Deep Reinforcement Learning-Based Collision-Free Navigation for Magnetic Helical Microrobots in Dynamic Environments

Huaping Wang, *Member, IEEE*, Yukang Qiu, Yaozhen Hou, Qing Shi, *Senior Member, IEEE*, Hen-Wei Huang, *Member, IEEE*, Qiang Huang, *Fellow, IEEE*, and Toshio Fukuda, *Life Fellow, IEEE*

***Abstract*— Magnetic helical microrobots have great potential in biomedical applications due to their ability to access confined and enclosed environments via remote manipulation by magnetic fields. However, achieving collision-free navigation for microrobots in complex and unstructured environments, particularly in highly dynamic settings, remains a challenge. In this paper, we present a novel deep reinforcement learning-based control framework for magnetic helical microrobots, focusing on the tasks of goal-reaching and dynamic obstacle avoidance. To streamline data collection, a specialized training environment capturing essential aspects of navigation for magnetic helical microrobots is devised. The robustness and adaptability of the trained policy are supported using a randomization technique within the training environment. To facilitate seamless integration with real-world magnetic actuation systems, a visual processing algorithm based on OpenCV is devised and incorporated to collect policy observations. Simulations and experiments in various scenarios validate the high robustness and adaptability of the method. The performance assessment revealed a success rate of 99% in navigating the microrobot around 4 dynamic obstacles of comparable speeds and a success rate of 90% in environments with 14 dynamic obstacles. The results indicate the potential for future applications of our method in unstructured, confined, and dynamic living environments.**

***Note to Practitioners*—The motivation of this work is to develop a robust and effective control scheme for collision-free navigation of magnetic helical microrobots in dynamic environments. The conventional navigation strategies in dynamic environments mainly include global path planning and local path replanning; thus, highly dynamic environments require frequent updates to the planned path, making it difficult to apply in highly dynamic environments. In this work, a deep reinforcement learning-based control framework is proposed that can guide microrobots through many dynamic obstacles to a series of locations without collisions. The simulation and experimental results validate the efficacy of the proposed control framework and the robustness and adaptability of the trained policy. The proposed control scheme enables better understanding of advanced motion control methods for magnetic microrobots.**

***Index Terms*— Magnetic helical microrobot, dynamic obstacle avoidance, electromagnetic actuation, deep reinforcement learning, sim-to-real transfer.**

I. Introduction

DUE to their ability to access confined and enclosed environments inside the human body via remote manipulation by magnetic fields, magnetic helical microrobots have drawn great attention for use in minimally invasive medicine [1], [2], [3], [4], [5], [6], [7], [8]. Since the unstructured liquid environment of the human body dynamically changes, microrobots inevitably encounter moving obstacles, such as cell clusters and shed tissue, when performing in vivo tasks [9], [10], [11], [12], [13]. This requires microrobots to perform autonomous navigation and dynamic obstacle avoidance to adapt to complex and unpredictable environments. Successful navigation in such environments requires comprehensive environmental mapping and real-time adaptability to changing obstacle configurations [14], [15], [16], [17], [18]. Microrobots need robust intelligence for instant obstacle recognition and prompt path adjustments, thereby ensuring efficient and safe traversal through complex spaces. Therefore, fast and accurate environment interpretation and decision-making are especially important.

The traditional navigation framework involves perceiving the environment, localizing the microrobot, mapping its surroundings, and finally employing path planning to determine the optimal route to the goal. In this framework, path planning

Received 18 September 2024; accepted 23 September 2024. This article was recommended for publication by Associate Editor C. Dai and Editor X. Liu upon evaluation of the reviewers' comments. This work was supported in part by the National Key Research and Development Program of China under Grant 2023YFB4705400; in part by Beijing Natural Science Foundation under Grant 4232055; in part by the National Natural Science Foundation of China under Grant 62073042, Grant 62222305, Grant 62403056, and Grant 62088101; in part by the Postdoctoral Fellowship Program of China Postdoctoral Science Foundation (CPSF) under Grant BX20230459; and in part by the Science and Technology Innovation Program of Beijing Institute of Technology under Grant 2022CX01019. *(Corresponding author: Yaozhen Hou.)*

Huaping Wang is with the Key Laboratory of Biomimetic Robots and Systems, Ministry of Education, Beijing Institute of Technology, Beijing 100081, China (e-mail: wanghuaping@bit.edu.cn).

Yukang Qiu and Yaozhen Hou are with the Intelligent Robotics Institute, School of Mechatronical Engineering, Beijing Institute of Technology, Beijing 100081, China (e-mail: callmeklausplease@gmail.com; houyaozhen@bit.edu.cn).

Qing Shi and Qiang Huang are with Beijing Advanced Innovation Center for Intelligent Robots and Systems, Beijing Institute of Technology, Beijing 100081, China (e-mail: shiqing@bit.edu.cn; qhuang@bit.edu.cn).

Hen-Wei Huang is with the School of Electrical and Electronic Engineering and the LKC School of Medicine, Nanyang Technological University, Singapore 639798 (e-mail: henwei.huang@ntu.edu.sg).

Toshio Fukuda is with the Department of Micro-Nano Systems Engineering, Nagoya University, Nagoya, Aichi 464-8603, Japan (e-mail: tofukuda@nifty.com).

This article has supplementary material provided by the authors and color versions of one or more figures available at https://doi.org/10.1109/TASE.2024.3470810.

Digital Object Identifier 10.1109/TASE.2024.3470810

plays a crucial role because it dictates the trajectory that the microrobot should follow in avoiding obstacles and achieving its objectives [19], [20], [21], [22]. Traditional global path planning methods, such as RRT [23], are proficient in exploring high-dimensional state spaces efficiently. These algorithms excel in generating feasible paths for microrobots, especially in complex maze-like environments [24], [25], [26]. However, their reliance on static maps renders these methods unsuitable for dynamic environments. Potential field methods, such as the artificial potential field (APF), involve real-time planning, which enables microrobots to quickly adjust their motion trajectories to adapt to dynamic changes in the environment [27], [28], [29]. However, such navigation methods are primarily based on local sensory information and do not incorporate global map information, which may limit their ability to find optimal paths in large-scale environments. To counterbalance these limitations and strike a balance between global optimality and local reactivity to dynamic changes, hybrid methods incorporating global path planning and local path replanning have been proposed to ensure that microrobots follow an efficient route and react quickly to dynamic changes in the environment [30]. However, these methods are sensitive to parameters and therefore lack adaptability. Additionally, intensive computations are needed, which may constrain the applicability of these methods in real-time scenarios, especially in highly dynamic environments. To reduce the computational burden, a hybrid method augmented with a fuzzy logic approach was proposed [31]. The fuzzy logic rule base utilizes expert experience to control microrobots, and real-time turn angles are determined through a fuzzy logic-based path planner to avoid incoming obstacles. Nevertheless, extending fuzzy logic systems to handle environments with a high number of dynamic obstacles poses challenges, as the rule base may become intricate and challenging to manage in such scenarios. In conclusion, the existing navigation strategies for microrobots are inadequate for operating effectively in highly dynamic environments. This inadequacy arises from the challenges associated with concurrently managing static and dynamic elements, as well as the computational burden imposed by complex algorithms. Therefore, an advanced and adaptive navigation strategy is needed for guiding microrobots through highly dynamic environments.

Recent advancements in robotics have employed deep reinforcement learning (DRL) as a promising approach for achieving more flexible, autonomous, and adaptive navigation in dynamic environments. DRL has been applied in various control scenarios, including adaptive motion planning [32], fleet management [33], and cooperative task execution [34]. These approaches allow agents to learn complex and nonlinear mappings, enabling them to adapt to intricate environments [35]. The workflow of DRL-based navigation typically involves iteratively training an agent to interact with its environment and learn a policy that maps observations to actions through trial and error. This process enables autonomous decision-making in dynamic and complex environments, thereby enhancing the agent's navigation capabilities. By applying DRL-based methods to the motion control of microrobots, more efficient and robust navigation in complex and dynamic environments can be achieved. However, due to scale effects and kinematic differences from the macroscale to the microscale, the motion behavior of microrobots is affected by various factors, such as fluid disturbances and material surface effects, making accurate modeling challenging. Moreover, integrating sensors at the microscale is impractical, resulting in insufficient feedback information for iteration, which is essential for DRL algorithms requiring extensive environmental, motion, and error analysis data for calculation and training. As a result, translating DRL-based methods to microrobot navigation encounters difficulties in modeling accuracy, data acquisition, and adaptation to real-world scenarios. Therefore, developing a DRL-based collision-free navigation strategy for microrobots in dynamic environments remains a significant challenge.

