Force-Vision Sensor Fusion Improves Learning-Based Approach for Self-Closing Door Pulling

YUFENG SUN[1], **LIN ZHANG**[2], **AND OU MA**[1]

[1]Department of Aerospace Engineering and Engineering Mechanics, University of Cincinnati, Cincinnati, OH 45221, USA
[2]Department of Physics and Astronomy, University of Central Arkansas, Conway, AR 72035, USA

Corresponding author: Yufeng Sun (sunyf@mail.uc.edu)

ABSTRACT Multiple sensors are often used in robotic applications for better situational awareness. Hence, sensor fusion becomes a key technology to manage multiple sources of information and plays a critical role to the success in robotic tasks such as object detection and tracking, autonomous navigation, and interaction with humans. With these capabilities, wheeled autonomous vehicles can be used to automate some public services. However, there are still challenges for wheeled vehicles to safely and agilely maneuvering in human-centered environments. One of these challenges is lacking the capability of autonomously opening doors and traversing doorways without using a general-purpose robotic arm (manipulator). An autonomous door-opening operation is a complex task consisting of identifying the door and door handle, navigating the vehicle to the door, operating the door handle, and pulling or pushing the door to open while traversing the doorway. A self-closing door adds significant difficulty for the last step because the door usually needs to be held open while the vehicle is traversing the doorway. This paper presents a method using force-vision sensor fusion to enhance a deep reinforcement learning (RL) process for a wheeled vehicle to perform the most difficult step of a door-opening and pass-through operation. That step is to pull and hold a self-closing door open while the vehicle is traversing the doorway. In our solution, the vehicle is equipped with a camera, a force sensor, and a concise door-opening mechanism. The method was simulated in Gazebo and the results demonstrated that the deep RL-based force-vision sensor fusion method can be successfully applied to the task of self-closing door pulling by a wheeled vehicle without using a robotic arm and without a pre-planned trajectory. The vehicle control was trained without using domain randomization, but it still works in variant environments.
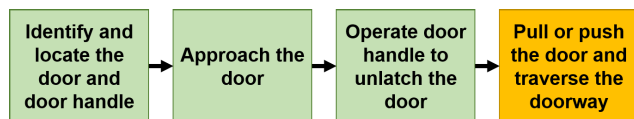
INDEX TERMS Force-vision sensor fusion, door opening, door pulling, deep reinforcement learning, autonomous mobile vehicle.

I. INTRODUCTION

Mobile vehicles equipped with multiple exteroceptive sensors (such as camera and LiDAR) for situational awareness, can be used in public services, such as material transportation, room cleaning, disinfection, and many other tasks. This is one of the promising solutions in the COVID-19 pandemic for reducing the need for human contact with the environment. However, it is still a big challenge to widely apply such servicing vehicles in human-centered environments, such as schools, office buildings, and hospitals. One of the challenging tasks is that the vehicles must be able to autonomously open doors and

The associate editor coordinating the review of this manuscript and approving it for publication was Hong-Mei Zhang.

traverse the doorways so that they can extend their working area to different rooms without human assistance. Because doors vary in size and types (such as push-/pull- doors, self-closing doors, and sliding doors), the strategies for opening a door are complicated and heavily rely on the vehicles and the sensor capabilities. Many studies [1]–[7] have investigated the door-opening problem with a mobile vehicle equipped with a manipulator, where the task was divided into four essential subtasks as shown in Fig. 1: 1) detecting the door handle, which is an object detection problem; 2) approaching the door and the door handle, which is a navigation problem; 3) operating the door handle and unlatching the door, which usually requires a multiple degrees of freedom (DOFs) manipulator to interact with the door handle; 4) opening the

FIGURE 1. Workflow of door opening by mobile vehicles.

door and traversing the doorway, which relies on the cooperation of the manipulator and the vehicle for pulling/pushing and holding the door. To address these problems in an autonomous door-opening task, most solutions adopted in the aforementioned studies relied on two costly conditions: (1) a general-purpose robotic arm of 6 or more DOFs and (2) precise modeling of the mobile manipulator and the environment. Such conditions make these solution techniques difficult to disseminate in real-world environments for broader impact due to the complexity of the control strategies and high cost of the 6-DOF manipulators. In addition, some solutions employed vision sensors for environment perception and some solutions used force sensors to address the problem of door handle operation, but none of these solutions combined the vision sensors and force sensors. We believe that the combination of the sensors can better serve in the task of autonomous door manipulation even in the most challenging case: pulling a self-closing door.

Pulling a self-closing door requires complex trajectory planning and coordinated control of both the manipulator and the mobile vehicle, thus, it was considered the most difficult part of the door-opening operation. Since a robotic system has to physically interact with the door in the door-pulling operation, force sensors can be employed as they are used in applications of high-precision manufacturing [8] and collaborative robots [9], [10] for collision detection and safety control. Vision sensors (such as cameras) are also necessary for observing the status of the door and its surrounding environment. In addition, combining the data from vision and force sensors will result in a more confident decision on the robot's motion control, just like the fact that a human relies not only on vision but also on hand sensation of force and touch when opening a door. Such a combination will benefit the robot control in terms of both safety and performance, because multiple sensors with data fusion technology can not only extract more features of the environment but also enhance understanding of the sensed environment [11], [12]. Based on this thinking, we propose a Deep Neural Network (DNN)-based force-vision sensor fusion method for enhancing a wheeled vehicle's learning to pull a self-closing door without using a general-purpose robotic arm. Instead, the vehicle is equipped with a camera, a force sensor, and a portable and cost-effective door-opening mechanism. The purpose of pulling the door is to allow the vehicle to traverse the doorway; however, we do not include the doorway traversing in this paper. One reason is that it is complicated and difficult to train the vehicle to pull the door while traversing the doorway at the same time; in addition, the vehicle does not really need to physically hold the door during a quick traversing of the already opened doorway before the door swings back; and hence, the force information may not be useful. Another reason is that the doorway traversing can be easier (as the above mentioned quick traversing without touching the already opened door), and it can be achieved by using a different approach without training. Focusing on door-pulling process, we adopt the Proximal Policy Optimization (PPO) [27] to train an end-to-end control policy (meaning from task definition and sensor raw data to vehicle control commands). Such a deep RL method does not require a system dynamics model and thus, it is naturally robust against small variations of the system's dynamics properties. The sensor fusion based control policy trained in a single simulated environment can generally work in varied environments with different visual appearances and physical properties.

