CPU memory as described in the Section 2. The knife was located in limited areas after 500 epochs of learning (Movie 4), and the area of movement was extended after 1500 epochs of learning (Movie 5). Even after extending the CPU memory, the task required a long amount of time, and thus, learning was not necessarily achieved. Thus, we modified Task 2 to Task 3.

In Task 2, the knife moved to areas where no reward could be obtained (upper right area), which was hypothesized to prevent appropriate learning (Movie 5). We considered that some areas are not significant in an actual hysterectomy and added inhibitory areas (Task 3) that would trigger the task to end (dark blue areas in Movie 6). By adding these inhibitory areas, the knife reached the lower areas where it had never reached after 1500 epochs of learning (Movie 6). However, even with these limitations, a long time for learning was required within the simulated environments. These results suggest that efficient learning can be achieved by setting appropriate rules such as adding inhibitory areas. All video materials referenced in this section are available at: https://www.intechopen.com/download/index/process/160/authtkey/f5524cabe6c92eff86119b2f0c58ba26

4. Discussion

In this study, we simulated and investigated whether reinforcement learning could be applied to surgery in the future and evaluated the types of hurdles that may exist.

Progress in AI has been remarkable, and AI is currently an essential part of the technology industry. Recent progress in AI can be attributed to the development of deep learning. Deep learning is a form of machine learning featured by the characteristic that the user does not have to choose features or representations while inputting data [10]. Deep learning is currently widely used in image recognition and audio recognition. While playing games with reinforcement learning has also been investigated, reinforcement learning has recently been combined with deep learning, resulting in drastic improvements in performance [4–6].

Thus, we sought to investigate whether reinforcement learning could be used in surgery in the future and developed appropriate simulations. Our team is composed of clinicians rather than engineers; however, we performed this study after studying deep learning and reinforcement learning.
In reinforcement learning, it is preferable to obtain a reward during the immediately following action rather than having expectations to obtain rewards in distant future actions [6]. This can be understood by the results of Task 1, in which the knife passed all the green areas in the shortest possible time, even though a rule never inhibited a more roundabout approach (Figure 1B). This can be expected to as a way to shorten the surgical time. Furthermore, choosing the shortest route was also of interest. The blue and red arrows in the circle in Figure 2 represent the shortest routes; the route of the blue arrows was chosen after several trials (Figure 1B). The method used to decide the next action in reinforcement learning is maximizing the expected values that would be obtained after subsequent actions [6]. Therefore, actions that have greater expected values are likely to be preferred. Therefore, we expected that the red arrows were not preferred because points would be deducted (red area) or the task would be ended (uterus) with subsequent actions. This approach is also considered superior in surgery. In other words, avoiding areas associated with point deductions would result in retaining the appropriate margins during surgery. Therefore, the characteristics of reinforcement learning seemed to be compatible with typical surgical approaches.

The other results obtained were expected. It was concluded that reinforcement learning could solve a simple 2D task; 3D models that more closely replicate surgery should be further considered. In addition, increasing the playable area provides significantly more options for actions, and high-performance computers and tuning of neural networks will be needed for more complex tasks. 3D operation models are currently being developed and are used in
practice for endoscopic surgery simulation [11–14]. The speed at which high-performance computers are being developed is astonishing, and deep learning is being thoroughly investigated by engineers both in academia and business [15–17]. Combining the progress in both fields may provide designs allowing AI to realistically perform surgery. In such areas, clinicians can contribute the essential aspects of the surgery, that is, the playable task areas and the appropriate scores. Although increasing reward areas or limiting playable areas would speed learning time, such restrictions could also prevent the AI from finding alternative paths. Thus, clinicians and engineers should work cooperatively define rules.

In this study, we applied the concept of reinforcement learning to surgical procedures and identified common points between reinforcement learning and the way surgeons approach an operation. Reinforcement learning is now exclusively used and studied in the context of game play; however, it could also be applied to performing surgeries now that robotic surgery is widely available. Although there are many hurdles to overcome, AI could be applied to surgery by setting the appropriate rules, such as defining rewards and playable (operable) areas. To realize this goal, it is important for clinicians to further study deep learning and reinforcement learning strategies.

Acknowledgements

We greatly appreciate American Journal Experts for their generous help in editing the manuscript.

Conflict of interest

Authors declare there is no conflict of interests.