In this paper, we introduce a novel control framework based on deep reinforcement learning (DRL) designed for magnetic helical microrobots, focusing on the tasks of goal-reaching and dynamic obstacle avoidance. The policy, trained through simulated interactions between the microrobot and a dynamic environment, enables effective navigation through dynamic obstacles to reach predefined goals. To enhance the efficiency of data collection for training, we constructed a customized training environment that captures essential aspects of navigation for magnetic helical micro swimmers. To fortify the robustness and adaptability of the trained policy in real-world scenarios, a randomization technique is incorporated into the training environment. For seamless transfer of the trained policy to real-world magnetic actuation systems, a visual processing algorithm based on OpenCV is developed and integrated with the system to collect policy observations. Finally, simulations and experiments in environments characterized by a high degree of variability are conducted to validate the robustness and adaptability of the proposed method, demonstrating its great potential for biomedical applications in unstructured, confined, and dynamic living environments.

The structure of this paper is outlined as follows. Section II provides an in-depth overview of the microdrill design and the DRL-based control framework. Section III gives the details of the designed environment and expounds upon the training specifics. The experimental results are presented in Section IV, and finally, Section V provides the conclusions of this paper.

II. Design and Modeling

A. Design and Fabrication of Microdrills

We designed a microrobot with a cylindrical core that is enveloped by a double helix structure. The 3D microdrill design is shown in Figure 2(a). The microdrill has a length of 75 μm, pitch of 60 μm, wavenumber of 1.25, inner diameter of 4.2 μm, and cord radius of 12.3 μm.

The microdrill was made of compounded biocompatible photoresist with a photoinitiator, which was fabricated by a high-precision 3D photolithography system (NanoScribe Photonic Professional GT) [36]. In addition, the microrobots were uniformly coated with magnetic nanoparticles for magnetic actuation.

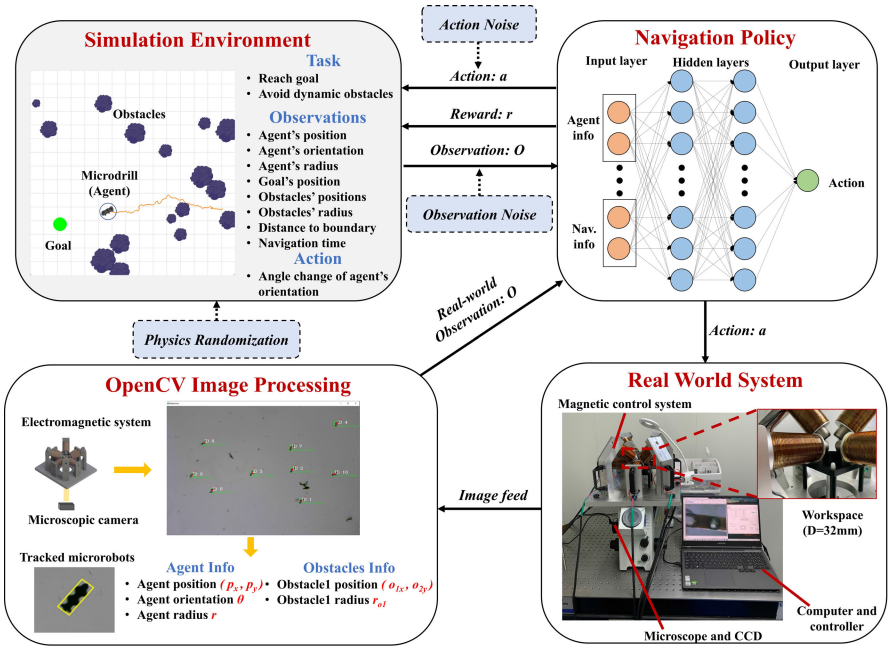

Fig. 1. Pipeline of the whole DRL-based control framework.

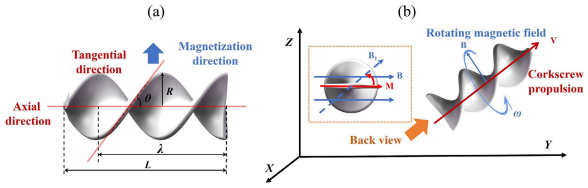

Fig. 2. (a) Structural design of the microdrill. (b) Motion diagram of the microdrill.

B. Actuation of Microdrills

To actuate the microdrills, a rotating magnetic field generated by an electromagnetic system comprising eight axial electromagnets is utilized. After the microdrill is magnetized perpendicular to its helical axis, the magnetization of the microdrill can be viewed as being directed perpendicular to its helical axis, as illustrated in Figure 2. When the microdrill is subjected to an external uniform magnetic field, the magnetic torque τ of the microdrill can be expressed as:

$$\tau = \mathrm{V}\mathbf{M} \times \mathbf{B} \tag{1}$$

where V represents the volume of the microrobot, M represents its magnetization, and $\mathbf{B}$ denotes the external magnetic field. The torque tends to align the magnetic moment with the applied field [37]. By continuously rotating the applied field $\mathbf{B}$ in a circle on a two-dimensional plane, the microdrill undergoes continuous rotation around its helical axis to achieve propulsion. The swimming direction and forward speed of the microdrill can be controlled by adjusting the rotation direction and frequency of the magnetic field.

C. DRL-Based Control Framework

Although the motion dynamics of magnetic helical microrobots can be theoretically modeled, when microrobots are actuated under external magnetic fields to perform dynamic tasks in real environments, factors such as fluid viscosity, impurities, and surface roughness significantly affect the microrobots' velocity, swimming direction, and motion stability. Not all the influencing factors can be applied or tested in every experiment because this would require considerable time and reduce the control efficiency. Therefore, it is necessary to construct a customized test environment model that includes key aspects of real-world environments as a training environment for achieving precise motion control of microrobots. Additionally, instead of modeling the entire dynamics of the environment in a simulator, certain key aspects of the real world are abstracted, and specific simulators addressing specific requirements and constraints are customized for simplicity. To simplify the simulation environment, the following key points should be considered:

1. The speed of the microdrill is controlled by the frequency of the rotating magnetic field, and we maintain a constant speed in the simulation since the microrobot typically operates at a fixed frequency lower than the step-out frequency.
2. The direction of the microdrill motion is controlled by the direction of the rotating magnetic field, which is the primary parameter we control during navigation.
3. The microdrill must maintain a safe distance from dynamic obstacles; we check for collisions in the simulation by comparing the distance between the microdrill and an obstacle to the sum of the radii of their circumscribed circles.

Following the construction of the custom training environment, a DRL-based control pipeline is developed. As depicted in Figure 1, the framework comprises four main modules: the simulation environment, the navigation policy, the real-world magnetic actuation system, and the OpenCV-based image processing algorithm. The simulation environment abstracts the navigation task for magnetic helical microrobots, and the

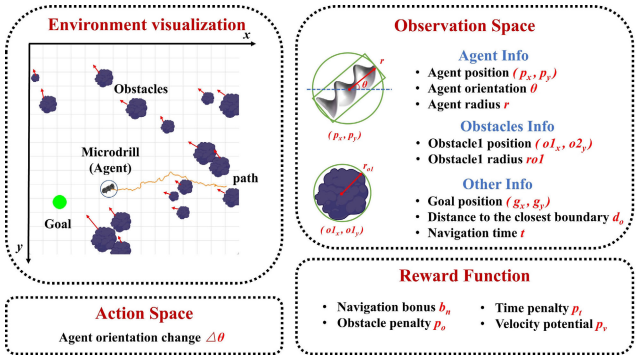

Fig. 3. Illustration of the key parts of the custom training environment, including the observation space, the action space and the reward function.

navigation policy is trained within this custom simulator. The real-world magnetic actuation system processes actions from the policy and generates magnetic fields to control the microrobot. The image processing algorithm serves as a feedback module, providing observations during training. A sim-to-real transfer method is employed to minimize the discrepancies between the simulation and real-world environments. To enhance robustness, the training phase incorporates both observation and action noise to mitigate tracking and actuation discrepancies. Furthermore, physical randomization is employed to ensure that the policy is adaptable across a spectrum of dynamics, thereby broadening its applicability to real-world scenarios.

