2023 IEEE International Conference on Robotics and Automation (ICRA 2023)
May 29 - June 2, 2023. London, UK

Automatic Cell Rotation Method Based on Deep Reinforcement Learning

Huiying Gong, Yujie Zhang, Yaowei Liu, *Member, IEEE*,
Qili Zhao, *Member, IEEE*, Xin Zhao, *Member, IEEE*, Mingzhu Sun*, *Member, IEEE*,

***Abstract*— Cell rotation is widely used to adjust cell posture in sub-cellular micromanipulations. The trajectory planning of the injection micropipette is needed, so that the cells can be rotated with the minimum deformation to reduce cell damage and keep cell viability. Due to the uncertainty of cell properties and manipulation environment, it is difficult to identify the parameters of the mechanical models in traditional robotic cell rotation methods. In this paper, deep reinforcement learning is introduced into cell manipulation for the first time to perform trajectory planning of the micropipette. We first abstract the cell rotation process by using the mechanical model and microscopic vision techniques and build a cell rotation simulation environment. Then we design a reward function by combining various factors of cell rotation and implement a reinforcement learning framework based on deep Q-learning (DQL). Finally, we train the cell rotation process based on the deep reinforcement learning algorithm. The simulation results indicate the proposed DQL agent achieved an average success rate of 97% without useless exploration. Moreover, the proposed method rotated the cells in a way that causes less mechanical damage than humans, demonstrating the DRL ability for cell rotation with high efficiency and low cell damage.**

I. Introduction

Cell manipulation is widely used in various fields of life science, such as transgenic, assisted reproduction and animal cloning. Different from non-living manipulation objects, cells are the basic units of life, which makes cell manipulation a kind of complex cell surgery [1]. Therefore, we should focus on the cell damage caused by cell manipulations to keep cell viability. As a typical cell manipulation, cell rotation is the key technique for adjusting cell posture in sub-cellular micromanipulations, such as intracytoplasmic sperm injection (ICSI) [2], preimplantation genetic screening (PGS) [3], and somatic cell nuclear transfer (SCNT) [4]. In manual cell rotation, the cell is normally aspirated by a holding micropipette with negative pressure and rotated by an injection micropipette by applying friction between the cell and the pipette. This rotation method has the advantages of no additional equipment and high accuracy, compared with other rotation methods, such as in-house developed holding stage [5], [6] or electric field-driven [7], [8], magnetic field-driven [9], [10], microfluidic flow [11], [12], optical tweezer [13], [14], acoustic field [15], [16], etc.

This research was jointly supported by National Natural Science Foundation of China (62273185, 62027812, 62003174).

The authors are with the Institute of Robotics and Automatic Information System (IRAIS) and the Tianjin Key Laboratory of Intelligent Robotic (tjKLIR), Nankai University, Tianjin 300350, China and the Institute of Intelligence Technology and Robotic Systems, Shenzhen Research Institute of Nankai University, Shenzhen 518083, China.*Corresponding Author: Mingzhu Sun. (e-mail: sunmz@nankai.edu.cn).

During cell rotation, the injection micropipette's indentation depth on the cell should be large enough to overcome the resistance and be minimal to avoid possible mechanical damage. In the past decade, a series of studies have been reported to optimize the indentation for robotic cell rotation [17]–[19]. These methods calculated the appropriate indentation depth based on mechanical models and designed the trajectory of the injection micropipette, which have proven to be effective. However, the measurement of cell mechanical properties is challenging, since Young's modulus of cells is different due to the cell characteristics and health status [20], and the friction factors between different micropipettes and cells are also varied. Due to the complexity of the cell properties, it is difficult to accurately identify the parameters of the mechanical models, which limits the performance and robustness of the model-based control strategy. Besides, human experience and skills have not been used to guide robotic cell rotation.

Recently, the rapid development of reinforcement learning (RL) technology has roundly impacted the traditional control field. Aiming of maximizing the rewards obtained from the environment, the RL agents perform a series of actions and learn behavioral strategies in an unknown environment [21]. The recent appearance of deep learning using multilayer artificial neural networks has promoted the application of reinforcement learning in dealing with high-dimensional state space problems, which is called deep reinforcement learning(DRL). DRL could realize end-to-end control from high-dimensional original input to output without the mechanical model. The experimental results show that the DRL-based methods perform better than humans in tasks such as video games [22] and board games [23], and make great achievements in robot control problems such as mobile robot navigation, industrial robot grasping, and multi-legged robot motion [24]–[26]. Most recently, DRL has been introduced to control micromanipulation devices with high accuracy and robustness [27]–[29]. These DRL methods show great application potential for model-free control, which overcomes the difficulties of mechanical modeling and model parameter identification in cell manipulation.