Supporting information

**Movie 1.** Result of Game 1 after 10 epochs of learning.

**Movie 2.** Result of Game 1 after 500 epochs of learning.

**Movie 3.** Result of Game 1 after 1500 epochs of learning.

**Movie 4.** Result of Game 2 after 500 epochs of learning.

**Movie 5.** Result of Game 2 after 1500 epochs of learning.

**Movie 6.** Result of Game 3 after 1500 epochs of learning.

https://www.intechopen.com/download/index/process/160/authkey/f5524cabe6c92eff86119b2f0c58ba26
S7 Text. The game codes for Game 1, 2 and 3 (python script). The code of agent.py, memory.py and game.py could be obtained in the link in Section 2.

```python
# gamel_test.py

from keras.models import Sequential
from keras.layers import *
from game1 import OP
from keras.optimizers import *
from agent import Agent

grid_size = 11
nb_frames = 5
nb_actions = 5

model = Sequential()
model.add(Convolution2D(16, 2, 2, activation='relu', input_shape=(nb_frames, grid_size, grid_size)))
model.add(Convolution2D(32, 3, 3, activation='relu'))
model.add(Flatten())
model.add(Dense(256, activation='relu'))
model.add(Dense(nb_actions))
model.compile(RMSprop(), 'MSE')

operation = OP(grid_size)

agent = Agent(model=model, memory_size=-1, nb_frames=nb_frames)
agent.train(operation, batch_size=64, nb_epoch=500, gamma=0.8)
agent.play(operation)

# gamel.py

__original_author__ = "Fariz Rahman"
# https://github.com/farizrahman4u/qlearning4k

# Modified by "Masakazu Sato" on 10th December, 2016.
import numpy as np
from game import Game

actions = {0:'left', 1:'right', 2:'up', 3:'down', 4:'idle'}
forbidden_moves = [(0, 1), (1, 0), (2, 3), (3, 2)]

class OP(Game):

    def __init__(self, grid_size=11, knife_length=1):
        self.grid_size = grid_size
        self.knife_length = knife_length
        self.reset()
        self.state_changed = True
```
@property
def name(self):
    return "OP"

@property
def nb_actions(self):
    return 5

def play(self, action):
    assert action in range(5), "Invalid action."
    self.plus_scored = False
    self.minus_scored = False
    self.non_scored = False
    self.moveKnife(action)
    tip = self.knife[0]

    if self.hit_artery():
        self.minus_scored = True
        self.art_border.remove(tip)
    elif self.hit_marker():
        self.plus_scored = True
        self.mark_border.remove(tip)
    elif self.hit_peritoneum():
        self.non_scored = True
        self.per_border.remove(tip)
    elif self.hit_border():
        self.game_over = True
    else:
        self.minus_scored = True

def arrange(self):
    self.artery = (1,4),(1,6),(2,4),(2,6),(4,1),(4,9),(5,1),(5,3),(5,7),
    (5,9),(9,4),(9,6)
    self.marker = (3,4),(3,6),(4,2),(4,3),(4,7),(4,8),(5,2),(5,8),(7,4),
    (7,6),(8,4),(8,6)
    self.uterus = (4,4),(4,5),(4,6),(5,4),(5,5),(5,6),(6,4),(6,5),(6,6)
    self.peritoneum = (1,1),(1,2),(1,3),(1,5),(1,7),(1,8),(1,9),(2,1),
    (2,2),(2,5),(2,7),(2,8),(2,9),(3,1),(3,2),(3,3),(3,5),(3,7),
    (3,8),(3,9),(6,1),(6,2),(6,3),(6,7),(6,8),(6,9),(7,1),(7,2),
    (7,3),(7,5),(7,7),(7,8),(7,9),(8,1),(8,2),(8,3),(8,5),(8,7),
    (8,8),(8,9),(9,1),(9,2),(9,3),(9,5),(9,7),(9,8),(9,9)

