Learning-Based Control and Localization of Magnetic Soft Millirobots

Von dem Stuttgarter Zentrum für Simulationswissenschaften (SC SimTech) der Universität Stuttgart zur Erlangung der Würde eines Doktor-Ingenieurs (Dr.-Ing.) genehmigte Abhandlung

Vorgelegt von

Sinan Özgün Demir

aus Istanbul, Türkei

Hauptberichter:	Prof. Dr. Metin Sitti
Mitberichter:	Prof. Dr. Syn Schmitt
Mitberichter:	Prof. Dr. Georg Martius

Tag der mündlichen Prüfung: 23.04.2024

Max-Planck-Institut für Intelligente Systeme

2024

Learning-Based Control and Localization of Magnetic Soft Millirobots

A thesis accepted at the Stuttgart Center for Simulation Science in fulfillment of the requirements for the degree
Doctor Engineer (Dr.-Ing.)

Presented by
Sinan Özgün Demir
from Istanbul, Türkiye

Main Reporter:	Prof. Dr. Metin Sitti
Co-reporter:	Prof. Dr. Syn Schmitt
Co-reporter:	Prof. Dr. Georg Martius
Date of the oral exam:	23.04.2024

Max-Planck-Institute for Intelligent Systems
2024

Aileme

Abstract

Soft millirobots have promising biomedical applications due to mechanical compliance, absorbing excess forces without additional computational effort, and multifunctionality. Especially with wireless multimodal locomotion capabilities, magnetic soft millirobots (i.e., ≤1 cm) have emerged as potential minimally invasive medical robotic platforms as they can access confined and hard-to-reach spaces in the human body (e.g., distal vascular regions), and carry out medical applications, such as on-demand drug delivery, sensing, and embolization, in a target location. For such potential biomedical applications, the adaptivity of the robot control is essential to ensure the continuity of the operations, as task environment conditions show dynamic variations that can alter the robot's performance. However, fabrication-, material-, physical-interaction-dependent variations, and complex kinematics with virtually infinite degrees of freedom arising from the nature of their soft material structure limit the applicability of the conventional modeling and control methods.

The main objective of this dissertation is to establish a data-efficient adaptive multimodal locomotion framework for the targeted application scenarios of magnetic soft millirobots. To this end, a probabilistic learning approach leveraging Bayesian optimization (BO) and Gaussian processes (GPs) is introduced to address the controller adaptation challenge. First, the efficacy of the BO to fabrication variabilities is shown on three different robots fabricated following the same steps. Next, through augmented tests on benchmark datasets, it is shown that

transferring the posterior mean learned by one robot as the prior mean to the other robots and test cases improves the learning performance of BO by achieving quicker gait adaptation. Afterward, the controller adaptation method employing the proposed transfer learning approach is demonstrated in various task spaces with varying surface adhesion, surface roughness, and medium viscosity properties. To further improve the adaptation performance by including multimodal locomotion, the sim-to-real transfer learning method is developed in the third study. In this regard, a data-driven simulation environment is designed, and its accuracy is demonstrated by comparing the simulated results to the physical experiments. Leveraging the simulated experience and BO based transfer learning, it is demonstrated that sim-to-real transfer learning provides efficient locomotion learning. Furthermore, the adequacy of the automated locomotion adaptation through the Kullback-Leibler divergence-based domain recognition approach is shown to changing environmental conditions.

As the secondary objective, a new localization method using electrical impedance tomography is introduced. The applicability of the proposed approach is demonstrated for stationary and moving cases in environments with and without any obstacles.

With these contributions, this thesis proposes a domain-adaptive locomotion learning framework enabling the soft millirobot locomotion to quickly and continuously adapt to environmental changes while exploring actuation space.