The key contributions of this research are: 1) uses a cost-effective simple mechanism instead of an expensive robotic arm to help pull a self-closing door; 2) addresses the problem of door pulling using a deep RL algorithm; 3) proposes a DNN-based force-vision sensor fusion method to enhance the learning of the vehicle's control policy for the door-pulling operation; 4) demonstrates that the proposed method is robust against environment variations, sensor noise and small camera installation offsets. To our best knowledge, this is the first end-to-end solution for a wheeled vehicle to accomplish such a task.

II. RELATED WORKS
A. DOOR PULLING

Researchers have investigated the door-opening problem and developed different solutions for wheeled vehicles (with a robotic arm) [1]–[7], humanoid robots [13], [14], and legged robots [15]. In the solutions using wheeled vehicles, Chitta *et al.* presented a door-opening method by using a 7-DOF manipulator [1]. They used a graph-based representation of the three-dimensional search space to plan the trajectory of the mobile base and the robotic arm. For pulling the door open, the system checked for collision of the arm with the door and released the handle on one side of the door if a collision was detected, then moved the arm so that it can grasp the handle on the opposite side of the door. The operation only requires arm movement while keeping the mobile base stationary. Such a mobile manipulator-based "re-grasp" solution was also used in [3] for pulling a door and an extra flipper was used to hold the door open against possible self-closing forces. Different solutions were presented in [2], [4], the manipulator and wheeled vehicle were controlled simultaneously to generate a circular trajectory in Cartesian space for pulling the door open with an assumption of knowing the radius and the open angle of the door. Such solutions based on modeling of the environments are expensive and less practical in real-world applications because an accurate model of the environment is difficult to obtain. In addition, among all the above prior studies, only [3] considered self-closing doors. In fact, the self-closing function of a door adds significant

difficulty to the door-opening task, because the door has to be held open by the arm while the arm's mobile base (the vehicle) is traversing the doorway.

To address the difficulties in dynamics modeling and/or trajectory planning for door opening operations, solutions leveraging the advancement of RL techniques were presented [16], [17]. These solutions employed 7-DOF robotic arms to learn control policies for operating the door handles and pushing/pulling the door without modeling the dynamics of the robotic system and the environment nor requiring task-specific knowledge. A simulation environment "DoorGym" was also presented in [17], which supported a variety of randomized domains and difficulties. A control policy for unlatching a door using the Berkley BLUE robot arm was trained in this virtual environment and a zero-shot sim-to-real policy transfer was performed to check the robustness of the trained policy. However, these studies only focused on opening normal doors by controlling a robotic arm, they considered neither the mobile base control nor self-closing door opening.

All the above studies employed multi-DOF general-purpose robotic arms for door opening. A general-purpose robotic arm of 6 or more DOFs is a solution for multi-purpose applications which require flexibility, but it also requires complex control effort and thus it is unnecessary if the robotic arm is solely used for door opening. In addition, general-purpose robotic arms are expensive and need significant programming effort for complex tasks such as self-closing door pulling. This is one of the reasons that none of the robotic-arm based solutions has been adopted in practical services or the commercial market. Instead, a simpler and more cost-effective modular device can be designed for operating the door handle and opening the door. In this study, we assume that the door has been unlatched by such a modular device, which is a necessary operation before pulling a door but it is not a focus of this paper. In this paper, we focus on the self-closing door pulling by leveraging the vehicle's locomotion. To demonstrate this solution, we propose a method of using a skid-steering vehicle to pull a self-closing door with a passive side bar attached to the vehicle to help holding the door. Thus, we only need to develop the vehicle's control policy which can be learned from a RL process taking the information from a camera and a force sensor installed on the modular door-opening device.

B. SENSOR FUSION

Sensor fusion was defined as the cooperative use of the information provided by multiple sensors to improve accuracy and quality content and thus enhance the performance of the system [18]. It has been widely studied in object recognition and autonomous navigation, and it has boarder applications in Internet of Things (IoT), automotives, drones, computer vision, virtual reality, and healthcare domains due to its advantages of richer semantic and higher resolution on observation, better confidence in certainty and accuracy

of the data, and more comprehensive knowledge of the environment [11]. There are three fundamental ways of fusing sensor data: 1) complementary: combining the data from each sensor which provides data about different aspects or attributes of the environment; 2) competitive: fusing the data from several sensors which measure the same or similar attributes of the environment; 3) cooperative: deriving information of the environment from the data from two or more independent sensors in the system [12], [18]. Dasarathy *et al.* classified sensor fusion architectures into three levels depending on the input/output characteristics [19], namely data-level fusion, feature-level fusion, and decision-level fusion. Traditional sensor fusion algorithms, such as Kalman filter and particle filter, focus on state estimation, and Bayesian inference technique and Dempster-Shafer theory of evidence are decision fusion methods. These traditional methods of sensor fusion may suffer from problems depending on the types of sensor information to be fused. For example, 1) different sensor data needs to be transformed into a common reference frame; 2) diverse formats of the data may introduce noise and ambiguity in the fusion process; 3) the level of detail from different sensors is rarely similar. Compared with these traditional methods, artificial neural network-based sensor fusion techniques are more powerful and more adaptive for robotics applications [11].

For robot autonomy in both indoor and outdoor environments, cameras are the most popular sensors as they provide rich information. Hence, sensor fusion techniques for combining vision information and depth information from LiDAR or ultrasonic distance data were usually applied to autonomous navigation [20], [21] using the DNN-based sensor fusion approach due to its capability of multi-level feature representation. Deep RL techniques were also utilized to teach mobile vehicles to avoid obstacles and navigate in indoor environments using sensors such as cameras and laser range finders [22], [23]. Force-torque sensors were usually installed on end-effectors or joints to indirectly measure the gripper contact force for both safety control and force control on workpieces [9], [10] in manufacturing applications and collaborative robots. They are sometimes also combined with vision sensors for better contact controlling of robot arms. A hybrid force-vision control law was presented in [24] for a robotic arm to perform grasping tasks, which used a Convolutional Neural Networks (CNN)-based vision controller in the reaching stage, and a force-contact Proportional Integral (PI) controller in grasping stage of opening a drawer. However, this sensor fusion method must be incorporated with a low-level control law based on a computed torque technique which requires a dynamics model of the robotic arm. Similar hybrid force-vision feedback control methods for robotic arm control were also presented in [25], [26], where the vision systems were used to extract features for estimating the systems' motion states in closed-loop controls which also involved the contact forces of the end-effectors. These studies concluded that the hybrid force-vision control methods could overcome the uncertainties of kinematics,
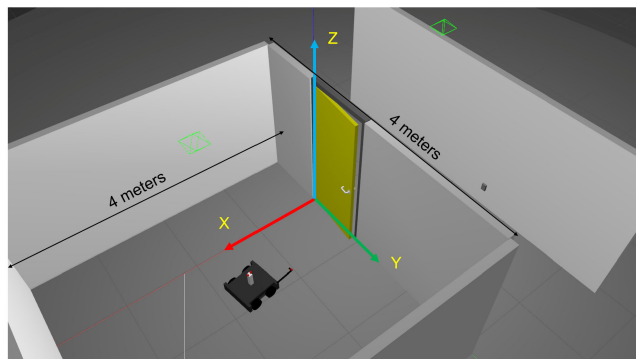
FIGURE 2. A simulated environment for door-pulling task.