def new_arrange(self):
    self.artery = (1,4),(1,6),(2,4),(2,6),(4,1),(4,9),(5,1),(5,3),(5,7),
    (5,9),(9,4),(9,6)
    self.marker = (3,4),(3,6),(4,2),(4,3),(4,7),(4,8),(5,2),(5,8),(7,4),
    (7,6),(8,4),(8,6)
    self.uterus = (4,4),(4,5),(4,6),(5,4),(5,5),(5,6),(6,4),(6,5),(6,6)
    self.peritoneum = (1,1),(1,2),(1,3),(1,5),(1,7),(1,8),(1,9),(2,1),
    (2,2),(2,5),(2,7),(2,8),(2,9),(3,1),(3,2),(3,3),(3,5),(3,7),
    (3,8),(3,9)
(6,1), (6,2), (6,3), (6,7), (6,8), (6,9), (7,1), (7,2), (7,3), (7,5), (7,7), (7,8), (7,9), (8,1), (8,2), (8,3), (8,5), (8,7), (8,8), (8,9), (9,1), (9,2), (9,3), (9,5), (9,7), (9,8), (9,9)

    self.art_border = []
    self.art_border += self.artery
    self.mark_border = []
    self.mark_border += self.marker
    self.per_border = []
    self.per_border += self.peritoneum

def move_knife(self, action):
    if action == 4 or (action, self.previous_action) in forbidden_moves:
        action = self.previous_action
    else:
        self.previous_action = action
    head = self.knife[0]
    if action == 0:
        p = (head[0] - 1, head[1])
    elif action == 1:
        p = (head[0] + 1, head[1])
    elif action == 2:
        p = (head[0], head[1] - 1)
    elif action == 3:
        p = (head[0], head[1] + 1)
    self.knife.insert(0, p)
    self.knife.pop()

def get_state(self):
    canvas = np.ones((self.grid_size, ) * 2)
    canvas[1:-1, 1:-1] = 0.0
    for seg in self.knife:
        canvas[seg[0], seg[1]] = 0.8
    for seg in self.art_border:
        canvas[seg[0], seg[1]] = 0.9
    for seg in self.mark_border:
        canvas[seg[0], seg[1]] = 0.5
    for seg in self.per_border:
        canvas[seg[0], seg[1]] = 0.7
    return canvas

def get_score(self):
    if self.game_over:
        score = -3
    elif self.plus_scored:
        score = +3
    elif self.minus_scored:
        score = -2
    else:
        score = +1
else:
    score = -1
return score

def reset(self):
    grid_size = self.grid_size
    knife_length = self.knife_length
    self.knife = [(3,5)]
    self.game_over = False
    self.plus_scored = False
    self.minus_scored = False
    self.non_scored = False
    self.arrange()
    if np.random.randint(2) == 0:
        self.previous_action = 0
    else:
        self.previous_action = 1
    self.border = []
for z in range(grid_size):
    self.border += [(z, 0), (z, grid_size - 1), (0, z), (grid_size - 1, z)]
self.ut_border = []
self.ut_border += self.uterus
self.art_border = []
self.art_border += self.artery
self.mark_border = []
self.mark_border += self.marker
self.per_border = []
self.per_border += self.peritoneum

def left(self):
    self.play(0)
def right(self):
    self.play(1)
def up(self):
    self.play(2)
def down(self):
    self.play(3)
def idle(self):
    self.play(4)

def hit_border(self):
    return self.knife[0] in self.border or self.knife[0] in self.ut_border
def hit_artery(self):
    return self.knife[0] in self.art_border
def hit_marker(self):
    return self.knife[0] in self.mark_border
def hit_peritoneum(self):
    return self.knife[0] in self.per_border

def is_over(self):
    return self.hit_border()
def is_won(self):
    return len(self.mark_border) < 1
# game2_test.py

from keras.models import Sequential
from keras.layers import *
from game2 import OP
from keras.optimizers import *
from agent import Agent

grid_size = 50
nb_frames = 5
nb_actions = 5

model = Sequential()
model.add(Convolution2D(16, 2, 2, activation='relu', input_shape=(nb_frames, grid_size, grid_size)))
model.add(Convolution2D(32, 3, 3, activation='relu'))
model.add(Flatten())
model.add(Dense(256, activation='relu'))
model.add(Dense(nb_actions))
model.compile(RMSprop(), 'MSE')

operation = OP(grid_size)

agent = Agent(model=model, memory_size=5, nb_frames=nb_frames)
agent.train(operation, batch_size=64, nb_epoch=500, gamma=0.80)
agent.play(operation)