III. Methods

In this section, we design a control framework aimed at training a policy to guide a microdrill toward a goal while avoiding dynamic obstacles. A customized training environment that follows the Open AI gym interface [38] is built, as illustrated in Fig. 3. Our approach involves training the policy in simulation and then applying it to real-world environments. The key components of the control framework are described in this section.

A. Customized Training Environment

1) Task: Our task is to navigate a microdrill actuated by an external electromagnetic system, as depicted in Figure 2, to a specific goal location while avoiding dynamic obstacles. The microdrills are actuated by a rotating magnetic field, and the direction and speed of motion are controlled by the direction and frequency of the external rotating magnetic fields, respectively. In the simulator, the goal is randomly generated within the environment's boundaries in each episode and the speed of the agent is set as v. And a fixed number of obstacles with random sizes and speeds are randomly generated along the bottom or the rightmost edge of the environment, while avoiding placement near the bottom-left and upper-right corners, as obstacles in these areas would quickly move outside the boundary. The position of a generated obstacle i can be expressed as follows:

$$\mathbf{p_i} = \begin{cases} x \sim U[\frac{1}{15}W, W], \;\; y = 0 & \text{if } r = -1 \\ x = W, \;\; y \sim U[0, \frac{14}{15}H] & \text{if } r = 1 \end{cases} \tag{2}$$

where $U[a, b]$ denotes a uniform distribution between a and b, r is a random variable drawn from a discrete uniform distribution $\{-1, 1\}$, and W and H are the width and height of the environment, respectively.

The direction of each obstacle's movement varies in each timestep, but generally, the obstacles move toward the upper left corner of the environment. The speed of an obstacle i in each timestep is:

$$\mathbf{v_i} = \begin{pmatrix} v_{ix} \\ v_{iy} \end{pmatrix} = \begin{pmatrix} U[-v, v/3] \\ U[v/3, v] \end{pmatrix}, \quad i = 1, 2, \ldots, n \tag{3}$$

The number of obstacles in the environment is fixed during training, and the environment replaces any obstacles that move outside the boundaries with new obstacles generated according to (2). The microdrill fails if it collides with an obstacle or the environment's boundaries and succeeds upon reaching the goal.

2) Observation: The observation space is a continuous box with lower and upper bounds defined by the width, height, and other parameters of the environment:

$$O = [p_x, p_y, \theta, r, g_x, g_y, o_{1x}, o_{1y}, r_1, \ldots, d_o, t] \tag{4}$$

where p_x is the x-coordinate of the microdrill's current position, p_y is the y-coordinate of the microdrill's current position, θ is the orientation (in radians) of the microdrill's heading, r is the radius of the circumscribed circle of the microdrill, g_x and g_y represent the position of the goal, o_{nx}, o_{ny} and r_n denote the position and radius of the n-th obstacle, d_o is the distance to the closest boundary of the environment, and t denotes the number of timesteps in the episode.

3) Action: In this environment, we focus on the key parameters that govern the motion of a helical micro swimmer, namely, the direction of the rotating magnetic field. The action space A is a 1-dimensional box with lower and upper bounds of -1 and 1, respectively. The action is used to adjust the orientation of the microdrill, which is updated by adding the product of the action and $\pi/6$ to the current orientation. In this way, we limit the change in the orientation of the microdrill to one timestep within the range of $[-\pi/6, \pi/6]$.

$$A = [\theta_d] \tag{5}$$

4) Reward: The reward function r calculates a reward signal based on four terms: navigation bonus b_n, obstacle penalty p_o, time penalty p_t, and velocity potential p_v.

$$r = b_n + p_o + p_t + p_v \tag{6}$$

The navigation bonus b_n, determined by the attractive potential, encourages the microdrill to move toward the goal by providing a reward proportional to the inverse of the distance to the goal. The closer the microdrill is to the goal, the greater the reward.

$$b_n = c_a \cdot \frac{1}{d_{2g}} \tag{7}$$

where c_a is the attraction coefficient and d_{2g} is the distance between the microdrill and the goal position.

The obstacle penalty p_o encourages the microdrill to avoid obstacles by providing a penalty proportional to the inverse of

the distance to the obstacles. The closer the microdrill is to the obstacles, the higher the penalty:

$$p_o = -c_r \cdot \sum_i^n \left(\frac{1}{d_{2i}} - \frac{1}{d_{safe}} \right) \tag{8}$$

where n is the number of obstacles and d_{2i} is the distance between the microdrill and the ith obstacle position; d_o is included in d_{2i} in this equation. The repulsive coefficient c_r determines the strength of the repulsive force, while the safe distance d_{safe} determines the distance at which the repulsive force starts to take effect.

The time penalty p_t discourages the microdrill from taking too much time to reach the goal. It provides a negative reward that is proportional to the time taken by the microdrill.

$$p_t = -k_t \tag{9}$$

where k_t is a constant penalty.

The velocity potential p_v encourages the microdrill to move away from obstacles and penalizes it for moving toward obstacles. $\vec{v}_{rel}$ is the relative velocity between the microdrill and the i-th obstacle, $\vec{a}_i$ is the unit vector pointing from the i-th obstacle to the microdrill, and n is the number of obstacles. The dot product parameter determines the direction of the microdrill's velocity relative to the obstacle, while the velocity potential parameter k_v determines the strength of the penalty. A positive dot product indicates that the microdrill and obstacle are moving in roughly the same direction, and the velocity potential is set to 0, which means that no reward or penalty is given. A negative dot product indicates that the microdrill is moving toward the obstacle, and the velocity potential is set to a negative value, which indicates that the microdrill is penalized.

$$p_v = \sum_i^n \begin{cases} 0, & \text{if } \vec{v}_{rel} \cdot \vec{a}_i > 0 \\ kv \cdot \vec{v}_{rel} \cdot \vec{a}_i, & \text{otherwise} \end{cases} \tag{10}$$

B. Randomization

Despite the meticulous design of the custom environment for the task, the simulation remains an approximation of the physical setup. Even slight deviations from the simulation can result in undesirable outcomes in real-world experiments. For instance, a microdrill can drift due to interactions with the solid boundary beneath it [39], causing a discrepancy between its actual motion and the intended input action. Moreover, this discrepancy varies across different microrobots and rotating frequencies, making accurate modeling challenging. These disparities between the simulation and real-world setup create a gap with reality, meaning that policies trained solely in the simulator will struggle with effective transfer to real-world scenarios. Therefore, the transfer of DRL policies from simulation environments to reality has become a crucial step toward the application of complex robotic systems, and the most widely used method for learning transfer is domain randomization. Domain randomization involves extensively randomizing the simulation parameters rather than meticulously modeling every aspect of the real world; this ensures that the real distribution of the real-world data is covered despite the mismatch between the model and real world.

TABLE I
STANDARD DEVIATION OF OBSERVATION NOISE

Observation	Value
Microdrill position	10 μm
Microdrill orientation	0.1 rad
Microdrill radius	5 μm
Obstacles positions	10 μm
Obstacles radius	5 μm

In real-world scenarios, data acquisition systems such as cameras often introduce noise in measurements due to various factors such as sensor inaccuracies, environmental disturbances, or electronic noise. By adding noise to policy observations, we can create a more realistic training environment that better mimics the noise present in real-world situations. By exposing the RL microdrill to various sources of noise during training, the microdrill is forced to learn policies that are more robust and adaptive. To achieve this, we incorporate Gaussian noise with a mean μ of 0 and a user-defined standard deviation σ, as outlined in TABLE I, the noise level (σ) is set based on the expected level of image capturing and processing inaccuracy acquired from experimentation.