However, it is difficult to provide sufficient and strictly controlled training data to converge the neural network, which is the main difficulty of the DRL application. On one hand, obtaining the cell manipulation data from real cell experiments would be costly in money, labor, material, and time. On the other hand, there is no reported simulation environment of cell manipulation suitable for reinforcement

learning training at present, which greatly limits the application of DRL in cell rotation.

In this paper, we introduce deep reinforcement learning to perform cell rotation for the first time. By using the mechanical model and microscopic vision techniques, we abstract the cell rotation process and build a cell rotation simulation environment, which could provide sufficient data for training the agent. To overcome the difficulty of DRL training caused by reward sparsity, we design a reward function by combining various factors of cell rotation and propose a deep reinforcement learning framework based on deep Q-learning. Simulation experiments show that the proposed method achieved an average success rate of 97% without useless exploration after 9000 episodes of learning. Meanwhile, the agent caused less stress and cell deformation than humans during cell rotation, which helps to reduce cell mechanical damage.

II. METHOD

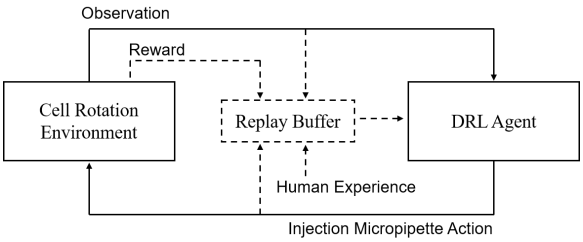

Fig. 1. Framework of cell rotation based on DRL.

The proposed method aims to perform cell rotation with high autonomy and efficiency. The framework of the method is shown in Fig. 1. Solid arrows represent the process of cell rotation using the DRL agent. At each time step t, the microscopic vision state s_t of the cell rotation environment is observed. Taking the current state s_t as input, the DRL agent predicts the expected value of each injection micropipette action and generates the optimal action a_t. The implemented action acts on the environment, changing the observation results and obtaining a reward r_t and a new state s_{t+1}. The agent then makes the next decision. Meanwhile, the replay buffer stores the operating experience of the agent and human for DRL agent training, as shown by dotted arrows.

To realize this framework, the proposed method consists of three parts. Firstly, mechanical analysis is introduced to calculate the interaction between the injection micropipette and the cell, so that the DRL agent obtains sufficient and effective operating experience. Further, cell manipulation images are represented in a simple and effective way using microscopic vision techniques. By abstracting the cell rotation process, the simulation environment of cell rotation is constructed. Finally, we train the DRL agent based on the deep Q-learning method. To improve the training performance and speed, we design an elaborate reward function and introduce human experience for guidance.

A. Mechanical Analysis of Cell Rotation

In cell rotation, the holding micropipette aspirates the target cell using an appropriate negative pressure to prevent the cell from dropping. The force applied by the injection micropipette is used to push the cell for rotation. The force needs to be sufficient to overcome the friction between the cell and the holding micropipette so that the movement of the injection needle can rotate the cell without deviation or sliding.

The cell rotation process is modeled by establishing a connection between the movement of the injection micropipette and the rotation of the cell: when the injection micropipette is in contact with the cell, a small angle of the injection micropipette movement around the cell center $\Delta\varphi_{in}$, and the corresponding angle of cell rotation $\Delta\varphi_{cell}$, are associated in a functional relationship $\Delta\varphi_{cell} = f(\Delta\varphi_{in})$. The rotation angle of the cell is reflected by the position of the polar body. Due to the occlusion of the cytoplasm, which is not transparent in the bright field of the microscope, the polar body can be observed only when it is close to the focal plane. In this paper, we assume that the polar body has appeared on the focal plane, so that the injection micropipette only needs to move on the X-Y plane to rotate the cell and make the polar body reach the target position.

Since the rotation speed of the cell is very low, the cell is considered in a state of static equilibrium. We analyze the cell rotation process using rigid body static according to the principle of solidification [30]. The force model of an ellipsoidal cell is established during cell rotation. By determining the relationship between the current rotation force exerted by the injection micropipette and the minimum rotation force required for rotation, we determine whether each small movement of the injection micropipette could rotate the cell.

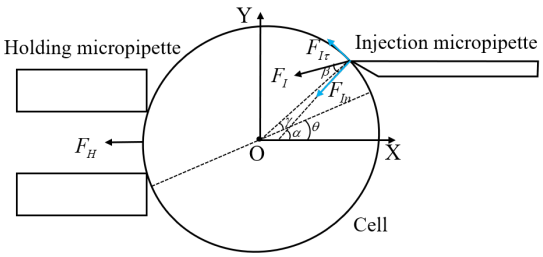

Fig. 2. Force model of in-plane ellipsoidal cell rotation.