The right side bar is released for pulling a right-hand swing door in a room

FIGURE 3. Configuration of the wheeled vehicle.

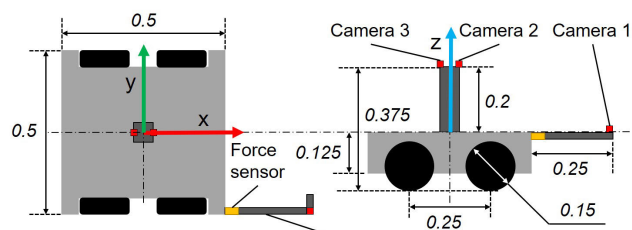dynamics and camera models and resulted in a robust motion control of the robotic arm.

C. POLICY GENERALIZATION

Deep RL methods have been successfully applied in board games, video games, and simulated control problems, but they are inefficient and the learning process often requires millions of attempts to solve a complex task. This is often impossible for a real-world robotic application because the cost of acquiring data is extremely high in the real world and the number of duty cycles of robot hardware is also limited. One solution is to train a control policy in a simulated environment, and then transfer the learned policy to a physical robot. This is known as simulation to reality (Sim2Real) transfer [28], [29], which is still an open research problem. To ensure the generalization capability of learned policies and bridge the gap between the simulation and the reality, domain randomization [30] is often used. However, it suffers from high sample complexity and requires customizing software which is capable of simulating the intricacies (i.e. visual appearances and dynamics) of the real world, such as the simulation environment "DoorGym". For the task of pulling a self-closing door by a wheeled vehicle, it is cumbersome to simulate visual appearances. However, the forces exerted on the door and the vehicle do not vary too much. Based on this thinking, we used the force-vision sensor fusion method to improve the policy transferability. Using both vision and force information, we trained the vehicle in a single simulated environment without randomizing the properties of the environment and found out that the learned policy can be applied to the environments that were never experienced before.

III. METHOD
A. ENVIRONMENT SETUP

Without loss of generality, a 4 meters × 4 meters room with a self-closing and right-hand swing door was built in Gazebo. The standard door is 4.5 centimeters thick, 0.9 meters wide, and 2.1 meters high. It has a uniformly distributed mass of 10 kilograms and needs to be pulled to open from inside of the room as shown in Fig. 2. The door hinge axis is aligned with the Z axis of the global coordinate system where the

X axis is pointing to the inside of the room. The dynamics of the door is modeled with a torsional spring for its self-closing function. The door can swing about its hinge axis and the door position can be represented by the hinge angle with 0 radians in the fully closed position and 0.5π radians in the fully open position. We assume that the door has been unlatched and pulled to be slightly open by modular door-opening device, which is realistically achievable but not the focus of this paper. With this initial condition of door position, a skid-steering wheeled vehicle equipped with a camera, a force sensor, and a spring-loaded passive side bar (for door holding) is placed at a certain position with small uncertainty, but the side bar must be placed near the gap of the door and door frame as the initial condition for pulling the door.

B. WHEELED VEHICLE CONFIGURATION

Instead of using a mobile manipulator which was adopted in other studies for the door-pulling operation, a pair of passive side bars were installed on a wheeled vehicle to help pull the door in our study. As shown in Fig. 3, the vehicle is assumed to be in a square shape with 0.5 meters in length and width. A 0.25 meters long side bar with a small hook is attached on the vehicle for holding the door against its self-closing force. Each spring-loaded side bar is normally folded in in its home position parallel with the vertical edge of the vehicle body, so that it will not touch the environment when not in use. It will be released to its horizontal work configuration, as shown in the figure, by the door-pulling controller using a simple solenoid device. After completion of a door opening task, the passive side bar can be retracted and locked to its vertical home position. As a door can be either left-hand swing or right-hand swing, a side bar can be installed at each side of the vehicle. A force sensor is installed near the joint of the side bar and the vehicle body, which can detect the contact forces exerted on the side bar. A camera (camera 1) is installed at the end of the side bar looking upward to observe the status of the door, which is used in the force-vision sensor fusion method for the door-pulling task. Only for analytic purposes, we temporarily install two other cameras (camera 2 and camera 3) on the vehicle body looking forward and backward respectively and we assume that all the cameras are well calibrated. We also developed a method of using observations from all three cameras, a method of using force only, and a method of using single camera (camera 1)
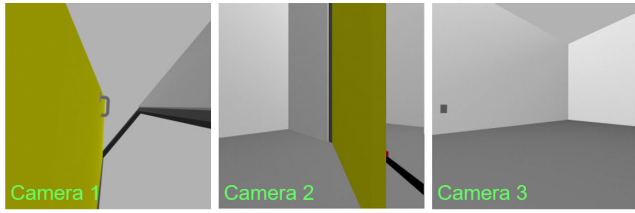
FIGURE 4. Camera views of the door status.

to control the vehicle for door pulling. The raw observation from each camera is a three-channeled RGB image, which partially represents the environment, as shown in Fig. 4. The image is converted to a grayscale and fed into a CNN which is specifically designed for door pulling. The output of the neural network is a high-level control command in the vehicle frame for maneuvering the vehicle in the 3-DOF planar space. The control command is a desired velocity; however, the actual velocity of the vehicle will be affected by its physical properties and the environment. The uncertain behaviors of the vehicle when responding to the control commands is also expected to be handled in our DNN-based sensor fusion method.