In addition to introducing observation noise, we incorporate the randomization of physical parameters in our training environment to enhance the adaptability of the trained microdrill to real-world conditions. Before each episode begins, a set of physical parameters is uniformly sampled within a defined range. The ranges of the parameters are determined by measurements from our actual experimental setup, and the values of the parameters are kept constant throughout the entire episode. The randomization helps the agents generalize across different environmental conditions and increases the adaptability by ensuring they don't overfit to a specific set of physics parameters. The specific ranges for physical parameter randomizations are detailed in TABLE II, which correspond to the expected variation in the real-world situations where the policy should work. Specifically, the "Number of obstacles" indicates the trained policy is expected to guide the microrobot to navigate through this range of obstacles. The "Microdrill speed" ranges from 20 μm/s to 40 μm/s, corresponding to a rotating frequency of 4 Hz to 12 Hz for our microrobots [36]. And the "Goal radius" reflects the variability in task criteria for determining goal achievement. Unlike the other physical parameters listed, the speed and radius of each obstacle are not initialized before each episode. Instead, the speed of each obstacle is dynamically adjusted according to formula (3). Since some obstacles may move beyond the observable boundary of the simulated environment, new obstacles are generated to ensure a constant number of obstacles throughout an episode, and the radius of each obstacle is uniformly sampled within a specified range upon its generation.

A physical microdrill exhibits rolling (drifting) induced by an imbalance in drag forces resulting from its interaction with the solid boundary beneath it [36]. Rolling introduces

TABLE II
RANGES OF PHYSICAL RANDOMIZATION

Parameter	Range
Number of obstacles	U [8, 18]
Microdrill speed	U [20 μm/s, 40 μm/s]
Goal radius	U [60 μm, 90 μm]

a deviation between the actual motion direction of the microdrill and the intended input direction. To accommodate this deviation, the final direction of the rotating magnetic field is adjusted by adding 16°, a value predetermined by the microdrill's design, before it is input to the microdrill. Furthermore, to address the signal transmission delays in the magnetic actuation system and the external environmental disturbance of the microrobot, we introduce Gaussian noise with a mean of 0 and a standard deviation of 10% of the action range into the action space.

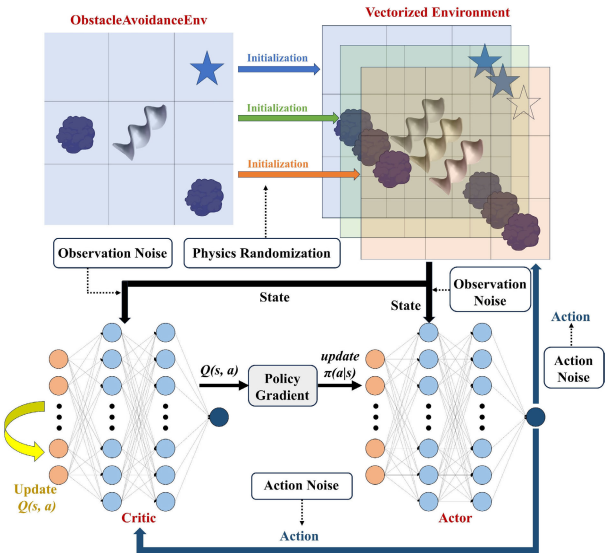

Fig. 4. Illustration of the training process with randomization.

C. Training

Given that both the action space and observation space in our tailored environment are continuous, we opt for proximal policy optimization (PPO) as our training algorithm. The PPO is selected for its sample efficiency and employs a clipped surrogate objective, mitigating the risk of overly drastic policy updates and thereby preventing divergence issues. Its stochastic policy ensures an appropriate exploration-exploitation trade-off, aiding the microdrill in discerning the optimal policy given the complexities of our dynamic environment.

Fig. 4 provides a high-level overview of our training process. Training begins by initializing an instance of the "ObstacleAvoidanceEnv" environment, which is specifically designed for microdrill navigation toward a goal while avoiding obstacles. Subsequently, a vectorized environment is generated with multiple parallel instances of the obstacle avoidance environment to enhance sample efficiency during training. Our chosen reinforcement learning algorithm is PPO, and the policy function employs a multilayer perceptron (MLP). This neural network, acting as the actor, takes the environment state as input and outputs the probability distribution over actions. Parallel to this, a critic network is used to estimate the value function, which predicts the expected return (or future rewards) from a given state. The critic's value estimation helps in computing the advantage function, which evaluates the action produced by the actor. Additionally, an evaluation callback method is configured to periodically assess the model's performance in the environment and log relevant metrics. Throughout training, the actor network guides the microdrill's actions at each timestep, while the critic network evaluates these actions to refine the policy. The PPO utilizes a combination of policy and value iteration for iterative policy improvement. The use of a clipped surrogate objective ensures stable policy updates by constraining the objective function to prevent large updates that might compromise training stability. The actor network is updated to improve the policy, and the critic network is updated to refine its value predictions. After training, the PPO model is saved for subsequent evaluation, testing, and deployment.

Throughout training, we utilize two key metrics, namely, the "episode length mean" and "episode reward mean." These metrics represent the average number of time steps (or actions) taken by the microdrill and the average cumulative reward obtained through an episode, respectively. Paramount to the training outcomes is the parameter configuration within the reward function, which defines the feedback the microdrill receives based on its motions in the environment. Notably, careful tuning is essential for preventing undesirable behaviors. For example, an excessively large penalty may incentivize the microdrill to crash into boundaries for fast but undesired termination, while an overly generous navigation bonus may lead the microdrill to circle the goal endlessly without completing the task. After iteratively tuning the reward function over hundreds of iterations, a balanced formulation consistent with our objectives was achieved. The policy involves training for 1 million total time steps on an NVIDIA GeForce RTX 2060 with 22 GB of memory, which was completed in 38 minutes and 41 seconds. The training metrics, as depicted in Fig. 5, reveal that the "episode length mean" continually decreases after the initial exploration phase, and the "episode reward mean" steadily increases until it reaches an extreme value at approximately 200k time steps. This trend signifies that the microdrill progressively refines its strategy for more efficient goal achievement, ultimately attaining optimal performance at approximately 200k time steps.

D. Simulation Results

To evaluate the obstacle avoidance efficacy of the trained policy, simulation tests are conducted. To evaluate the robustness of the trained policy, the microdrill and goal are randomly positioned within an environment featuring a variable number (ranging from 0 to 30) of static or dynamic obstacles. The failure criteria include instances where the microdrill moves outside the boundaries, collides with obstacles, or exceeds the time limit. The success rate, which

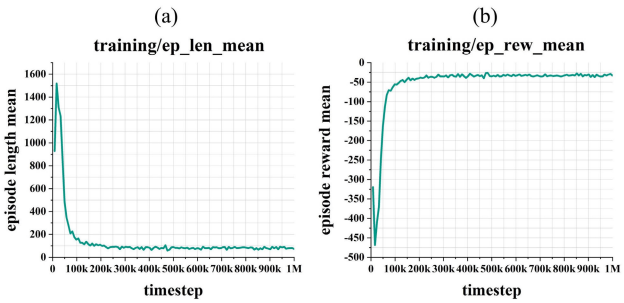

Fig. 5. Evaluation metrics during training. (a) The average length of episodes during training. (b) The average reward of episodes during training.

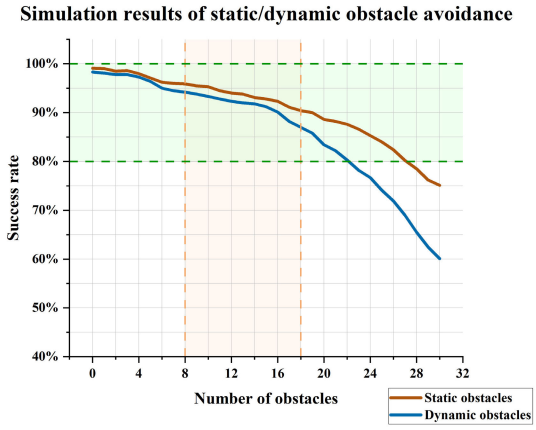

Fig. 6. Success rate of static and dynamic obstacle avoidance with various numbers of obstacles in the simulation.

quantifies the ability of a microdrill to reach its goal without colliding with obstacles, is computed over a thousand simulation test episodes. As shown in Fig. 6, the policy, initially designed for dynamic obstacle avoidance, is effectively extended to cases of static obstacles as well. The orange segment indicates cases with a number of obstacles ranging from 8 to 18, consistent with the training range. Notably, the average success rates for static and dynamic obstacle avoidance within this range are 94.0% and 92.2%, respectively. Considering success rates exceeding 80% as acceptable, the green section denotes the acceptable range and is notably broader than the originally trained range.