Since the deformation of the cell is local and small during cell rotation, the relationship between the indentation force applied by the injection micropipette F_{In} and the measured indentation depth d is estimated using the Hertzian model [31]:

$$F_{In} = \frac{2}{\pi}\tan\lambda\frac{E}{1-\upsilon^2}d^2, \tag{1}$$

where λ is the half-apex angle of the injection micropipette, E is Young's modulus of the cell, and υ is Poisson's ratio with a typical value of 0.5.

As shown in Fig. 2, the friction force $F_{I\tau}$ is generated by the indentation force F_{In}, when the friction force reaches the maximum. The maximal friction F_I exerted by the injection micropipette can be derived as:

$$F_I = \frac{F_{In}}{\cos\beta}, \tag{2}$$

where β is the inclined angle of F_I and F_{In}, which can be calculated by the friction coefficient μ_I between the injection micropipette and the cell: $\beta = \arctan\mu_I$.

Based on the mechanical analysis [19], the minimum force $F_{I\_\min}$ applied by the injection micropipette, which is required to rotate the cell, is determined as:

$$F_{I\_\min} = \frac{\mu_H F_H}{\sqrt{1+\mu_H^2}\sin[\alpha - \arctan\mu_I - \arctan\mu_H]}, \tag{3}$$

where F_H is the aspiration force of the holding micropipette, which is estimated as $F_H = \pi R_{H\_in}^2 P_H$, where R_{H_in} is the inner radius of the holding micropipette, P_H is the aspiration pressure. α is the angle between F_{In} and X-axis, which is calculated as $\alpha = \arctan(\frac{a^2}{b^2}\tan\gamma) + \theta$, where a and b are the lengths of semimajor and semiminor axes of the ellipse, γ is the angle between its connecting line with the ellipse center O and the ellipse's major axis, and θ is the angle between the ellipse's major axis and the X-axis.

After calculating F_I and $F_{I\_\min}$, the relationship between $\Delta\varphi_{in}$ and $\Delta\varphi_{cell}$ is determined as:

$$\begin{cases} \Delta\varphi_{cell} = \Delta\varphi_{in}, & F_I \geq F_{I\_\min} \\ \Delta\varphi_{cell} = 0, & F_I < F_{I\_\min} \end{cases}. \tag{4}$$

B. Simulation Environment of Cell Rotation

To represent the simulation environment simply and effectively, microscopic vision techniques are used to abstract the cell rotation images. Taking the porcine oocyte manipulation as an example, the redundant information, including liquid background and impurities, is ignored, and the positions and shapes of micropipettes, the cell and the polar body are extracted. As shown in Fig. 3, the procedure is described as follows:

1) Cytoplasmic detection. Otsu's thresholding[see Fig. 3(b)] is performed on the original image [see Fig. 3(a)], then the position and radius of the cytoplasm are obtained by using Hough circle detection [see Fig. 3(c)].
2) Holding micropipette detection. After the cytoplasm image on the left is cropped, a disc structure with a size of 10 is used for morphological open operation [see Fig. 3(d)]. The contour of the holding micropipette is represented by a rectangle [see Fig. 3(e)].
3) Injection micropipette detection. Make the contour of the injection micropipette close by modifying the data in the rightmost columns of the binary image [see Fig. 3(f)]. Then the detected contours are filtered according to the aspect ratio and the area of the contours. The preserved region is considered the injection micropipette region [see Fig. 3(f)]. The upper left and lower left vertices of the region are determined to fit the micropipette contour [see Fig. 3(i)].
4) Zona pellucida detection. The edge of the zona pellucida is obtained by subtracting the contours of the cytoplasm and the holding micropipette from the Canny edge detection image [see Fig. 3(j)]. The morphological operations are used to remove the contour of the injection micropipette [see Fig. 3(k)] and close the edge of the zona pellucida [see Fig. 3(l)]. Finally, the contour of the zona pellucida is obtained by ellipse fitting after area filtering [see Fig. 3(m)].
5) Polar body detection. The polar body is small and similar to the cytoplasm in texture, which makes the traditional image processing methods not suitable for polar body detection. In this paper, an improved U-net is used to detect the polar body, which was described in detail in our previous work [32]. After obtaining the ROI image of the cell [see Fig. 3(n)], the U-net is used for prediction. The output image [see Fig. 3(o)] is then processed using binarization and non-maximum suppression to get the contour. Finally, the center position and radius of the polar body are obtained using contour calculation [see Fig. 3(p)].
6) After extracting the state information of cell rotation, the microscopic image is converted into a simple abstract image [see Fig. 3(q)].