Zusammenfassung

Weiche Milliroboter haben vielversprechende biomedizinische Anwendungen, da sie mechanisch nachgiebig sind, überschüssige Kräfte ohne zusätzlichen Rechenaufwand absorbieren und multifunktional sind. Insbesondere mit ihren kabellosen multimodalen Fortbewegungsfähigkeiten haben sich magnetische weiche Milliroboter (d. h. $\leq$1 cm) als potenzielle minimalinvasive medizinische Roboterplattformen herauskristallisiert, da sie Zugang zu engen und schwer zugänglichen Stellen im menschlichen Körper (z. B. distale Gefäßregionen) haben und medizinische Anwendungen wie die bedarfsgerechte Verabreichung von Medikamenten, Sensorik und Embolisierung an einem Zielort durchführen können. Für solche potenziellen biomedizinischen Anwendungen ist die Anpassungsfähigkeit der Robotersteuerung von entscheidender Bedeutung, um die Kontinuität des Betriebs zu gewährleisten, da die Bedingungen der Aufgabenumgebung dynamische Schwankungen aufweisen, die die Leistung des Roboters verändern können. Fertigungs-, material- und wechselwirkungsabhängige Variatnen sowie eine komplexe Kinematik mit praktisch unendlichen Freiheitsgraden, die sich aus der Natur ihrer weichen Materialstruktur ergeben, schränken jedoch die Anwendbarkeit herkömmlicher Modellierungs- und Steuerungsmethoden ein.

Das Hauptziel dieser Dissertation ist es, ein dateneffizientes adaptives multimodales Fortbewegungskonzept für die angestrebten Anwendungsszenarien magnetischer weicher Milliroboter zu entwickeln. Zu diesem Zweck wird ein probabilistischer Lernansatz eingeführt, der die

Bayes'sche Optimierung (BO) und Gauß'sche Prozesse (GPs) nutzt, um die Herausforderung der Regleranpassung anzugehen. Zunächst wird an drei verschiedenen Robotern, die in den gleichen Schritten hergestellt wurden, die Wirksamkeit der BO bei Fertigungsschwankungen gezeigt. Anschließend wird gezeigt, dass die Übertragung des von einem Roboter gelernten posterioren Mittelwerts als priorer Mittelwert auf die anderen Roboter und Testfälle die Lernleistung von BO verbessert, indem eine schnellere Ganganpassung durch die Durchführung erweiterter Tests auf Benchmark-Datensätzen erreicht wird. Anschließend wird die Methode der Regleranpassung, die den vorgeschlagenen Transfer-Learning-Ansatz verwendet, in verschiedenen Aufgabenräumen mit unterschiedlichen Oberflächenhaftungen, Oberflächenrauhigkeiten und mittleren Viskositätseigenschaften demonstriert. Um die Anpassungsleistung durch Einbeziehung multimodaler Fortbewegung weiter zu verbessern, wird in der dritten Studie die Sim-to-Real-Transferlernmethode entwickelt. In diesem Zusammenhang wird eine datengesteuerte Simulationsumgebung entworfen, deren Genauigkeit durch den Vergleich der simulierten Ergebnisse mit den physikalischen Experimenten nachgewiesen wird. Durch die Nutzung der simulierten Erfahrungen und des auf der BO basierenden Transfer-Lernens wird gezeigt, dass das Sim-zu-Real-Transfer-Lernen ein effizientes Lernen der Fortbewegung ermöglicht. Darüber hinaus wird die Angemessenheit der automatischen Bewegungsanpassung durch den auf der Kullback-Leibler-Divergenz basierenden Domänenerkennungsansatz für sich ändernde Umweltbedingungen gezeigt.

Als sekundäres Ziel wird eine neue Lokalisierungsmethode unter Verwendung der elektrischen Impedanztomographie eingeführt. Die Anwendbarkeit des vorgeschlagenen Ansatzes wird für stationäre und bewegte Fälle in Umgebungen mit und ohne Hindernisse demonstriert.

Mit diesen Beiträgen schlägt diese Arbeit ein domänenadaptives Lernsystem für die Fortbewegung vor, das die sanfte Fortbewegung von Millirobotern in die Lage versetzt, sich schnell und kontinuierlich an Umweltveränderungen anzupassen, während sie den Aktionsraum erkunden.

Acknowledgments

Since the beginning of my doctoral research in 2018, I have had the chance to meet and work with great people whose presence and collaboration have influenced my personal and professional life. I would like to thank them for their personal, philosophical, and scientific support in completing this dissertation.