C. INITIAL CONDITION

In this paper, we only focus on the subtask of pulling a self-closing door with a wheeled vehicle and a side bar, so we assume that the door is already unlatched and slightly open, which is the initial condition of the door-pulling task. Such an initial condition can be achieved by controlling the vehicle and the modular device which is a task less challenging than the door-pulling task. Assuming the vehicle is in the room, it can be steered to search for the door and door handle based on the camera image, as shown in Fig. 11 in Appendix A, using a vision-based door and door handle detection method [7]. Once the door handle is identified in the camera image, the distance to the door handle can be measured by a RGB-D sensor, then feedback control can be applied to drive the vehicle to approach the door handle. The modular device with a force sensor is used to operate the door handle to unlatch the door based on the force information. After the door is unlatched, the vehicle can pull the door slightly open by driving slightly backward. Thus, the initial condition of pulling a self-closing door can be set after unfolding the side bar and rotating the vehicle to make the side bar touch the edge of the door. Such an operation of the vehicle for door unlatching is shown in Fig. 5.

D. DNN STRUCTURE FOR FORCE-VISION SENSOR FUSION

A DNN was designed for handling the vision and force inputs. The vision input is a grayscale image with a size of 64 × 64 pixels and the camera captures images at a frequency of 30 fps. The force input is a 3 × 1 vector which is the force between the side bar and door. The force sensor collects data at a frequency of 100 Hz and the force data is smoothed using a simple moving average method. At a frequency of 2 Hz,
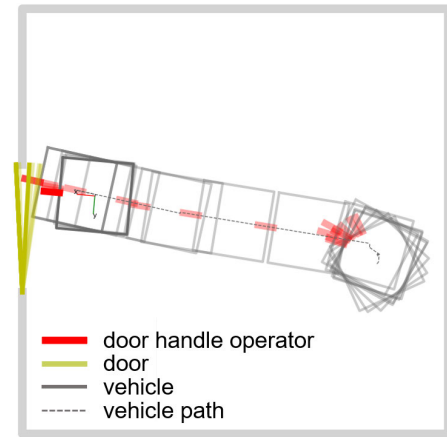

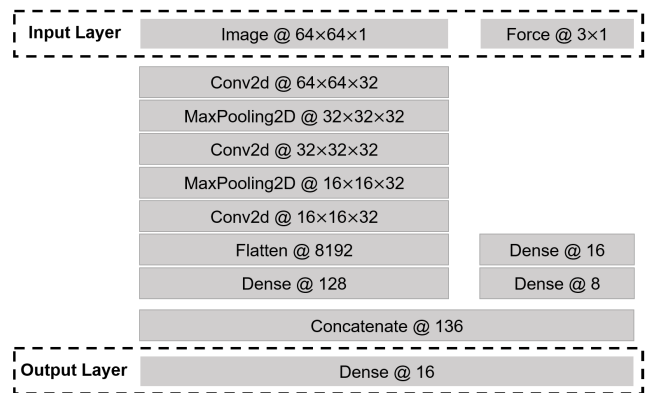**FIGURE 5.** A path of the wheeled vehicle for unlatching a door.

FIGURE 6. DNN structure of the force-vision sensor fusion based control policy.

the force-vision sensor fusion DNN takes the images and force data as the input. The output of the neural network is a probability of possible desired velocity commands for controlling the vehicle. The possible control commands are driving forward, driving backward, turning left, turning right, driving left forward, driving right forward, driving left backward, and driving right backward. Each control can be performed in a low-speed mode with a desired linear speed of 1 meter per second and desired angular speed of π radians per second and a high-speed mode with desired linear speed of 3 meters per second and desired angular speed of 3π radians per second, thus, the dimensions of the action space is 16 in total. The DNN structure is given in Fig. 6, which has 2 convolution layers for the image input followed with a max pooling operation, and a third convolution layer reducing the image size to 16 × 16 pixels before the flattening layer. The force input is connected to a 2-level hidden layer and then concatenated with a hidden layer for the image data after the flattening layer and finally connected to the output layer, resulting in more than one million trainable weights in the whole neural network. A lighter version of DNN with a much smaller number of weights was also designed and trained. The trained model can also be applied for pulling the self-closing

door in variant environments. However, the learning performance decreased and the trajectory of pulling the door became more costly.

E. RL CONTEXT

In our study, the vehicle starts from the initial position and maneuvers itself to pull the door until the door angle reaches 0.45π radians. For each time step, the vehicle observes the current state (the image from camera 3 and the force data) and takes an action (one of the 16 velocity commands) determined by a control policy, then transfers to a new state, and receives a reward at the next time step. The reward function (r) is defined as

$$r = c_1 \times 100 - c_2 \times 10 + c_3 \times (10 \times \Delta\alpha - p - 0.1) \quad (1)$$

where c_1 is 1 when the door-pulling is successful (i.e., the door angle reaches 0.45π), otherwise it is 0. c_2 is 1 when the door-pulling failed (i.e., the door angle is less than 0.45π and the side bar is away from the door with a distance greater than 5 centimeters), otherwise it is 0. c_3 is 1 if $c_1 = c_2 = 0$, it is 0 if $c_1 = 1$ or $c_2 = 1$. $\Delta\alpha$ denotes the door angle change after each time step of driving the vehicle. A simple step penalty with a constant value of 0.1 is added for achieving a goal of optimizing control with fewer steps in an episode. A force penalty p is added, where the value of p is 1 if the magnitude of the detected force in any direction exceeds a maximum value of 70 Newtons (a value established by ADAAG, ICC/ANSI A117.1 Standard on Accessible and Usable Buildings and Facilities [31]), otherwise, p is 0, which means no force related penalty will be applied. By considering the force in the reward function, we expect a smooth motion control in the door-pulling operation after training. We use a PPO algorithm [27] to train the force-vision DNN model, because PPO algorithms are more general and much simpler to implement and they have better sample complexity compared with other policy gradient algorithms. We adopted Actor-Critic implementation for the PPO agent. The "Actor" network represents the control policy, its architecture is shown in Fig. 6. The "Critic" network is for state value evaluation, which has the same architecture as the "Actor" network except the output layer. The output layer of "Critic" network is a scalar which estimates how good the state is.