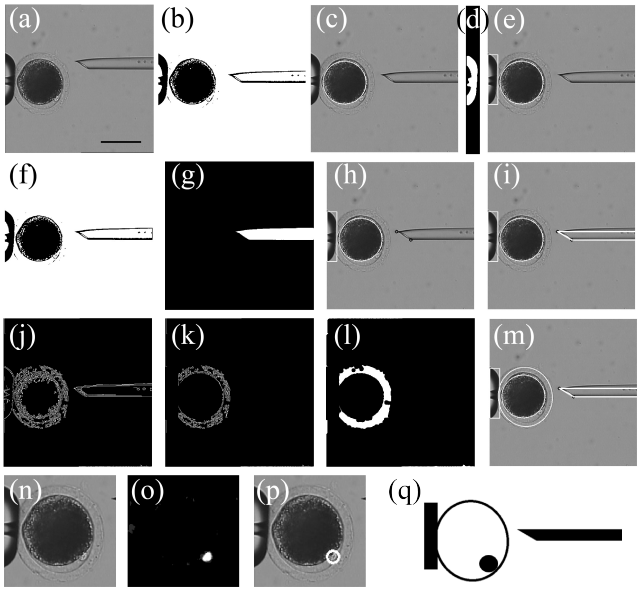

Fig. 3. Image processing and abstracting of cell rotation images. (a) Original image. (b)-(c) Cytoplasmic detection. (d)-(e) Holding micropipette detection. (f)-(i) Injection micropipette detection. (j)-(m) Zona pellucida detection. (n)-(p) Polar body detection. (q) Abstracted image. Scale bar: 100μm

Based on the converted image, a simulation environment of cell rotation is developed by imitating porcine oocyte manipulations. The shape and size of the objects are set according to reality. The holding micropipette and injection

micropipette are represented by a rectangle and right-angled trapezoid, wherein the external diameter of the holding micropipette and the injection micropipette is set to 100μm and 20μm, respectively. An elliptical cell with a short axis of [120, 130]μm and an ellipticity of [1, 1.05] is fixed on the right side of the holding micropipette and tangent to the surface of the holding micropipette. The polar body represented by a circle with a diameter of [15, 20]μm is located at the edge of the cell. The structural correlation between the simulation image and the microscopic image ensures that the rotation strategies trained in the simulation environment can be transferred to the real environment in the future.

Before cell rotation simulation, the polar body position is initialized in the random direction of the cell, and the injection micropipette is also randomly initialized on the right side of the field of vision. The operator has five discrete actions: remain still, up, down, left and right with a moving step of 1μm. The injection micropipette is expected to approach, poke and rotate the cell until the polar body is within 5° of the three o'clock direction of the cell, which is defined as a successful operation. The task is judged as failed if the rotation is not completed within 500 steps or the injection micropipette moves out of sight.

To obtain reliable interaction information on cell rotation, the relationship between the current rotation force and the required minimum rotation force is calculated in real-time at each step of the simulation. The mechanical parameters are initialized before each rotation simulation, to maintain the diversity of cell properties and manipulation environment. Notably, the simulation environment can be adjusted by modifying the mechanical parameters to model the rotation operations of various mammalian oocytes. In this study, we used the mechanical parameters of mature porcine oocytes as an example, based on the measurement data from [17]. Specifically, Young's modulus of the mature porcine oocytes is set to [8000, 12000]Pa, and the friction coefficients of the holding micropipette and the injection micropipette are set in the range of [0.04, 0.065] and [0.11, 0.2] respectively. According to the experiments, the negative pressure of the holding micropipette is set to [3000, 5000]Pa for holding and allowing cells to rotate. By varying the mechanical parameters, the simulation environment can be customized to investigate cell rotation in different cell species or under diverse experimental conditions.

C. Deep Reinforcement Learning for Cell Rotation

Considering the large-scale state space of cell rotation, a classical DRL algorithm, deep Q-learning (DQL), is used for training the agent. The term "Q" refers to the expected rewards for an action taken in a given state. DQL, which combines deep learning and reinforcement learning, uses an artificial neural network to replace Q-functions in traditional Q-learning [22], [33]. To improve the sample efficiency, the network is trained using an experience replay buffer, which stores and replays previous environment interactions experienced by the agent. After exploring the environment, this algorithm could automatically find features through the network and infer the optimal action without modeling. Fig. 4 shows the DQL network structure, all convolution layers and fully connected layer are followed by a ReLU activation function. The network inputs four consecutive frames of 80 * 80 abstracted microscopic images, and outputs the predicted Q-values of five actions. The action corresponding to the highest Q-value is regarded as the optimal action.