IV. Experiments

A. System and Experimental Setup

In real-world experiments, our microdrills were placed in Petri dishes filled with DI water. Due to the challenge of enclosing a microdrill with obstacles of predefined sizes and speeds, we generated artificial obstacles through image processing to assess the microdrill performance. The radius of the obstacles ranged from 37.5 μm to 150 μm, and the size of the environment was 840 μm $\times$ 1125 μm. The workflow depicting the integration of the magnetic actuation system with the deep RL policy is presented in Fig. 7. The detailed hardware specifications can be found in our prior publication [36], [40]. Throughout the experiments, an inverted microscope was used to capture real-time images, while an OpenCV-based image processing algorithm was developed to extract essential observation data, including the position, orientation, and radius

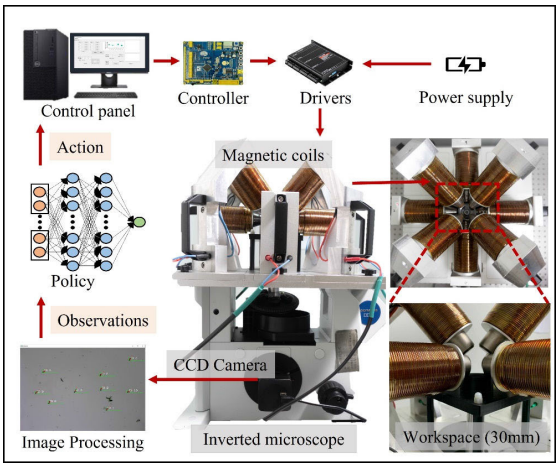

Fig. 7. Schematic of the electromagnetic control system combined with the deep RL policy.

of the microdrill, as well as the positions and radii of artificial obstacles within the field of view. Each image frame is cropped to eliminate extraneous areas, thereby reducing noise. Gaussian blurring is then applied for smoothing and to suppress minor noise, followed by morphological opening operations for additional noise reduction. The image is converted from the BGR to the HSV color space, which enhances detection of the black microrobot. Contour matching is utilized to locate the microrobot, and shape fitting is employed to determine its position and orientation. A black mask is applied between frames to improve detection accuracy and reduce computational complexity. The extracted observations are subsequently input into the trained policy, which controls the microdrill's motion by adjusting the direction of the rotating magnetic field. The policy's output is sent to the control panel, where the necessary current for each electromagnetic coil is calculated to generate the desired magnetic field. A subordinate microcontroller receives these data and converts them into pulse-width modulation (PWM) signals with adjustable duty cycles. These signals are transmitted to current drivers, which then deliver the programmed current to each electromagnet, generating the desired magnetic field to control the microdrill's motion.

B. Real-Time Obstacle Avoidance in Three Scenarios

To assess the robustness and adaptability of our trained policy in obstacle avoidance, three distinct scenarios were designed for experimentation. These scenarios encompass environments featuring static obstacles, dynamic obstacles with entirely random motions, and dynamic obstacles following a specified flow direction, as illustrated in Fig. 8 (a), (c), and (e). In each scenario, the microdrill, obstacles (depicted as purple blobs), and goal (depicted as a green star) are initially placed at random positions. The green circles around them represent the circumcircles. Collisions with obstacles and goal attainment are identified when the distance between two objects is smaller than the sum of their circumradii. The red arrow signifies the motion direction of each obstacle. The microdrill is tasked with pursuing the goal while avoiding obstacles. Upon reaching a goal, a new goal is randomly generated for the microdrill. Notably, the model

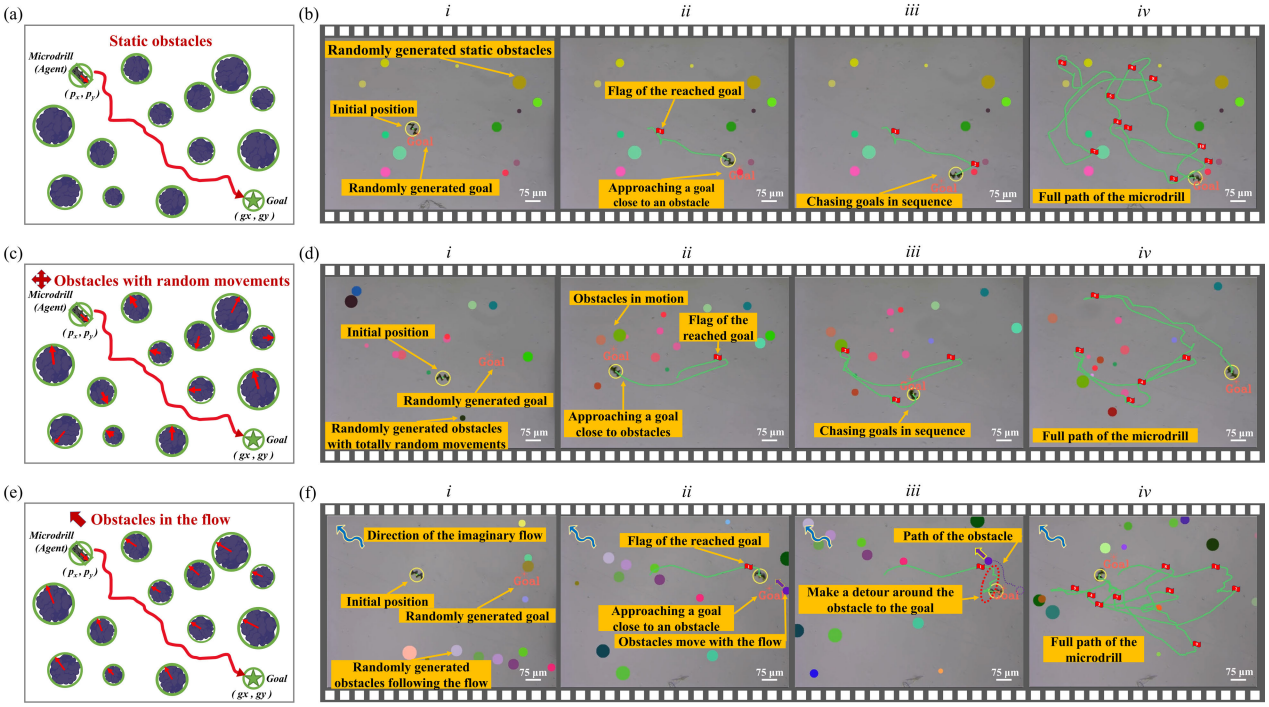

Fig. 8. Schematic illustrations and experimental images depicting obstacle avoidance and goal reaching of the microdrill in various scenarios.

was initially trained exclusively in environments similar to those in Fig. 8 (e) with randomizations. Experiments across these three scenarios validate the obstacle avoidance ability and adaptability of our model.