# game2.py
__original_author__ = "Pariz Rahman"
@Modified by "Masakazu Sato" on 10th December, 2016.
import numpy as np
from game import Game

actions = {0:'left', 1:'right', 2:'up', 3:'down', 4:'idle'}
forbidden_moves = [(0, 1), (1, 0), (2, 3), (3, 2)]

class OP(Game):
    def __init__(self, grid_size=50, knife_length=1):
        self.grid_size = grid_size
        self.knife_length = knife_length
        self.reset()
        self.state_changed = True

@property
    def name(self):
        return "OP"

@property
    def nb_actions(self):
        return 5
Can Reinforcement Learning Be Applied to Surgery?

http://dx.doi.org/10.5772/intechopen.76146

def play(self, action):
    assert action in range(5), "Invalid action."
    self.plus_scored = False
    self.minus_scored = False
    self.non_scored = False
    self.move_knife(action)
    tip = self.knife[0]

    if self.hit_artery():
        self.minus_scored = True
        self.art_border.remove(tip)
    elif self.hit_marker():
        self.plus_scored = True
        self.mark_border.remove(tip)
    elif self.hit_peritoneum():
        self.non_scored = True
        self.per_border.remove(tip)
    elif self.hit_border():
        self.game_over = True
    else:
        self.minus_scored = True

def arrange(self):
    self.artery = [1, 21], [1, 22], [2, 21], [2, 22], [3, 21], [3, 22], [4, 21], [4, 22],
                   [5, 21], [5, 22], [6, 21], [6, 22], [7, 21], [7, 22], [8, 21], [8, 22], [1, 27], [1, 28], [2, 27], [2, 28], [3, 27], [3, 28], [4, 27], [4, 28], [5, 27], [5, 28], [6, 27], [6, 28], [7, 27], [7, 28], [8, 27], [8, 28], [11, 21], [11, 22], [11, 23], [12, 21], [12, 22], [12, 23], [13, 21], [13, 22], [13, 23], [14, 21], [14, 22], [14, 23], [15, 21], [15, 22], [15, 23], [16, 21], [16, 22], [16, 23], [11, 26], [11, 27], [11, 28], [12, 26], [12, 27], [12, 28], [13, 26], [13, 27], [13, 28], [14, 26], [14, 27], [14, 28], [15, 26], [15, 27], [15, 28], [16, 26], [16, 27], [16, 28], [21, 1], [21, 2], [22, 1], [22, 2], [23, 1], [23, 2], [24, 1], [24, 2], [25, 1], [25, 2], [26, 1], [26, 2], [27, 1], [27, 2], [28, 1], [28, 2], [29, 2], [23, 3], [23, 4], [23, 5], [24, 3], [24, 4], [24, 5], [25, 3], [25, 4], [25, 5], [26, 3], [26, 4], [26, 5], [26, 6], [24, 7], [24, 8], [24, 9], [25, 6], [25, 7], [25, 8], [25, 9], [23, 12], [23, 13], [23, 14], [24, 12], [24, 13], [24, 14], [25, 12], [25, 13], [25, 14], [26, 12], [26, 13], [26, 14], [21, 15], [21, 16], [22, 15], [22, 16], [23, 15], [23, 16], [24, 15], [24, 16], [25, 15], [25, 16], [26, 15], [26, 16], [27, 15], [27, 16], [28, 15], [28, 16], [21, 33], [21, 34], [22, 33], [22, 34], [23, 33], [23, 34], [24, 33], [24, 34], [25, 33], [25, 34], [26, 33], [26, 34], [27, 33], [27, 34], [28, 33], [28, 34], [29, 35], [29, 36], [29, 37], [24, 35], [24, 36], [24, 37], [25, 35], [25, 36], [25, 37], [26, 35], [26, 36], [26, 37], [24, 40], [24, 41], [24, 42], [24, 43], [25, 40], [25, 41], [25, 42], [25, 43], [23, 44], [23, 45], [23, 46], [24, 44], [24, 45], [24, 46], [25, 44], [25, 45], [25, 46], [26, 44], [26, 45], [26, 46], [21, 47], [21, 48], [22, 47], [22, 48], [23, 47], [23, 48], [24, 47], [24, 48], [25, 47], [25, 48], [26, 47], [26, 48], [27, 47], [27, 48], [28, 47], [28, 48], [33, 21], [33, 22], [33, 23], [34, 21], [34, 22], [34, 23], [35, 21], [35, 22], [35, 23], [40, 21], [40, 22], [40, 23], [41, 21], [41, 22], [41, 23], [42, 21], [42, 22], [42, 23], [43, 21], [43, 22], [43, 23], [44, 21], [44, 22], [44, 23], [45, 21], [45, 22], [45, 23], [46, 21], [46, 22], [46, 23], [47, 21], [47, 22], [47, 23], [48, 21], [48, 22], [48, 23], [43, 26], [33, 27], [33, 28], [34, 26], [34, 27], [34, 28], [35, 26], [35, 27], [35, 28], [40, 26], [40, 27], [40, 28], [41, 26], [41, 27], [41, 28], [42, 26], [42, 27], [42, 28], [43, 26], [43, 27], [43, 28], [44, 26], [44, 27], [44, 28], [45, 26], [45, 27], [45, 28], [46, 26], [46, 27], [46, 28], [47, 26], [47, 27], [47, 28], [48, 26], [48, 27], [48, 28]
Can Reinforcement Learning Be Applied to Surgery?