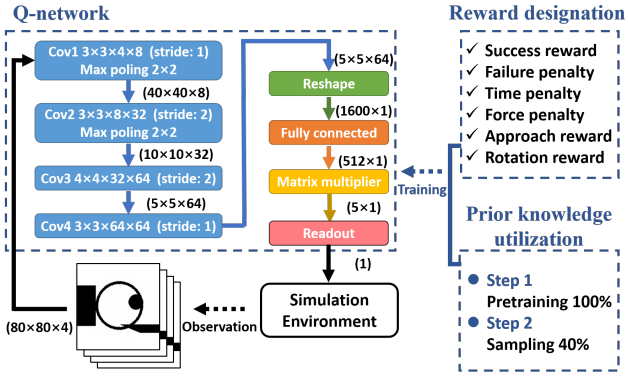

Fig. 4. Reinforcement learning framework for cell rotation based on DQL.

The RL agent is guided only by positive and negative rewards to optimize its behavior. The setting of the reward function is crucial since the goal of the RL agent is to maximize the total rewards. However, cell rotation is a sparse reward task with a large-scale state space, which makes learning difficult. To alleviate this problem, we introduce prior knowledge to design an appropriate reward function. By combining various factors of cell rotation, the reward function of one step has six parts:

$$R = R_{suc} + R_{fail} + R_{time} + R_{force} + R_{app} + R_{rot}. \quad (5)$$

Where the success reward, $R_{suc} = 20$, and failure penalty, $R_{fail} = -2$, are related to the completion of cell operation. To encourage the agent to operate faster, a constant time penalty $R_{time} = -0.01$ is added for each step. To minimize cell damage, the force penalty R_{force} is related to the rotation force F_I as $R_{force} = 50000 \cdot F_I$. Since it is difficult to describe successful cell rotation quantitatively, two intermediate rewards, including approach reward R_{app} and rotation reward R_{rot}, are introduced to alleviate the problem of sparse rewards. When the micropipette is far from the cell center (120μm), R_{app} of 0.1/μm is used to reward/punish the action. Similarly, R_{rot} of 30/rad is used to reward/punish the rotation. The designed reward function ensures continuous valuable reward feedback during cell rotation.

In the large-scale state space, it is difficult for random agents to explore strategies with high rewards. The experience replay buffer accumulates useless data, making agents easy to fall into local optimization. Inspired by imitation learning, we introduce the prior knowledge of human experts to help DQL obtain stable control effects in early training.

Firstly, the prior knowledge acquired by humans is recorded as transition sequences. Each sequence contains the old state s_t, the action a_t, the reward r_t, the new state s_{t+1} and the done signal d_t, which is considered to imply the knowledge and skills acquired by humans. The collected prior knowledge is used for two-step training. In the first step, the agent does not interact with the environment, and only uses the transition sequences to train the network for warm-up. In the second step, the agent then interacts with the environment, and utilizes the network parameters trained in the first step to initialize Q-network. In addition, the network is updated using a certain proportion of collected transition sequences, so that the agent has learned the prior knowledge continuously during training.

III. Experiment

A. Training

The DQL agent was trained and evaluated in the cell rotation simulation environment. Firstly, the field experts spent 30 minutes performing 200 episodes of cell rotation in the simulation environment, obtaining 50000 hand-crafted transition sequences for prior knowledge collection. The prior knowledge was then used for two-step DQL training, which was implemented using TensorFlow as the coding platform (Intel(R) Xeon(R) Silver 4316 CPU, Nvidia GeForce RTX3080TI GPU).

In the first training step, the experience replay buffer was filled with all the hand-crafted transitions. The agent was trained with the replay buffer. It took about 7.5 hours to train the Q-network 1000000 times with a batch of 64. The network parameters obtained in the first training step were utilized to initial network parameters for the second training step.

In the second step, the DQL agent would interact with the simulation environment to fine-tune the Q-network, including an initial random phase of 5000 frames for agent observation and an annealing phase of 200000 frames for exploration (number of frames to change from 20% random actions to 0.01% random actions). The size of the replay memory buffer was set to 50000, storing the most recent experiences collected from interactions of the agent with the environment. In our experiments, each time the network was updated using 64 transitions, 60% of which were sampled in replay memory and 40% in prior knowledge, improving the efficiency and stability of training.