IV. SIMULATION AND RESULTS

The door-pulling operation by a wheeled vehicle was simulated in Gazebo. We used a DNN-based force-vision sensor fusion method (camera 1 and the force sensor were used in our method) for developing the door-pulling control policy. We used a specific simulated environment with unchangeable visual appearances (e.g., color, texture), fixed door size, fixed number of springs, and fixed friction coefficients between wheels and ground for training. But the initial conditions can change, we randomly initialized the pose (adding small variations in x, y and orientation) of the vehicle during the training process. To verify the robustness of the force-vision sensor fusion method, we also compared it with three other methods:
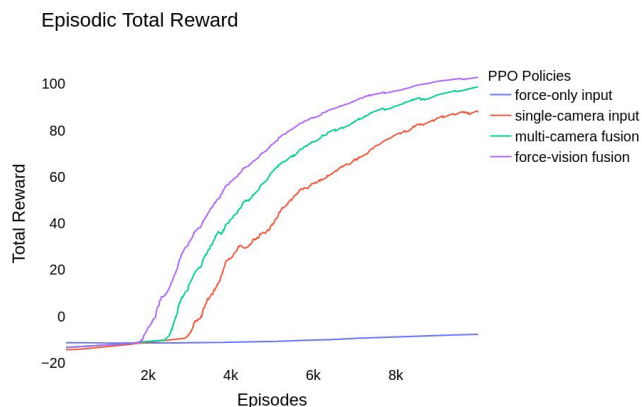
FIGURE 7. Learning curves of different control policies.

1) using force sensor only; 2) using single camera (camera 1) only; and 3) using multiple cameras (all three cameras), in terms of the training performance, control optimization and policy generalization in varied simulated environments with different visual appearances and physical properties.

A. COMPARISON OF TRAINING PERFORMANCE

All the training was performed on a desktop computer with an Intel i7-8700 CPU, an Nvidia GeForce GTX 1070 GPU and 32 GB memory. Each training was set to run 10,000 episodes using the hyper-parameters listed in Table 4 in Appendix B.

The training using only the force sensor data failed to come up with a useful policy for controlling the vehicle to pull the door open due to the insufficient information provided to the vehicle. The training for the other three methods were successful for the door-pulling task. Overall, the method of using force-vision fusion outperformed that of using multi cameras and the method of using a single camera in terms of both training efficiency and gained episodic total rewards, as shown in Fig. 7. With more information from the environment, the vehicle learned faster. As we can see, the force information played an important role for the vehicle to understand the environment from a different perspective even with less vision information.

B. CONTROL POLICY ASSESSMENT

For the three successfully trained policies, we tested them in the same training environment. We tested the performance of each control policy for 100 times with a small variation to the initial pose of the vehicle at each time. The success rate and average steps required to complete the task can be used to assess the control policy. As shown in Table 1, all the three control policies achieve a high success rate over 98% in the training environment. The simulation parameters are listed in Table 3 and environment appearances are shown in Fig. 12 to Fig. 17 in Appendix C.

As we discretized the actions of moving the vehicle and considered a constant step penalty for both angular and linear movements, the trained door-pulling process should require

TABLE 1. Success rate and number of required steps for door pulling.

Policy	Success rate	Average number of steps	Range of steps
Single camera input	100%	16	13-20
Multi-camera fusion	98%	13	10-47
Force-vision fusion	100%	11	10-13

TABLE 2. Shortest travel distance.

Policy	Vehicle	Side bar
Single camera input	0.835 m	1.247 m
Multi-camera fusion	0.893 m	1.108 m
Force-vision fusion	0.571 m	1.105 m
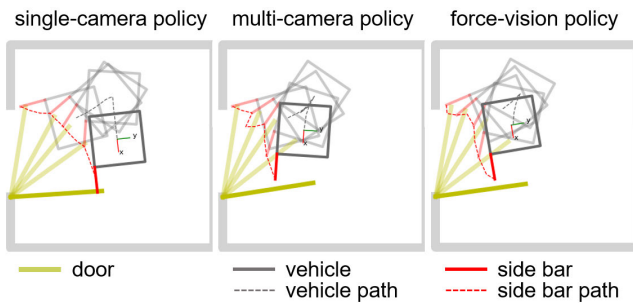
FIGURE 8. Traces of the vehicle and side bar motion configurations for maneuver distance comparison.

FIGURE 9. Force data of the test scenarios under different control policies.
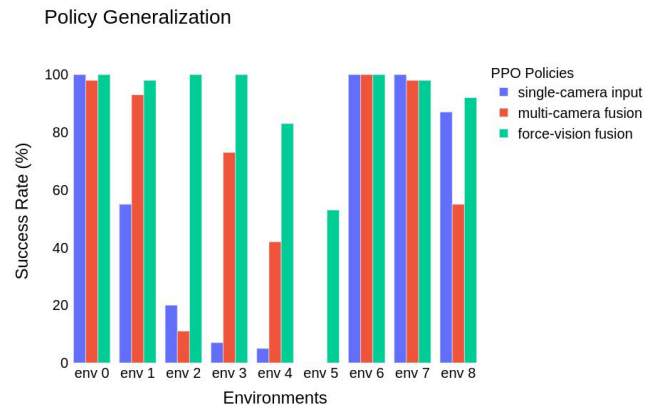
FIGURE 10. Success rate of different control policies.

as few steps as possible. From Table 1, the force-vision fusion method required the least average number of steps and the smallest range from minimum to maximum steps for completing the task, which was significantly more optimal than the method of using a single camera. The method of using all three cameras required fewer average steps than the method of using a single camera but had the largest variance of steps, which indicated that using multiple sensors of the same type was not always better for a task such as door pulling.

By analyzing the test scenarios with the shortest number of steps under the control of all the three policies, as shown in Table 2 and Fig. 8, we also found that the maneuver of pulling a self-closing door under the policy of the force-vision fusion method requires the shortest displacement of both the vehicle itself and the end of the side bar, which infers that the force-vision fusion based control policy provides the best result in terms of travel distance.