First, I would like to thank Prof. Dr. Metin Sitti for giving me the opportunity to work with him in the Physical Intelligence Department at the Max Planck Institute for Intelligent Systems, Stuttgart. I thank him for all the supervision, support, guidance, and freedom he has provided to pursue my research and curiosity.

I would like to express my deepest gratitude to the Max Planck Society and Republic of Türkiye Ministry of National Education for supporting me and my research throughout my Ph.D.

I would also like to thank my co-examiners, Prof. Dr. Syn Schmitt and Prof. Dr. Georg Martius, for their contributions to my dissertation with their invaluable feedback.

I want to express my special thanks to Prof. Dr. Sebastian Trimpe for his scientific support, stimulating discussions, and influence on my research.

I would like to express my gratitude to Dr. Utku Culha. His mentorship, continuous support, and scientific and philosophical insights have been a constant source of motivation and enlightenment throughout this journey.

I also want to thank my colleagues and friends who were with me through my Ph.D. I am especially grateful to Patricia Martinez and Janina Schwartz for always being there and helping with professional and personal problems. I thank the undergraduate students I worked with, Melike Yıldırım, Ceyda Akyol, Burak Aydın, Aras Güngöre, and Mete Erdoğan, for contributing to my research. I thank the great friends I made during my Ph.D., Jack Saud, Alp. C. Karacakol, Cem B. Dayan, Jie Han, Ugur Bozuyuk, M. Birgul Akolpoglu, Zhiqiang Zheng, Gaurav Gardi, Y. Eren Suyolcu, Amir Aghakhani, Abdon Pena-Francesch, Isabella Bianchini, Ferro F. Ceylan, 'Le Petit Coq bar' family, and many others for all the good memories. I also thank Eylül Suadiye for his heartful friendship, all-time support, and coffee breaks. I especially thank M. Efe Tiryaki for his heartful friendship, all-time support, and endless questions about life, research, and technology.

I would like to extend my grateful appreciation to my parents, Hülya and Sadık Demir, and my brother, U. Eren Demir, for their endless love and support.

Last but most importantly, I would like to state my heartfelt thankfulness to my wife, Saadet F. Baltacı, for her all-time support, endless patience, continuous effort to motivate me, and for sharing all my happiness and grief since we met. Without her, none of these would be possible.

Science is the only true guide in life.

-Mustafa Kemal Atatürk-

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1. Magnetic Soft Millirobots

Soft robots made of highly flexible materials can withstand and undergo both elastic and plastic deformations[1]. This characteristic enables them to adapt to their surroundings, absorb the excess forces without additional computation effort, and achieve various functions while maintaining a reduced system complexity compared to their rigid counterparts[2–4]. Therefore, they have been utilized in various fields, such as object manipulation, wearables, and biomedical devices[5–9]. Meanwhile, small-scale (i.e., ≤1 cm) soft robots with wireless locomotion capability have become a promising approach for medical applications as they can access previously hard-to-reach environments non-invasively[10–12], including distal vascular regions and carry out medical applications, such as on-demand drug delivery, sensing, and embolization[13–16], in a target location.

Among the existing external actuation methods, such as heat[17], light[17,18], electric[19], and magnetic fields[17,20], magnetic actuation stands out due to its high precision, dexterity, speed, penetration depth, and biological safety features[21]. Moreover, magnetic actuation can be used with different imaging modalities, such as camera[22], electrical impedance tomography[23], x-ray[13], and ultrasound[24] without disturbing their functionality. Furthermore, researchers demonstrated that magnetic soft millirobots can perform various locomotion modes, such as walking, rolling, crawling, jumping, tumbling, swimming, and climbing[14,15,24]. While multiple robots with different magnetic profiles can generate these

gaits separately[14], a single robot design with a preprogrammed magnetic profile can also achieve multimodal locomotion under different periodic actuation signals[15–17,25], which extends the adaptation capabilities and targeted application areas of the magnetic soft millirobots.

Despite their exciting capabilities and potential medical use, soft robots face challenges arising from the nature of their soft materials, i.e., fabrication-, material-, and physical-interaction-dependent variations[26]. Moreover, the inherent high degrees of freedom, complex dynamics, contact mechanics, and dynamically changing task environments restrict the application of conventional modeling and control methods widely used in rigid robotic systems[12,27].