B. Experimental Results and Comparison

Fig. 5 shows the mean rewards for every 100 training episodes. The purple dotted line represents the human performance level, which was calculated from the 50000 hand-crafted transition sequences, and the gray dotted line represents the random performance level. By introducing prior knowledge in steps 1 and 2, our method steadily optimized the rotation strategy with about 3000 episodes and obtained an average score of 31.06 after 9000 episodes of learning, which is greater than the human performance level of 29.62.

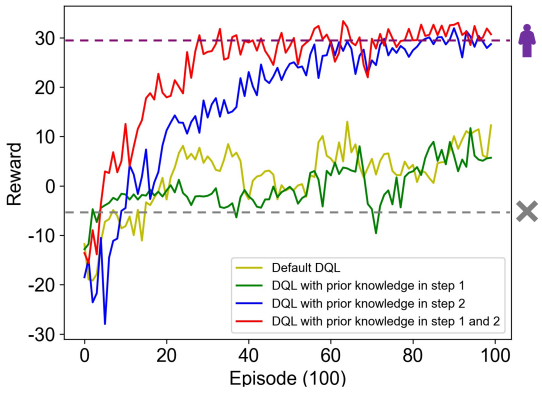

Fig. 5. Total rewards of an episode relative to the training episodes. The human performance level is indicated with a purple dotted line, and random performance level is indicated with a gray dotted line. For clarity, the results are averaged for every 100 training episodes.

We compared our method (DQL with prior knowledge in steps 1 and 2) with default DQL, DQL with prior knowledge only in step 1, and DQL with prior knowledge only in step 2. Compare with the low reward of the DQL agent without prior knowledge in step 2 (yellow line and green line), the introduced human transitions during step 2 obviously improved the performance of the DQL agent (red line and blue line) and avoided the agent falling into low reward behavior, reaching a score that is similar to or even higher than human. In addition, due to pretraining of step 1, our method (red line) achieved a high reward with about 3000 episodes (9.5 hours), while the DQL with prior knowledge only in step 2 needs double training time to reach the same score, confirming the improvement of the learning speed by prior knowledge pretraining.

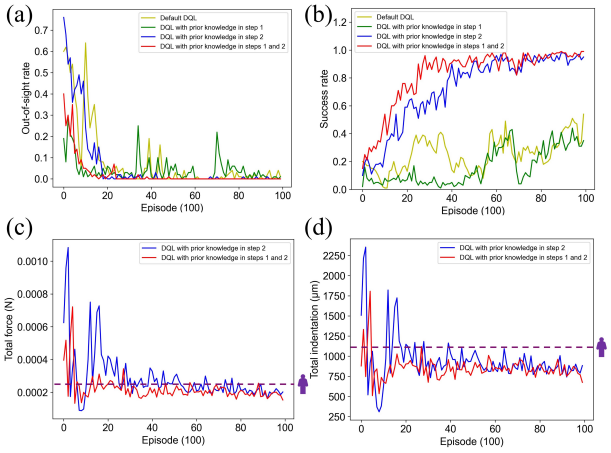

Fig. 6. Results of an episode relative to the training episodes. (a) Out-of-sight rate: the ratio of the agent to move the pipette out of the field of view. (b) Success rate. (c) Total force: accumulation of forces exerted on the cell by injection micropipette during successful rotations. (d) Total indentation: accumulation of the cell indentation during successful rotations. The human performance level is indicated with a purple dotted line. For clarity, the results are averaged for every 100 training episodes.

We evaluated the proposed cell rotation strategy using rotation efficiency and the stress exerting on the cell. As shown in Fig. 6(a) and (b), the out-of-sight rate and the success rate were calculated respectively to assess the rotation efficiency. The out-of-sight rate describes the proportion of the agent moving the micropipette out of the field of view, which is a useless exploration and will lead to rotation failure. After 3000 episodes of learning, the proposed method stopped the useless exploration and its behaviors approximately converged into stable controls. After 9000 episodes of learning, the method achieved an average out-of-sight rate of 0% and a success rate of 97% in 1000 episodes. Compared with other methods, our method showed a faster learning speed and better performance.

Our trained agent was found to adjust the position of the injection micropipette according to the initialized position of the polar body and determine the correct direction of rotation. In addition, the agent avoided contacting the cell near three o'clock, since the indentation force applied here cannot drive the cell to rotate according to mechanical analysis. These self-study skills ensure the rotation efficiency of the agent. However, the 3% failure cases typically occurred when the initial angle of the cell's polar body was far from the target direction, making it difficult for the agent to complete the rotation task within 500 steps. Additionally, the agent getting stuck in local optimal behavior can also lead to rotation failure. Encouraging the agent to select unexplored actions may help achieve better exploration-exploitation trade-offs.