In real-world experiments, image processing is employed to generate artificial obstacles at random positions. Across all scenarios, the number of obstacles within the field of view is consistently maintained at 12, with radii uniformly distributed between 37.5 μm and 150 μm—corresponding to half to twice the length of the microrobot. The rotational frequency is set at 10 Hz, corresponding to a forward velocity v of 38.75 μm/s for the microrobot. In the scenario involving dynamic obstacles with entirely random motion, the x and y coordinates of each obstacle randomly change within the range of $[-|v|, +|v|]$ every second, resulting in a vibrating motion pattern. In the scenario with dynamic obstacles following a specified flow direction, the x coordinates change randomly within the range of $[-|v|, +|v|/3]$, while the y coordinates vary within the range of $[+|v|/3, +|v|]$ each second, creating the appearance of obstacles moving in a flow from the bottom right to the upper left. Fig. 8. (b) shows actual experimental screenshots of the microdrill navigating through static obstacles. In the initial state, static obstacles of varying sizes and the first goal are randomly generated, prompting the microdrill (in a yellow circle) to initiate navigation. Upon reaching a goal, a red flag with an index indicating the order of the tracked goals is marked on the green path, and a new goal is generated. Images *ii* and *iii* in Fig. 8 (b) show that as the microdrill approaches a goal in close proximity to obstacles, it prioritizes reaching the goal over obstacle avoidance by moving directly toward the goal. Image *iv* illustrates the full path of the microdrill following a sequence of goals through static obstacles, highlighting its adaptability to static obstacles despite having been trained in environments with dynamic directional obstacles, as shown in Fig. 8 (e). Fig. 8. (d) shows an image sequence of the microdrill navigating through randomly moving obstacles, displaying navigation performance akin to that of the scenario with static obstacles. When the goals are near obstacles, the microdrill exhibits a similar tendency to prioritize goal-reaching. In Fig. 8 (f), obstacles are generated from the right and bottom edges, following an imaginary flow toward the upper left, and their speeds are expressed by (3). Unlike in the previous scenarios, images *ii* and *iii* in Fig. 8 (f) show that when the microdrill approaches a goal close to an obstacle, it strategically avoids a direct collision. By anticipating the obstacle's path, the obstacle is circumvented, demonstrating the ability to navigate dynamically. This difference arises from the velocity potential component in our reward function (10), which incentivizes the microdrill to move in a manner that prevents obstacles from approaching it while penalizing movement toward obstacles moving away. The results align with our expectations. In scenarios with predictable obstacle motion (Fig. 8 (e)), the microdrill prioritizes obstacle avoidance. In scenarios with static or randomly moving obstacles, collisions are inevitable in reaching the goal, prompting the microdrill to prioritize goal attainment. These experimental results affirm that the trained model exhibits effective obstacle avoidance with directional dynamic obstacles and adapts well to static and random dynamic obstacles.

C. Performance of Dynamic Obstacle Avoidance

To comprehensively assess the performance of our trained policy, we conducted experiments to observe and record the dynamic obstacle avoidance capabilities of the microdrill. Initially, we roughly positioned the microdrill and its goal

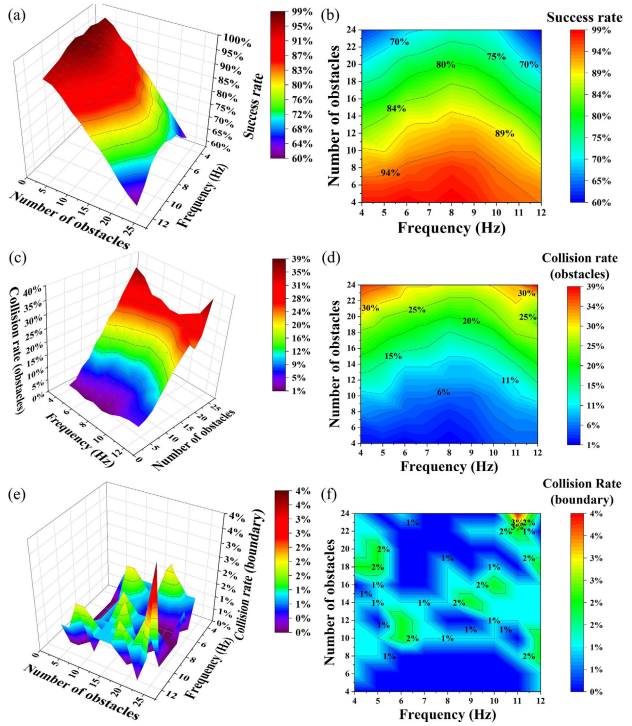

Fig. 9. Performance metrics of dynamic obstacle avoidance with varying rotating frequency and obstacle density. (a) 3D surface plot view of the success rate. (b) 2D contour map view of the success rate. (c) 3D surface plot view of the collision rate with obstacles. (d) 2D contour map view of the collision rate with obstacles. (e) 3D surface plot view of the collision rate with boundary. (f) 2D contour map view of the collision rate with boundary.

symmetrically within the environment, placing the microdrill at the upper left corner and the goal at the bottom right, with a distance of approximately 1140 μm between them. The number of obstacles and the magnetic field frequency were set as variables. The number of obstacles (n), reflecting the environmental complexity, ranged from 4 to 24 in increments of 2. Additionally, the rotating frequency (f), which represents the speed of the microdrill, ranged from 4 to 12 in increments of 1. For each group of n and f, we initialized 100 test cases in which each had an environment containing n randomly generated obstacles moving within the flow, and the microdrill was tasked with reaching the goal without collision. The successful instances were recorded to yield the success rate, which is graphically presented in Fig. 8. The trend depicted in Fig. 9 reveals a decline in the success rate as the number of obstacles increases; this aligns with expectations, as a greater number of obstacles increases the degree of environmental complexity. Additionally, a discernible slope is observed in Fig. 9 (b), where the success rate of navigation first increases as the rotating frequency increases and then decreases to some degree. Within the microdrill's step-out frequency, the forward velocity and the rotating magnetic field frequency exhibit a linear relationship. A lower rotating frequency results in slower motion for the microdrill, making collisions almost inevitable when pursued by faster obstacles. Conversely, a higher frequency endows the microdrill with greater speed and agility but reduces the available action steps when encountering obstacles, yielding a slope in the success rate in Fig. 9 (b). Furthermore, Fig. 9 (c) and (d) show an inverse relationship between the obstacle collision rate and the success rate during navigation, indicating the increased difficulty of dynamic obstacle avoidance as the number of obstacles in the environment increases. Fig. 9 (e) and (f) depict the rate of collision with boundaries, which represents the likelihood of the microrobot colliding with the boundary or straying outside the field of view during navigation. The results indicate that the collision rate with boundaries remains below 2% across most settings and exhibits no significant correlation with the frequency of the rotating magnetic field or the number of obstacles. This trend may be attributed to the penalty imposed on boundary collisions in the reward function. In this function, the boundary is treated as a stationary obstacle with four sides for the microrobot. Consequently, during the training phase, the microrobot encounters the boundary frequently, resulting in penalization and prompting the rapid learning of a collision avoidance strategy. The data plot underscores the effectiveness of the navigation strategy in avoiding boundary collisions and proficiently controlling the microrobot's movement within the designated range. To optimize the obstacle avoidance performance of the microdrill, it is crucial to determine an appropriate rotating frequency, contingent on the structure of the microdrill and the average speed of the surrounding obstacles. Through the experiments, the average speed of the dynamic obstacles was determined to be approximately 28.3 μm/s. The microdrill achieved its best navigation performance at a rotating frequency of 8 Hz, corresponding to a forward velocity of 31 μm/s. At this frequency, the microdrill achieved a success rate of 99% with 4 obstacles in the environment, 90% with 14 obstacles and 80% with 20 obstacles.