http://dx.doi.org/10.5772/intechopen.76146
Can Reinforcement Learning Be Applied to Surgery?

http://dx.doi.org/10.5772/intechopen.76146
def move_knife(self, action):
    if action == 4 or (action, self.previous_action) in forbidden_moves:
        action = self.previous_action
    else:
        self.previous_action = action
    head = self.knife[0]
    if action == 0:
        p = (head[0] - 1, head[1])
    elif action == 1:
        p = (head[0] + 1, head[1])
    elif action == 2:
        p = (head[0], head[1] - 1)
    elif action == 3:
        p = (head[0], head[1] + 1)
    self.knife.insert(0, p)
    self.knife.pop()}

def get_state(self):
    canvas = np.ones((self.grid_size, ) * 2)
    canvas[1:-1, 1:-1] = 0.0
    for seg in self.knife:
        canvas[seg[0], seg[1]] = 0.8
    for seg in self.art_border:
        canvas[seg[0], seg[1]] = 0.9
    for seg in self.mark_border:
        canvas[seg[0], seg[1]] = 0.5
    for seg in self.ut_border:
        canvas[seg[0], seg[1]] = 0.7
    for seg in self.per_border:
        canvas[seg[0], seg[1]] = 0.1
    return canvas

def get_score(self):
    if self.game_over:
        score = -1
    elif self.plus_scored:
        score = +1
    elif self.minus_scored:
        score = -1
elif self.non_scored:
    score = 0
else:
    score = -1
return score

def reset(self):
    grid_size = self.grid_size
    knife_length = self.knife_length
    self.knife = [(10,24)]
    self.game_over = False
    self.plus_scored = False
    self.minus_scored = False
    self.non_scored = False
    self.arrange()
    if np.random.randint(2) == 0:
        self.previous_action = 0
    else:
        self.previous_action = 1
    self.border = []
    for z in range(grid_size):
        self.border += [(z, 0), (z, grid_size - 1), (0, z), (grid_size - 1, z)]
    self.ut_border = []
    self.ut_border += self.uterus
    self.art_border = []
    self.art_border += self.artery
    self.mark_border = []
    self.mark_border += self.marker
    self.per_border = []
    self.per_border += self.peritoneum

def left(self):
    self.play(0)
def right(self):
    self.play(1)
def up(self):
    self.play(2)
def down(self):
    self.play(3)
def idle(self):
    self.play(4)

def hit_border(self):
    return self.knife[0] in self.border or self.knife[0] in self.ut_border
def hit_artery(self):
    return self.knife[0] in self.art_border
def hit_marker(self):
    return self.knife[0] in self.mark_border
def hit_peritoneum(self):
    return self.knife[0] in self.per_border

def is_over(self):
    return self.hit_border()

def is_won(self):
    return len(self.mark_border) < 1

# game3_test.py
from keras.models import Sequential
from keras.layers import *
from game3 import OP
from keras.optimizers import *
from agent import Agent

grid_size = 50
nb_frames = 5
nb_actions = 5

model = Sequential()
model.add(Convolution2D(16, 2, 2, activation='relu', input_shape=(nb_frames, grid_size, grid_size)))
model.add(Convolution2D(32, 3, 3, activation='relu'))
model.add(Flatten())
model.add(Dense(256, activation='relu'))
model.add(Dense(nb_actions))
model.compile(RMSprop(), 'MSE')

operation = OP(grid_size)

agent = Agent(model=model, memory_size=5, nb_frames=nb_frames)
agent.train(operation, batch_size=64, nb_epoch=500, gamma=0.80)
agent.play(operation)