To address these challenges, one proposed solution involves simplifying the governing models for soft millirobots, such as dividing the robot into sub-parts and using constant curvature (CC) approximation with well-established beam theories to model the kinematics and dynamics of each sub-part[28,29]. However, CC approximation fails when significant loads are applied to the system, and its accuracy highly depends on the discretization of the robot[30]. Although the Cosserat rod model offers improved accuracy, it suffers from the computationally heavy model solutions[31].

Numerical approaches, on the other hand, such as finite-element methods (FEMs), voxel-based representations, and discrete differential geometries, improve the computation time by modeling the continuum robot structures using a chain of rigid elements connected with tunable spring-damper mechanisms at the expense of nonlinear dynamics

precision[32–34]. While these models and simulation tools perform well for large-scale soft robots[35], the lack of detailed and continuous body deformation sensing and computational complexity restrict their use in small-scale[36]. Alternatively, many soft millirobot applications rely on manual tuning to control the periodic actuation signals[15,16,24,37].

Besides manual methods, data-driven machine learning methods may also be used to control soft robots in the lack of analytical or numerical models[38]. One common approach is training neural network (NN) architecture with the data collected from physical experiments[39–41]. However, since they require a large training dataset, and the soft millirobots tend to have variations in their performance over prolonged usage due to material degradation, their use is also limited. Conversely, Bayesian optimization (BO) enables a data-efficient optimization, i.e., using a small number of physical experiments[42,43]. Furthermore, using BO together with Gaussian processes (GPs) improves the data efficiency and optimization performance by incorporating information as probabilistic priors[44], which is demonstrated on different-size robots to optimize their locomotion performance[45–48], including untethered soft millirobots[49].

Although data-efficient machine learning methods can be used for adaptation problems, their performance decreases as the search space enlarges since they rely on experimental data. As a solution, sim-to-real transfer learning-based control approaches have attracted significant attention in various complex real-world robotics problems. Control algorithms trained in high-fidelity rigid body simulation environments

now could beat human pilots in drone racing[50,51], plan and control complex legged locomotion[52–56], and perform complex object manipulation tasks[57–60]. On soft robotic systems, Hiller *et al.* proposed a FEM-based simulation environment to model complex material structures and multi-stimulus actuation[61]. The same simulation environment is further utilized to learn shape and control policy pairs in a given environment for a pneumatically actuated soft robot[62]. However, the sim-to-real transferability of multiple shapes and behaviors to a single robot remains an unsolved problem due to the gap between simulation and real-world performance.

As locomotion learning and adaptation to environmental changes are important challenges for magnetic soft millirobots, the proposed solutions' efficacy relies on the feedback systems' performance. Therefore, various technologies well established in biomedical imaging have been explored for tracking small-scale robots, such as magnetic-field-based techniques[63], ultrasound[64], optical techniques[65], and ionizing-radiation-based methods[66]. Despite their successful results, they face some limitations in terms of cost, scanning speed, penetration depth, and adverse health effects. As an alternative to these approaches, localization via electrical impedance variations in the environment has been demonstrated[67]. Electrical impedance tomography (EIT) exploits electrical stimulations and measurements through the surface of an object to noninvasively image the interior of a domain by estimating the internal distribution of the electrical properties. EIT has been used in medical imaging for various physiological phenomena, such as breathing[68], cardiac function[69], and brain activity[70]. While it does not cause any risk

to the patient, unlike the prolonged use of ionizing radiations, such as X-rays, it can image deep tissues, contrary to optical techniques, at high speed[71] and with low cost[72]. Moreover, since it is decoupled from the magnetic actuation, it can be used without damaging the utilized robots' magnetic profile, unlike magnetic resonance imaging (MRI).

1.2. Thesis Objectives

There is a growing interest in the field of soft millirobots and their application in biomedical targeting scenarios. However, there is a noticeable gap in addressing the control aspect of these robots, especially given their intended use in dynamic biomedical environments. Moreover, soft millirobots exhibit differences from one another due to fabrication variances and face performance lost over time due to material degradation.