The sensed force data of the test scenarios can also tell us whether the door-pulling process was smooth or not. As shown in Fig. 9, by taking the force information as input, the door-pulling process trained with the force-vision sensor fusion policy has smoother force curves in all the three directions, as there are less fluctuations in the contact force. It also indicates that the actions with the force-vision sensor fusion policy are more concise and efficient than those of the other two policies using cameras only.
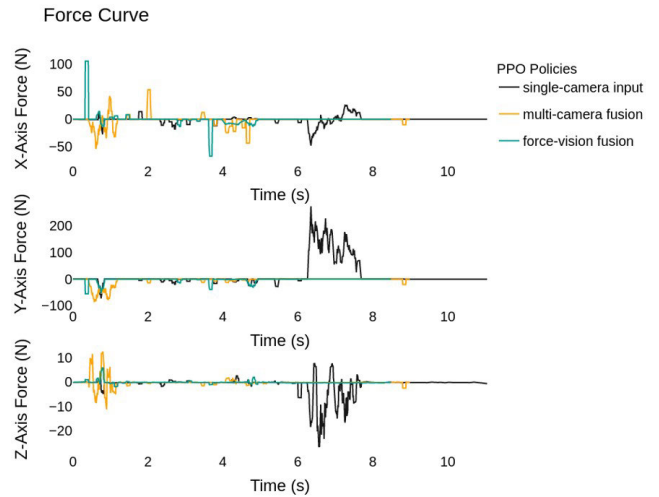
C. POLICY GENERALIZATION

In order to verify the robustness of the trained policies against environmental uncertainties, we did the same test in 8 different simulated environments (simulation parameters are listed in Table 3) besides the original training environment (env 0). These environments vary in door appearance (e.g., color, texture), door width, number of door springs, ground friction, and lighting conditions. Camera noise and small translational and rotational offsets were also introduced in the tests. As we trained the policies to pull a right-hand swing door, we also tested the policies to pull a left-hand swing door without retraining the network.

From the test results shown in Fig. 10, the single-camera policy has the worst generalization capability as the vehicle barely succeeded to pull the door open in the other environments. The vehicle performed better in most environments using the multi-camera policy. However, in environment 2 which has more complex wall texture and worse lighting conditions, the performance of the multi-camera policy is even worse, which further indicates that it is not always better to use multiple sensors of the same type. Unsurprisingly,

TABLE 3. Simulation parameters.

Properties	env 0	env 1	env 2	env 3	env 4	env5	env 6	env 7	env 8
door mass	10 kg	10 kg	20 kg	15 kg	30 kg	20 kg	10 kg	10 kg	10 kg
door width	0.9 m	0.9 m	0.9 m	0.75 m	1.05 m	0.9 m	0.9 m	0.9 m	0.9 m
door color	yellow	wood	red	pallet	gray	red	yellow	yellow	yellow
number of spring	2	2	3	2	3	3	2	2	2
door frame color	gray	wood	black	black	black	black	gray	gray	gray
door swing direction	right	right	right	right	right	right	right	right	left
door handle color	white	white	white	white	gold	white	white	white	white
wall color	white	painted	bricks	yellow	green	bricks	white	white	white
lighting constant	1	0.2	1	0.5	0.2	1	1	1	1
ground friction μ_1	0.98	0.98	0.5	0.7	0.7	0.7	0.98	0.98	0.98
ground friction μ_2	0.98	0.98	0.5	0.98	0.98	0.98	0.98	0.98	0.98
camera translation	no	no	no	no	no	no	yes	no	no
camera rotation	no	no	no	no	no	no	no	yes	no
camera noise	no	no	no	no	no	yes	no	no	no

the force-vision sensor fusion policy shows the best generalizing capability as it achieved over 98% success rate in most environments except environments 4 and 5. In environment 4, the door width was increased to 1.05 meters, and the number of springs were increased as well. Thus, the vehicle demanded a larger force to pull the door open, so that it failed in some test cases but still achieved an 83% success rate. Environment 5 is the same as environment 2, but with noise added to the cameras (a Gaussian distribution noise mode with a zero mean and 0.02 variance in images). It turned out that the polices based on vision alone cannot handle the camera noise at all, but the force-vision fusion policy still achieved 53% chance to pull the door open. This also tells us that the force information from the interaction is very important in the task of door pulling by a wheeled vehicle. It is worthy of mentioning that, by slightly changing the height (5 centimeter lower) and position (2 centimeters back) of camera 1, as in environment 6, the vehicle could still succeed under the control of all policies. And in environment 7, a small angular offset of camera 1 was considered to verify the robustness of the trained policies. By rotating the camera about the z-axis by 5 degrees, the success rate of the originally trained single-camera input policy remained 100%, while the success rates of the originally trained multiple-camera fusion and the force-vision sensor fusion policies slightly dropped to 98%. These results indicate that the trained policies are robust against small translational and rotational offsets of the camera pose. In addition, we tested the trained policies in environment 8 with a left-hand swing door by using the same vehicle with the side bar installed on the left. We simply flipped the image from camera 1, reversed the measured force value with respect to the y-axis, and also flipped the rotation direction of the vehicle. The test results show that the resulting success rate of the force-vision sensor fusion policy is 92%, which is higher than the single-camera input policy (87%), and the multi-camera fusion policy (55%). The lower success rate with the multi-camera fusion policy may be caused by the other two cameras having different

observations in the new test environment. These results indicate that, for the left-hand side bar control, the policy trained with the right-hand side bar case can be reused as is (without retraining), providing simply "mirroring" the inputs and outputs.

Different from the approach of using domain randomization to achieve better policy generalization, our force-vision sensor fusion method can achieve good policy generalization by only training in a single simulated environment. This is extremely important for DRL-based approaches in robotic applications, where domain randomization currently is a popular solution for bridging the Sim2Real gap although it suffers from high sample complexity.

V. CONCLUSION

This paper presents an autonomous wheeled vehicle using reinforcement learning and deep neural network-based force-vision sensor fusion to address the problem of pulling a self-closing door, which is the most challenging task in a door-opening operation. Our approach relies only on a simple passive side bar attached to the vehicle instead of an expensive multi-DOF manipulator. The method can successfully train a robust control policy from a specifically simulated environment without using domain randomization or other main-stream transfer learning techniques. The learned force-vision sensor fusion-based control policy can successfully perform the operation in other environments with different visual appearances, door width, ground friction, number of door springs, and lighting conditions than those in the training environment. In addition, the tests have also shown that the force-vision sensor fusion-based control policy is robust against sensor noise and small translational and rotational offsets of the camera. It can also be reused as is to pull a left-hand swing door, which is different from the one used in training the policy, simply by mirroring the inputs and outputs without retraining the policy. Further, the trained policy does not need a preplanned motion trajectory for the vehicle to perform the operation.
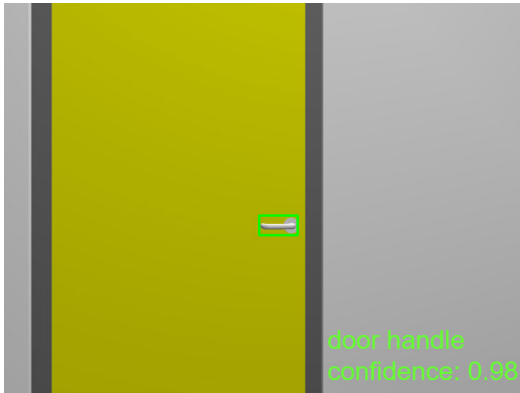
FIGURE 11. Door handle detection in a camera image.