# game3.py
__original_author__ = "Fariz Rahman"

# Modified by "Masakazu Sato" on 10th December, 2016.
import numpy as np
from game import Game

actions = {0:'left', 1:'right', 2:'up', 3:'down', 4:'idle'}
forbidden_moves = [(0, 1), (1, 0), (2, 3), (3, 2)]

class OP(Game):
    def __init__(self, grid_size=50, knife_length=1):
        self.grid_size = grid_size
Can Reinforcement Learning Be Applied to Surgery?

http://dx.doi.org/10.5772/intechopen.76146
Can Reinforcement Learning Be Applied to Surgery?

http://dx.doi.org/10.5772/intechopen.76146
Can Reinforcement Learning Be Applied to Surgery?

http://dx.doi.org/10.5772/intechopen.76146
def move_knife(self, action):
    if action == 4 or (action, self.previous_action) in forbidden_moves:
        action = self.previous_action
    else:
        self.previous_action = action
    head = self.knife[0]
    if action == 0:
        p = (head[0] - 1, head[1])
    elif action == 1:
        p = (head[0] + 1, head[1])
    elif action == 2:
        p = (head[0], head[1] - 1)
    elif action == 3:
        p = (head[0], head[1] + 1)
    self.knife.insert(0, p)
    self.knife.pop()

def get_state(self):
    canvas = np.ones((self.grid_size, ) * 2)
    canvas[1:-1, 1:-1] = 0.0
    for seg in self.knife:
        canvas[seg[0], seg[1]] = 0.8
for seg in self.art_border:
    canvas[seg[0], seg[1]] = 0.9
for seg in self.mark_border:
    canvas[seg[0], seg[1]] = 0.5
for seg in self.ut_border:
    canvas[seg[0], seg[1]] = 0.7
for seg in self.per_border:
    canvas[seg[0], seg[1]] = 0.1
return canvas

def get_score(self):
    if self.game_over:
        score = -1
    elif self.plus_scored:
        score = +1
    elif self.minus_scored:
        score = -1
    elif self.non_scored:
        score = 0
    else:
        score = -1
    return score

def reset(self):
    grid_size = self.grid_size
    knife_length = self.knife_length
    self.knife = [(10, 24)]
    self.game_over = False
    self.plus_scored = False
    self.minus_scored = False
    self.non_scored = False
    self.arrange()
    if np.random.randint(2) == 0:
        self.previous_action = 0
    else:
        self.previous_action = 1
    self.border = []
    for z in range(grid_size):
        self.border += [(z, 0), (z, grid_size - 1), (0, z), (grid_size - 1, z)]
    self.border += self.preborder
    self.ut_border = []
    self.ut_border += self.uterus
    self.art_border = []
    self.art_border += self.artery
    self.mark_border = []
    self.mark_border += self.marker
    self.per_border = []
    self.per_border += self.peritoneum
def left(self):
    self.play(0)
def right(self):
    self.play(1)
def up(self):
    self.play(2)
def down(self):
    self.play(3)
def idle(self):
    self.play(4)

def hit_border(self):
    return self.knife[0] in self.border or self.knife[0] in self.ut_border
def hit_artery(self):
    return self.knife[0] in self.art_border
def hit_marker(self):
    return self.knife[0] in self.mark_border
def hit_peritoneum(self):
    return self.knife[0] in self.per_border

def is_over(self):
    return self.hit_border()
def is_won(self):
    return len(self.mark_border) < 1

Author details
Masakazu Sato*, Kaori Koga, Tomoyuki Fujii and Yutaka Osuga
*Address all correspondence to: masakasatou-ty@umin.ac.jp

Department of Obstetrics and Gynecology, Graduate School of Medicine, The University of Tokyo, Bunkyo-ku, Tokyo, Japan

References

[1] Ravi D, Wong C, Deligianni F, Berthelot M, Andreeu Perez J, Lo B, et al. Deep learning for health informatics. IEEE Journal of Biomedical and Health Informatics. Jan 2017;21(1): 4-21. DOI: 10.1109/JBHI.2016.2636665. PubMed PMID: 28055930

[2] Scholkopf B. Artificial intelligence: Learning to see and act. Nature. 2015;518(7540):486-487. DOI: 10.1038/518486a. Epub 2015/02/27, PubMed PMID: 25719660