To evaluate the stress and deformation caused by rotation, we calculated the force and the indentation exerted by the injection micropipette during successful rotations, and accumulated them in time steps of an episode. As shown in Fig. 6(c) and (d), the proposed method (red line) finally rotated cells in a way that causes less mechanical damage than humans, demonstrating the DRL ability for cell rotation with high efficiency and low cell damage.

IV. Conclusion

In this study, we proposed a model-free approach for cell rotation. It is the first time that DRL is introduced into cell manipulation. First, we built a cell rotation simulation environment based on the mechanical model and the abstracted microscopic images. Then we designed an elaborate reward function and introduced the prior knowledge to DQL for agent training. A series of simulation experiments were performed to evaluate the performance of the micropipette trajectory planning strategy for cell rotation. The results show that the proposed DQL agent reached an average success rate of 97% without useless exploration and rotated the cells in a way that causes less mechanical damage than humans, demonstrating the DRL ability for cell rotation with high efficiency and low cell damage. It has a guiding significance for the development of a more intelligent, more robust, and more bio-compliant micromanipulation robot. Future work will focus on the expansion of three-dimensional cell rotation and the implementation of cell rotation in the real environment by transfer learning.

References

[1] A. Shakoor, M. Xie, T. Luo, J. Hou, Y. Shen, J. K. Mills, and D. Sun, "Achieving automated organelle biopsy on small single cells using a cell surgery robotic system," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 8, pp. 2210–2222, 2018.

[2] S. C. Esteves, M. Roque, G. Bedoschi, T. Haahr, and P. Humaidan, "Intracytoplasmic sperm injection for male infertility and consequences for offspring," *Nature Reviews Urology*, vol. 15, no. 9, pp. 535–562, 2018.

[3] J. C. Harper, "Preimplantation genetic screening," *Journal of medical screening*, vol. 25, no. 1, pp. 1–5, 2018.

[4] P. Mrowiec and M. Bugno-Poniewierska, "Technical, biological and molecular aspects of somatic cell nuclear transfer–a review," *Annals of Animal Science*, vol. 22, no. 1, pp. 63–87, 2022.

[5] X. Liu, Z. Lu, and Y. Sun, "Orientation control of biological cells under inverted microscopy," *IEEE/ASME Transactions on Mechatronics*, vol. 16, no. 5, pp. 918–924, 2010.

[6] Z. Lu, X. Zhang, C. Leung, N. Esfandiari, R. F. Casper, and Y. Sun, "Robotic icsi (intracytoplasmic sperm injection)," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 7, pp. 2102–2108, 2011.

[7] K. Huang, I. A. Ajamieh, Z. Cui, J. Lai, J. K. Mills, and H. K. Chu, "Automated embryo manipulation and rotation via robotic ndeptweezers," *IEEE Transactions on Biomedical Engineering*, vol. 68, no. 7, pp. 2152–2163, 2020.

[8] S. V. Puttaswamy, N. Bhalla, C. Kelsey, G. Lubarsky, C. Lee, and J. McLaughlin, "Independent and grouped 3d cell rotation in a microfluidic device for bioimaging applications," *Biosensors and Bioelectronics*, vol. 170, p. 112661, 2020.

[9] G. Lin, Y. Liu, G. Huang, Y. Chen, D. Makarov, J. Lin, Z. Quan, and D. Jin, "3d rotation-trackable and differentiable micromachines with dimer-type structures for dynamic bioanalysis," *Advanced Intelligent Systems*, vol. 3, no. 2, p. 2000205, 2021.

[10] Y. Liu, G. Lin, and D. Jin, "Off-axis gyration induces large-area circular motion of anisotropic microparticles in a dynamic magnetic trap," *Applied Physics Letters*, vol. 119, no. 3, p. 034102, 2021.

[11] X. Liu, Y. Li, L. Li, M. Kojima, Q. Shi, Q. Huang, T. Fukuda, and T. Arai, "Noncontact 3-d orientation control at microscale: Hydrodynamic out-of-plane rotation and in-plane rotation by compacted rotational stage," *IEEE/ASME Transactions on Mechatronics*, 2022.

[12] O. Fuchiwaki, Y. Tanaka, H. Notsu, and T. Hyakutake, "Multi-axial non-contact in situ micromanipulation by steady streaming around two oscillating cylinders on holonomic miniature robots," *Microfluidics and Nanofluidics*, vol. 22, no. 8, pp. 1–13, 2018.

[13] J. Sun, N. Koukourakis, J. Guck, and J. W. Czarske, "Rapid computational cell-rotation around arbitrary axes in 3d with multi-core fiber," *Biomedical Optics Express*, vol. 12, no. 6, pp. 3423–3437, 2021.