TABLE 4. Hyper-parameters of PPO learning.

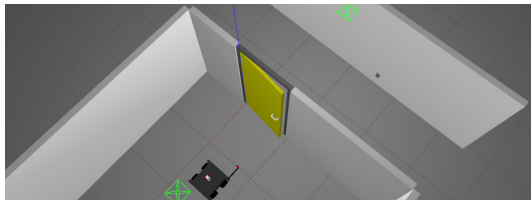| Parameter	Value
discount rate γ	0.99
clip ratio	0.2
batch size	500
learning rate of actor	1e-4
learning rate of critic	3e-4
cross entropy influence β	1e-3

FIGURE 12. Env 0 (for training) and env 6-7 (for test).

The future work will be to transfer the learned policies from the simulation domain to a real-world door pulling experiment, which is still an open research problem. To deploy the solution on a real-world mobile vehicle, a lighter version of the force-vision sensor fusion based DNN may be implemented using weight pruning techniques. We anticipate that the force-vision sensor fusion method can be transferred to a real hardware environment without too much tuning as it has been shown robust against environment variations, sensor noise and small camera installation offsets.

APPENDIX A
DOOR HANDLE DETECTION IN A CAMERA IMAGE
See Figure 11.

APPENDIX B
HYPER-PARAMETERS OF PPO LEARNING
See Table 4

APPENDIX C
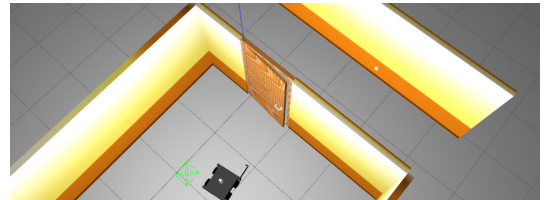## APPEARANCES OF SIMULATION ENVIRONMENTS
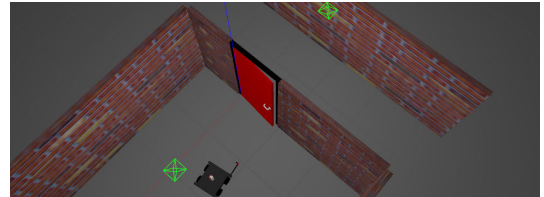See Figures 12–17.

FIGURE 13. Env 1 (for test.)
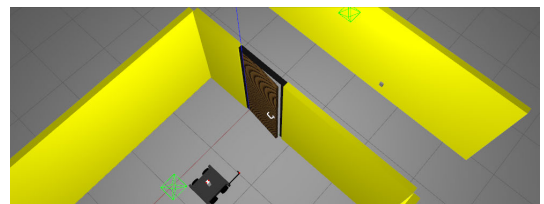

FIGURE 14. Env 2 (for test) and env 5 (for test).

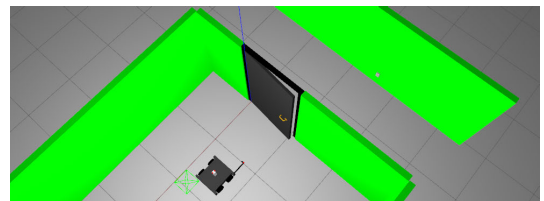**FIGURE 15.** Env 3 (for test).

FIGURE 16. Env 4 (for test).

FIGURE 17. Env 8 (for test).

REFERENCES

[1] S. Chitta, B. Cohen, and M. Likhachev, "Planning for autonomous door opening with a mobile manipulator," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2010, pp. 1799–1806.

[2] J. Li, J. Tao, L. Ding, H. Gao, Z. Deng, and K. Xia, "Twisting door handles and pulling open doors with a mobile manipulator," in *Proc. IEEE Int. Conf. Robot. Biomimetics (ROBIO)*, Dec. 2015, pp. 686–691.

[3] B. Axelrod and W. H. Huang, "Autonomous door opening and traversal," in *Proc. IEEE Int. Conf. Technol. Practical Robot Appl. (TePRA)*, May 2015, pp. 1–6.

[4] H. Gao, C. Ma, L. Ding, H. Yu, K. Xia, H. Xing, and Z. Deng, "Dynamic modeling and experimental validation of door-opening process by a mobile manipulator," *IEEE Access*, vol. 7, pp. 80916–80927, 2019.

[5] M. Stuede, K. Nuelle, S. Tappe, and T. Ortmaier, "Door opening and traversal with an industrial Cartesian impedance controlled mobile robot," in *Proc. Int. Conf. Robot. Autom. (ICRA)*, May 2019, pp. 966–972.

[6] S. A. Prieto, A. Adán, A. S. Vázquez, and B. Quintana, "Passing through open/closed doors: A solution for 3D scanning robots," *Sensors*, vol. 19, no. 21, p. 4740, Oct. 2019.

[7] M. Arduengo, C. Torras, and L. Sentis, "Robust and adaptive door operation with a mobile robot," 2019, *arXiv:1902.09051*. [Online]. Available: http://arxiv.org/abs/1902.09051

[8] G. Hao, M. Murphy, and X. Luo, "Development of a compliant-mechanism-based compact three-axis force sensor for high-precision manufacturing," in *Proc. ASME/IEEE Int. Conf. Mech. Embedded Syst. Appl.*, vol. 9, Aug. 2015, Art. no. V009T07A082.

[9] E. Mariotti, E. Magrini, and A. D. Luca, "Admittance control for human-robot interaction using an industrial robot equipped with a F/T sensor," in *Proc. Int. Conf. Robot. Autom. (ICRA)*, May 2019, pp. 6130–6136.

[10] W. Li, Y. Han, J. Wu, and Z. Xiong, "Collision detection of robots based on a force/torque sensor at the bedplate," *IEEE/ASME Trans. Mechatronics*, vol. 25, no. 5, pp. 2565–2573, Oct. 2020.

[11] M. L. Fung, M. Z. Q. Chen, and Y. H. Chen, "Sensor fusion: A review of methods and applications," in *Proc. 29th Chin. Control Decis. Conf. (CCDC)*, May 2017, pp. 3853–3860.