[3] Zhang YC, Kagen AC. Machine learning Interface for medical image analysis. Journal of Digital Imaging. Oct 2017;30(5):615-621. DOI: 10.1007/s10278-016-9910-0. PubMed PMID: 27730415

[4] Gibney E. DeepMind algorithm beats people at classic video games. Nature. 2015;518(7540): 465-466. DOI: 10.1038/518465a. Epub 2015/02/27, PubMed PMID: 25719643
[5] Gibney E. Google AI algorithm masters ancient game of Go. Nature. 2016;529(7587):445-446. DOI: 10.1038/529445a. Epub 2016/01/29, PubMed PMID: 26819021

[6] Mnih V, Kavukcuoglu K, Silver D, Rusu AA, Veness J, Bellemare MG, et al. Human-level control through deep reinforcement learning. Nature. 2015;518(7540):529-533. DOI: 10.1038/nature14236. PubMed PMID: 26017443

[7] Littman ML. Reinforcement learning improves behaviour from evaluative feedback. Nature. 2015;521(7553):445-451. DOI: 10.1038/nature14540. PubMed PMID: 26017443

[8] Wallace SK, Fazzari MJ, Chen H, Cliby WA, Chalas E. Outcomes and postoperative complications after hysterectomies performed for Benign compared with malignant indications. Obstetrics and Gynecology. 2016;128(3):467-475. DOI: 10.1097/AOG.0000000000001591. PubMed PMID: 27500339

[9] Rampasek L, Goldenberg A. TensorFlow: Biology's gateway to deep learning? Cell Systems. 2016;2(1):12-14. DOI: 10.1016/j.cels.2016.01.009. PubMed PMID: 27136685

[10] Schmidhuber J. Deep learning in neural networks: An overview. Neural Networks. 2015;61:85-117. DOI: 10.1016/j.neunet.2014.09.003. PubMed PMID: 25462637

[11] Beyer-Berjot L, Berdah S, Hashimoto DA, Darzi A, Aggarwal R. A virtual reality training curriculum for laparoscopic colorectal surgery. Journal of Surgical Education. 2016;73(6):932-941. DOI: 10.1016/j.jsurg.2016.05.012. Epub 2016/06/28; PubMed PMID: 27342755

[12] Khan ZA, Kamal N, Hameed A, Mahmood A, Zainab R, Sadia B, et al. SmartSIM—A virtual reality simulator for laparoscopy training using a generic physics engine. The International Journal of Medical Robotics + Computer Assisted Surgery: MRCAS. 2016;16:437. DOI: 10.1002/rcs.1771. Epub 2016/09/28; PubMed PMID: 27671920

[13] Li XL, Du DF, Jiang H. The learning curves of robotic and three-dimensional laparoscopic surgery in cervical cancer. Journal of Cancer. 2016;7(15):2304-2308. DOI: 10.7150/jca.16653. PubMed PMID: 27994668; PubMed Central PMCID: PMCPMC5166541

[14] Romero-Loera S, Cárdenas-Lailson LE, de la Concha-Bermejillo F, Crisanto-Campos BA, Valenzuela-Salazar C, Moreno-Portillo M. Skills comparison using a 2D vs. 3D laparoscopic simulator. Cirugía y Cirujanos. 2016;84(1):37-44. DOI: 10.1016/j.circen.2015.12.012. (English Edition)

[15] Kusy M, Zajdel R. Application of reinforcement learning algorithms for the adaptive computation of the smoothing parameter for probabilistic neural network. IEEE Transactions on Neural Networks and Learning Systems. 2015;26(9):2163-2175. DOI: 10.1109/TNNLS.2014.2376703. PubMed PMID: 25532211

[16] Senda K, Hattori S, Hishinuma T, Kohda T. Acceleration of reinforcement learning by policy evaluation using nonstationary iterative method. IEEE Transactions on Cybernetics. 2014;44(12):2696-2705. DOI: 10.1109/TCYB.2014.2313655. PubMed PMID: 24733037

[17] Xu B, Yang C, Shi Z. Reinforcement learning output feedback NN control using deterministic learning technique. IEEE Transactions on Neural Networks and Learning Systems. 2014;25(3):635-641. DOI: 10.1109/TNNLS.2013.2292704. PubMed PMID: 24807456