[14] M. Xie, "Autonomous robot-aided optical tweezer system for biological cell manipulation," *The International Journal of Advanced Manufacturing Technology*, vol. 105, no. 12, pp. 4953–4966, 2019.

[15] J. Zhu, Q. Zhang, F. Liang, Y. Feng, and W. Wang, "High-throughput acoustofluidic microchannels for single cell rotation," *Journal of Micromechanics and Microengineering*, vol. 31, no. 12, p. 124004, 2021.

[16] Q. Tang, F. Liang, L. Huang, P. Zhao, and W. Wang, "On-chip simultaneous rotation of large-scale cells by acoustically oscillating bubble array," *Biomedical Microdevices*, vol. 22, no. 1, pp. 1–11, 2020.

[17] Q. Zhao, M. Sun, M. Cui, J. Yu, Y. Qin, and X. Zhao, "Robotic cell rotation based on the minimum rotation force," *IEEE Transactions on Automation Science and Engineering*, vol. 12, no. 4, pp. 1504–1515, 2014.

[18] C. Zhao, Y. Liu, M. Sun, and X. Zhao, "Robotic cell rotation based on optimal poking direction," *Micromachines*, vol. 9, no. 4, p. 141, 2018.

[19] C. Dai, Z. Zhang, Y. Lu, G. Shan, X. Wang, Q. Zhao, C. Ru, and Y. Sun, "Robotic manipulation of deformable cells for orientation control," *IEEE Transactions on Robotics*, vol. 36, no. 1, pp. 271–283, 2019.

[20] M. Khalilian, M. Navidbakhsh, M. R. Valojerdi, M. Chizari, and P. E. Yazdi, "Estimating young's modulus of zona pellucida by micropipette aspiration in combination with theoretical models of ovum," *Journal of the Royal Society Interface*, vol. 7, no. 45, pp. 687–694, 2010.

[21] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

[22] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski *et al.*, "Human-level control through deep reinforcement learning," *nature*, vol. 518, no. 7540, pp. 529–533, 2015.

[23] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot *et al.*, "Mastering the game of go with deep neural networks and tree search," *nature*, vol. 529, no. 7587, pp. 484–489, 2016.

[24] A. Banino, C. Barry, B. Uria, C. Blundell, T. Lillicrap, P. Mirowski, A. Pritzel, M. J. Chadwick, T. Degris, J. Modayil *et al.*, "Vector-based navigation using grid-like representations in artificial agents," *Nature*, vol. 557, no. 7705, pp. 429–433, 2018.

[25] K. Fang, Y. Zhu, A. Garg, A. Kurenkov, V. Mehta, L. Fei-Fei, and S. Savarese, "Learning task-oriented grasping for tool manipulation from simulated self-supervision," *The International Journal of Robotics Research*, vol. 39, no. 2-3, pp. 202–216, 2020.

[26] J. Lee, J. Hwangbo, L. Wellhausen, V. Koltun, and M. Hutter, "Learning quadrupedal locomotion over challenging terrain," *Science robotics*, vol. 5, no. 47, p. eabc5986, 2020.

[27] K. Latifi, A. Kopitca, and Q. Zhou, "Model-free control for dynamic-field acoustic manipulation using reinforcement learning," *IEEE Access*, vol. 8, pp. 20 597–20 606, 2020.

[28] P. Leinen, M. Esders, K. T. Schütt, C. Wagner, K.-R. Müller, and F. S. Tautz, "Autonomous robotic nanofabrication with reinforcement learning," *Science advances*, vol. 6, no. 36, p. eabb6987, 2020.

[29] S. Liang, R. Xi, X. Xiao, and Z. Yang, "Adaptive sliding mode disturbance observer and deep reinforcement learning based motion control for micropositioners," *Micromachines*, vol. 13, no. 3, p. 458, 2022.

[30] H. Lamb, *Statics: including hydrostatics and the elements of the theory of elasticity*. The University Press, 1924.

[31] C. T. McKee, J. A. Last, P. Russell, and C. J. Murphy, "Indentation versus tensile measurements of young's modulus for soft biological tissues," *Tissue Engineering Part B: Reviews*, vol. 17, no. 3, pp. 155–164, 2011.

[32] H. Gong, L. Li, J. Qiu, Y. Yao, Y. Liu, M. Cui, Q. Zhao, X. Zhao, and M. Sun, "Automatic cell rotation based on real-time detection and tracking," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 7909–7916, 2021.

[33] C. J. Watkins and P. Dayan, "Q-learning," *Machine learning*, vol. 8, no. 3, pp. 279–292, 1992.