To comprehensively evaluate our method's performance, we compared its navigation success rate and average navigation time with those of two existing approaches for dynamic microrobot environments. The first approach utilizes fuzzy logic, which involves constructing a rule base from human experience to navigate through dynamic obstacles [31]. The second approach focuses on path planning, integrating the improved rapidly exploring random trees (IRRT) algorithm for global path planning and the improved artificial potential field (IAPF) algorithm for local path planning [30], [40], [41]. Two metrics, the success rate and average navigation time, were used for quantitative evaluation, representing collision-free navigation success and the average time taken for successful navigation, respectively. The experimental environment mirrored previous settings, with the microdrill positioned at the upper left corner and the goal at the bottom right. We examined 100 test cases for varying numbers of obstacles. To ensure fair comparison of navigation efficiency through the average navigation time, we maintained a constant microdrill speed by fixing the rotating frequency of the magnetic field at 8 Hz, at which the microdrill achieved its best performance as shown in Fig. 10. Fig. 10 (a) indicates that our DRL-based method achieved a higher success rate across most numbers of obstacles. While all methods achieved high success rates in less dynamic environments, when the number of obstacles exceeded 12, the success rates of the baseline methods significantly decreased, underscoring our method's ability to

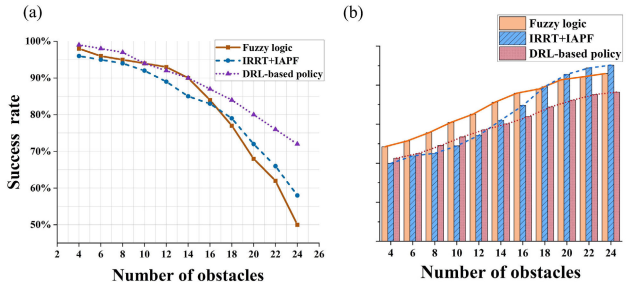

Fig. 10. Performance comparison of different methods. (a) Success rate. (b) Average navigation time of successful cases.

handle more dynamic and complex environments. Fig. 10 (b) illustrates the average navigation times of the three methods. With a fixed microdrill speed, the navigation time directly reflects the length of the actual navigation path. Compared with our method, the fuzzy logic approach exhibited longer navigation times than our method across all numbers of obstacles. Conversely, the "IRRT + IAPF" method showed slightly lower navigation times than our method with fewer obstacles due to efficient global path planning by the IRRT algorithm. However, as the environment grew more dynamic, particularly with more than 12 obstacles, the navigation time of the "IRRT + IAPF" method significantly increased, whereas our method found a more efficient path. Overall, our proposed method outperformed both existing methods in terms of success rate and navigation efficiency, particularly in highly dynamic and complex environments.

V. Conclusion

In this paper, we presented a DRL-based control framework for goal-reaching and dynamic obstacle avoidance via a microdrill. The overall control strategy was implemented with components including a custom training environment, the DRL algorithm, a magnetic actuation system and a real-time visual tracking method. First, we designed and fabricated a helical drill-like microrobot actuated by a rotating magnetic field. For effective data gathering, we constructed a custom DRL training environment adhering to the OpenAI Gym interface, abstracting the core physics of the navigation task. Based on the characteristics of our environment, we employed the PPO method to train the policy and integrated a randomization technique to enhance adaptability in real-world scenarios. To implement the policy in real-world systems, we developed a visual processing algorithm based on OpenCV, which, when combined with the magnetic actuation system, provides observations for the policy. The simulation and experimental results demonstrate that our method achieves the designed task of goal-reaching and dynamic obstacle avoidance with significant adaptability, and the method shows great potential in biomedical in vivo applications.

In future work, we aim to extend our methodology to three-dimensional space, enabling navigation in dynamic 3D environments. Furthermore, while our present approach is tailored for controlling a single microdrill during navigation, our future efforts will focus on scenarios involving multiple microrobots collaborating on tasks while actively avoiding collisions with each other.

References

[1] J. J. Abbott et al., "How should microrobots swim?" *Int. J. Robot. Res.*, vol. 28, nos. 11–12, pp. 1434–1447, Nov. 2009.
[2] X. Wang et al., "Microrobotic swarms for intracellular measurement with enhanced signal-to-noise ratio," *ACS Nano*, vol. 16, no. 7, pp. 10824–10839, Jul. 2022.
[3] X. Yan et al., "Multifunctional biohybrid magnetite microrobots for imaging-guided therapy," *Sci. Robot.*, vol. 2, no. 12, Nov. 2017, Art. no. eaaq1155.
[4] X. Dong, S. Kheiri, Y. Lu, Z. Xu, M. Zhen, and X. Liu, "Toward a living soft microrobot through optogenetic locomotion control of caenorhabditis elegans," *Sci. Robot.*, vol. 6, no. 55, Jun. 2021, Art. no. eabe3950.
[5] Y. Dong et al., "Magnetic helical micro-/nanomachines: Recent progress and perspective," *Matter*, vol. 5, no. 1, pp. 77–109, Jan. 2022.
[6] J. Miao et al., "Flagellar/ciliary intrinsic driven mechanism inspired all-in-one tubular robotic actuator," *Engineering*, vol. 23, pp. 170–180, Apr. 2023.
[7] S. Zhong et al., "Double-modal locomotion of a hydrogel ultra-soft magnetic miniature robot with switchable forms," *Cyborg Bionic Syst.*, vol. 5, p. 77, Jan. 2024.
[8] F. Wang et al., "Magnetic soft microrobot design for cell grasping and transportation," *Cyborg Bionic Syst.*, vol. 5, p. 109, Jan. 2024.
[9] L. Yang and L. Zhang, "Motion control in magnetic microrobotics: From individual and multiple robots to swarms," *Annu. Rev. Control, Robot., Auto. Syst.*, vol. 4, pp. 509–534, May 2021.
[10] J. Law, J. Yu, W. Tang, Z. Gong, X. Wang, and Y. Sun, "Micro/nanorobotic swarms: From fundamentals to functionalities," *ACS Nano*, vol. 17, no. 14, pp. 12971–12999, Jul. 2023.
[11] S. Tottori, L. Zhang, F. Qiu, K. Krawczyk, A. Franco-Obregón, and B. Nelson, "Micromachines: Magnetic helical micromachines: Fabrication, controlled swimming, and cargo transport," *Adv. Mater.*, vol. 6, no. 24, p. 709, 2012.
[12] X. Liu, K. Kim, Y. Zhang, and Y. Sun, "Nanonewton force sensing and control in microrobotic cell manipulation," *Int. J. Robot. Res.*, vol. 28, no. 8, pp. 1065–1076, Aug. 2009.
[13] B. Wang et al., "Low-friction soft robots for targeted bacterial infection treatment in gastrointestinal tract," *Cyborg Bionic Syst.*, vol. 5, p. 138, Jan. 2024.
[14] X. Wu, J. Liu, C. Huang, M. Su, and T. Xu, "3-D path following of helical microswimmers with an adaptive orientation compensation model," *IEEE Trans. Autom. Sci. Eng.*, vol. 17, no. 2, pp. 823–832, Apr. 2020.
[15] Y. Liu, H. Chen, Q. Zou, X. Du, Y. Wang, and J. Yu, "Automatic navigation of microswarms for dynamic obstacle avoidance," *IEEE Trans. Robot.*, vol. 39, no. 4, pp. 2770–2785, Aug. 2023.
[16] Q. Wang et al., "Ultrasound Doppler-guided real-time navigation of a magnetic microswarm for active endovascular delivery," *Sci. Adv.*, vol. 7, no. 9, Feb. 2021, Art. no. eabe5914.
[17] H. Xie et al., "Reconfigurable magnetic microrobot swarm: Multimode transformation, locomotion, and manipulation," *Sci. Robot.*, vol. 4, no. 28, Mar. 2019, Art. no. eaav8006.
[18] Y. Hou et al., "A review on microrobots driven by optical and magnetic fields," *Lab Chip*, vol. 23, no. 5, pp. 848–868, 2023.
[19] T. Xu, G. Hwang, N. Andreff, and S. Régnier, "Planar path following of 3-D steering scaled-up helical microswimmers," *IEEE Trans. Robot.*, vol. 31, no. 1, pp. 117–127, Feb. 2015.
[20] T. Xu, Y. Guan, J. Liu, and X. Wu, "Image-based visual servoing of helical microswimmers for planar path following," *IEEE Trans. Autom. Sci. Eng.*, vol. 17, no. 1, pp. 325–333, Jan. 2020.
[21] J. Liu et al., "3-D autonomous manipulation system of helical microswimmers with online compensation update," *IEEE Trans. Autom. Sci. Eng.*, vol. 18, no. 3, pp. 1380–1391, Jul. 2021.
[22] T. Xu, C. Huang, Z. Lai, and X. Wu, "Independent control strategy of multiple magnetic flexible millirobots for position control and path following," *IEEE Trans. Robot.*, vol. 38, no. 5, pp. 2875–2887, Oct. 2022.
[23] S. M. La Valle, "Rapidly-exploring random trees a new tool for path planning," Dept. Comput. Sci., Iowa State Univ., Ames, IA, USA, Tech. Rep., 1998.
[24] Q. Zou, X. Du, Y. Liu, H. Chen, Y. Wang, and J. Yu, "Dynamic path planning and motion control of microrobotic swarms for mobile target tracking," *IEEE Trans. Autom. Sci. Eng.*, vol. 20, no. 4, pp. 2454–2468, Oct. 2023.
[25] J. Liu, T. Xu, S. X. Yang, and X. Wu, "Navigation and visual feedback control for magnetically driven helical miniature swimmers," *IEEE Trans. Ind. Informat.*, vol. 16, no. 1, pp. 477–487, Jan. 2020.