[12] L. Kong, X. Peng, Y. Chen, P. Wang, and M. Xu, "Multi-sensor measurement and data fusion technology for manufacturing process monitoring: A literature review," *Int. J. Extreme Manuf.*, vol. 2, no. 2, Mar. 2020, Art. no. 022001.

[13] M. Zucker, Y. Jun, B. Killen, T.-G. Kim, and P. Oh, "Continuous trajectory optimization for autonomous humanoid door opening," in *Proc. IEEE Conf. Technol. Practical Robot Appl. (TePRA)*, Apr. 2013, pp. 1–5.

[14] N. Banerjee, X. Long, R. Du, F. Polido, S. Feng, C. G. Atkeson, M. Gennert, and T. Padir, "Human-supervised control of the ATLAS humanoid robot for traversing doors," in *Proc. IEEE-RAS 15th Int. Conf. Humanoid Robots (Humanoids)*, Nov. 2015, pp. 722–729.

[15] Z.-J. Chen, F. Gao, and Y. Pan, "Novel door-opening method for six-legged robots based on only force sensing," *Chin. J. Mech. Eng.*, vol. 30, no. 5, pp. 1227–1238, Sep. 2017.

[16] B. Nemec, L. Zlajpah, and A. Ude, "Door opening by joining reinforcement and intelligent control," in *Proc. 18th Int. Conf. Adv. Robot. (ICAR)*, Jul. 2017, pp. 222–228.

[17] Y. Urakami, A. Hodgkinson, C. Carlin, R. Leu, L. Rigazio, and P. Abbeel, "DoorGym: A scalable door opening environment and baseline agent," 2019, *arXiv:1908.01887*. [Online]. Available: http://arxiv.org/abs/1908.01887

[18] M. B. Alatise and G. P. Hancke, "A review on challenges of autonomous mobile robot and sensor fusion methods," *IEEE Access*, vol. 8, pp. 39830–39846, 2020.

[19] B. V. Dasarathy, "Sensor fusion potential exploitation-innovative architectures and illustrative applications," *Proc. IEEE*, vol. 85, no. 1, pp. 24–38, Jan. 1997.

[20] J. Kocic, N. Jovicic, and V. Drndarevic, "Sensors and sensor fusion in autonomous vehicles," in *Proc. 26th Telecommun. Forum (TELFOR)*, Nov. 2018, pp. 420–425.

[21] N. Patel, A. Choromanska, P. Krishnamurthy, and F. Khorrami, "Sensor modality fusion with CNNs for UGV autonomous driving in indoor environments," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2017, pp. 1531–1536.

[22] X. Ruan, D. Ren, X. Zhu, and J. Huang, "Mobie robot navigation based on deep reinforcement learning," in *Proc. Chin. control Decis. Conf. (CCDC)*, Jun. 2019, pp. 6174–6178.

[23] J. Xin, H. Zhao, D. Liu, and M. Li, "Application of deep reinforcement learning in mobile robot path planning," in *Proc. Chin. Autom. Congr. (CAC)*, Oct. 2017, pp. 7112–7116.

[24] N. Djelal, N. Saadia, and A. R. Cherif, "Hybrid force-vision control law based on convolutional neural networks," in *Proc. 4th Int. Conf. Smart City Appl.*, Oct. 2019, pp. 1–5.

[25] N. Djelal, N. Saadia, and A. Ramdane-Cherif, "Adaptive force-vision control of robot manipulator using sliding mode and fuzzy logic," *Autom. Control Comput. Sci.*, vol. 53, no. 3, pp. 203–213, May 2019.

[26] M. Dehghan, Z. Zhang, M. Siam, J. Jin, L. Petrich, and M. Jagersand, "Online object and task learning via human robot interaction," in *Proc. Int. Conf. Robot. Autom. (ICRA)*, May 2019, pp. 2132–2138.

[27] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," 2017, *arXiv:1707.06347*. [Online]. Available: http://arxiv.org/abs/1707.06347

[28] X. B. Peng, M. Andrychowicz, W. Zaremba, and P. Abbeel, "Sim-to-real transfer of robotic control with dynamics randomization," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2018, pp. 3803–3810.

[29] S. Hofer, K. Bekris, A. Handa, J. C. Gamboa, M. Mozifian, F. Golemo, C. Atkeson, D. Fox, K. Goldberg, J. Leonard, C. Karen Liu, J. Peters, S. Song, P. Welinder, and M. White, "Sim2Real in robotics and automation: Applications and challenges," *IEEE Trans. Autom. Sci. Eng.*, vol. 18, no. 2, pp. 398–400, Apr. 2021.

[30] J. Tobin, R. Fong, A. Ray, J. Schneider, W. Zaremba, and P. Abbeel, "Domain randomization for transferring deep neural networks from simulation to the real world," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2017, pp. 23–30.

[31] *Americans With Disabilities Act: Accessibility Guidelines for Buildings and Facilities (ADAAG)*, US Access Board, Architectural and Transportation Barriers Compliance Board, Washington, DC, USA, 1998.

YUFENG SUN received the M.S. degree in mechanical engineering from Huazhong of Science and Technology, Hubei, China, in 2007. He is currently pursuing the Ph.D. degree with the University of Cincinnati. From 2007 to 2018, he was a Software Engineer at Autodesk Inc., working on CAD software design and development. He is currently a Research Assistant at the University of Cincinnati. His research interests include robotics, computer vision, and intelligent control.

LIN ZHANG received the B.S. degree in automation from Harbin Institute of Technology, Heilongjiang, China, in 2007, and the Ph.D. degree in engineering from New Mexico State University, Las Cruces, NM, USA, in 2016. He is currently an Assistant Professor with the Department of Physics and Astronomy in University of Central Arkansas. His major research interests include deep reinforcement learning, robotics, and intelligent systems.

OU MA received the B.S. degree in mechanical engineering from Zhejiang University, in 1982, and the Ph.D. degree from McGill University, in 1991. From 1991 to 2002, he was a Senior Project Engineer and a Research and Development Lead with MDA Space Systems (MD Robotics). From 2002 to 2017, he was an Associate Professor, a Professor, and a John Nakayama and Tome Nakayama Professor with New Mexico State University. Since 2017, he has been the Alan B. Shepard Chair Professor with the Department of Aerospace Engineering and Engineering Mechanics, University of Cincinnati. His research interests include multi-body dynamics and control, impact-contact dynamics, intelligent control of robotics and autonomous systems, human–robot interaction and collaboration, and smart manufacturing.

● ● ●