[26] A. Short, Z. Pan, N. Larkin, and S. van Duin, "Recent progress on sampling based dynamic motion planning algorithms," in *Proc. IEEE Int. Conf. Adv. Intell. Mechatronics (AIM)*, Jul. 2016, pp. 1305–1311.
[27] P. Vadakkepat, K. Chen Tan, and W. Ming-Liang, "Evolutionary artificial potential fields and their application in real time robot path planning," in *Proc. Congr. Evol. Comput.*, vol. 1, 2000, pp. 256–263.
[28] S. S. Ge and Y. J. Cui, "Dynamic motion planning for mobile robots using potential field method," *Auto. Robots*, vol. 13, no. 3, pp. 207–222, 2002.
[29] H. Kim and M. J. Kim, "Electric field control of bacteria-powered microrobots using a static obstacle avoidance algorithm," *IEEE Trans. Robot.*, vol. 32, no. 1, pp. 125–137, Feb. 2016.
[30] Q. Fan, G. Cui, Z. Zhao, and J. Shen, "Obstacle avoidance for microrobots in simulated vascular environment based on combined path planning," *IEEE Robot. Autom. Lett.*, vol. 7, no. 4, pp. 9794–9801, Oct. 2022.
[31] T. Li et al., "Autonomous collision-free navigation of microvehicles in complex and dynamically changing environments," *ACS Nano*, vol. 11, no. 9, pp. 9268–9275, Sep. 2017.
[32] T. Haarnoja, S. Ha, A. Zhou, J. Tan, G. Tucker, and S. Levine, "Learning to walk via deep reinforcement learning," 2019, *arXiv:1812.11103*.
[33] C. Wang, J. Wang, Y. Shen, and X. Zhang, "Autonomous navigation of UAVs in large-scale complex environments: A deep reinforcement learning approach," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 2124–2136, Mar. 2019.
[34] T. Chu, J. Wang, L. Codeca, and Z. Li, "Multi-agent deep reinforcement learning for large-scale traffic signal control," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 3, pp. 1086–1095, Mar. 2020.
[35] W. Zhao, J. P. Queralta, and T. Westerlund, "Sim-to-real transfer in deep reinforcement learning for robotics: A survey," in *Proc. IEEE Symp. Ser. Comput. Intell. (SSCI)*, Dec. 2020, pp. 737–744.
[36] Y. Hou et al., "Design and control of a surface-dimple-optimized helical microdrill for motions in high-viscosity fluids," *IEEE/ASME Trans. Mechatronics*, vol. 28, no. 1, pp. 429–439, Feb. 2023.
[37] M. P. Kummer, J. J. Abbott, B. E. Kratochvil, R. Borer, A. Sengul, and B. J. Nelson, "OctoMag: An electromagnetic system for 5-DOF wireless micromanipulation," *IEEE Trans. Robot.*, vol. 26, no. 6, pp. 1006–1017, Dec. 2010, doi: 10.1109/TRO.2010.2073030.
[38] G. Brockman et al., "OpenAI gym," 2016, *arXiv:1606.01540*.
[39] I. C. Yasa, H. Ceylan, U. Bozuyuk, A.-M. Wild, and M. Sitti, "Elucidating the interaction dynamics between microswimmer body and immune system for medical microrobots," *Sci. Robot.*, vol. 5, no. 43, Jun. 2020, Art. no. eaaz3867.
[40] H. Wang et al., "Data-driven parallel adaptive control for magnetic helical microrobots with derivative structure in uncertain environments," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 54, no. 7, pp. 4139–4150, Jul. 2024.
[41] S. Zhong et al., "Spatial constraint-based navigation and emergency replanning adaptive control for magnetic helical microrobots in dynamic environments," *IEEE Trans. Autom. Sci. Eng.*, early access, Dec. 20, 2004, doi: 10.1109/TASE.2023.3339637.

Huaping Wang (Member, IEEE) received the B.S. degree in mechatronics and the Ph.D. degree in mechanical engineering from Beijing Institute of Technology, Beijing, China, in 2010 and 2015, respectively.

He has been a Professor with Beijing Institute of Technology, since 2022. His research interests include micro-nano robotics, micro-nano manipulation, and automation at micro-nano scales.

Yukang Qiu received the B.S. degree in mechatronics from Beijing Institute of Technology, Beijing, China, in 2021, where he is currently pursuing the M.S. degree in mechanical engineering.

His research interests include navigation control of micro-nanomachines.

Yaozhen Hou received the B.S. degree in mechatronics and the Ph.D. degree in mechanical engineering from Beijing Institute of Technology, Beijing, China, in 2017 and 2023, respectively.

He has been a Post-Doctoral Scientist in Mechanical Engineering with Beijing Institute of Technology, since 2023. His research interests include micro/nano robotics and micro/nano manupulation.

Qing Shi (Senior Member, IEEE) received the B.S. degree in mechatronics from Beijing Institute of Technology, Beijing, China, in 2006, and the Ph.D. degree in biomedical engineering from Waseda University, Japan, in 2012.

He was a Research Associate with GCOE Global Robot Academia, Waseda University, from 2009 to 2013. He is currently a Professor with the School of Mechatronical Engineering, Beijing Institute of Technology. His research interests are focused on bio-inspired robots, computer vision, and micro/nano robotics.

Hen-Wei Huang (Member, IEEE) received the B.S. and M.S. degrees in mechanical engineering from the National Taiwan University, Taiwan, in 2011 and 2012, respectively, and the Ph.D. degree in robotics technology from ETH Zürich in 2018.

He is currently an Assistant Professor with the School of Electrical and Electronic Engineering and the LKC School of Medicine, Nanyang Technological University, Singapore. His research interests include in vivo wireless sensor networks, personalized medicine, controlled drug delivery, robotics, and translational medicine.

Qiang Huang (Fellow, IEEE) was a Research Fellow with the National Institute of Advanced Industrial Science and Technology, Tokyo, Japan, from 1996 to 1999; and The University of Tokyo, from 1999 to 2000. He is currently a Professor with Beijing Institute of Technology, Beijing, China. He is the Director of the Key Laboratory of Biomimetic Robots and Systems, Ministry of Education of China.

Dr. Huang received the First Class Prize of the Ministry of Education Award for Technology Invention. He serves as the chairs for many IEEE conferences, such as the Organizing Committee Chair for the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems and the General Chair for the 2017 IEEE International Conference on Robotics and Biomimetics and the 2018 IEEE-RAS International Conference on Humanoid Robots.

Toshio Fukuda (Life Fellow, IEEE) received the B.S. degree from Waseda University, Tokyo, Japan, in 1971, and the M.S. and Ph.D. degrees from The University of Tokyo, Tokyo, in 1973 and 1977, respectively.

He is currently a Professor (1000 Foreign Experts Plan) with the School of Mechatronic Engineering, Intelligent Robotics Institute, Beijing Institute of Technology, Beijing, China, where he is mainly engaged in the research fields of intelligent robotic systems, cellular robotic systems, mechatronics, and micro/nano robotics. From 1977 to 1982, he was with the National Mechanical Engineering Laboratory, Tsukuba, Japan. From 1982 to 1989, he was with the Science University of Tokyo. In 1989, he was with Nagoya University, Nagoya, Japan, where he was a Professor with the Department of Micro System Engineering; and a Professor with Meijo University, Nagoya.

Dr. Fukuda was the President of IEEE Robotics and Automation Society (1998–1999), the Editor-in Chief of IEEE/ASME TRANSACTIONS ON MECHATRONICS (2000–2002), the Director of Division X: Systems and Control (2001–2002 and 2017–2018) and the Region 10 (2013–2014), and an IEEE Founding President of Nanotechnology Council (2002–2003 and 2005).