The primary objective of this thesis is to establish a data-efficient adaptive locomotion framework for targeted application scenarios. This framework leverages BO with GP to enhance locomotion learning in a data-efficient way. Additionally, a new transfer learning approach is proposed for GP-BO to further improve data efficiency and optimization performance, demonstrating efficacy for different robot and task environments.

To expand the applicability of the proposed adaptive locomotion approach for more complex test scenarios and larger search spaces, a data-driven simulation environment is designed to model the soft millirobot's behavior, and a sim-to-real transfer learning approach is

employed to optimize the robot's performance in the real world. This sim-to-real transfer learning approach also enables the discovery of new locomotion modes.

Finally, a localization method based on electrical impedance tomography is proposed as an alternative to common biomedical imaging techniques to track the soft millirobot's position in closed and nontransparent environments.

These objectives collectively contribute to developing soft millirobots and their application in dynamic and diverse scenarios.

1.3. Outline

This thesis is divided into three chapters and four appendices. This first chapter lays down the problem description and motivation of this thesis, including the formulation of the research goal and objectives. The second chapter starts with defining the magnetic soft millirobot design and magnetic actuation and feedback systems used in this thesis. Then, it continues by explaining the utilized methods: Gaussian processes, Bayesian optimization, data-driven magnetic soft millirobot simulation, and electrical impedance tomography. At last, the third chapter concludes by summarizing the contributions and presenting a future outlook for soft millirobots.

Each appendix presents a publication that supports the thesis objectives. The first publication [26] showed the performance differences between reported results in [24] and three robots fabricated following the same procedures by running repeated experiments. Then, GP-BO was

demonstrated for the walking gait adaptation of different robots. Moreover, it was shown that transferring the posterior mean learned by one robot as the prior mean to other robots improved learning performance by achieving quicker gait adaptation.

The second publication [22] starts by testing the efficacy of the proposed transfer learning approach in the first publication. To this end, it compared four possible transfer learning methods with each other and the standard GP-BO by running augmented tests on five benchmark datasets consisting of 3750 physical experiments. The results showed that the proposed posterior mean transfer method outperformed the rest by achieving quicker gait adaptation and larger stride lengths for all five test cases. Afterward, the efficacy of the adaptive learning method was demonstrated in a wide range of task spaces with varying (1) surface adhesion, (2) surface roughness, and (3) the medium viscosity properties.

The third study [73] addresses the challenges of domain adaptive magnetic soft millirobot locomotion using a sim-to-real transfer learning method. To this end, a data-driven simulation environment was designed. Leveraging the simulated experience and GP-BO based transfer learning, it was demonstrated that sim-to-real transfer learning provided efficient locomotion learning. Moreover, the adequacy of the automated locomotion adaptation through the Kullback-Leibler divergence-based domain recognition approach was shown in changing environmental conditions.

The last study [23] introduces using electrical impedance tomography to track the position of the magnetic soft millirobot. After showing that the

EIT and the magnetic actuation system are decoupled, the localization of the robot was demonstrated for stationary and moving cases in environments with and without any obstacles.

1.4. List of Publications

This dissertation combines four first-author peer-reviewed papers listed below. The papers are attached to the appendix of this dissertation.

Publication 1 U. Culha[*], **S. O. Demir**[*], S. Trimpe, M. Sitti, *"Learning of Sub-optimal Gait Controllers for Magnetic Walking Soft Millirobots"*, **Robotics: Science and Systems** 2020,
doi: 10.15607/RSS.2020.XVI.070.
([*] Co-first authors)

Publication 2 **S. O. Demir**[*], U. Culha[*], A. C. Karacakol, A. Pena-Francesch, S. Trimpe, M. Sitti, *"Task Space Adaptations via The Learning of Gait Controllers of Magnetic Soft Millirobots"*, **International Journal of Robotics Research**, 2021,
doi: 10.1177/02783649211021869.
([*] Co-first authors)

Publication 3 **S. O. Demir**, M. E. Tiryaki, A. C. Karacakol, M. Sitti, *"Learning Soft Millirobot Multimodal Locomotion with Sim-to-Real Transfer"*, **Advanced Science**, 2024,
doi: 10.1002/advs.202308881

Publication 4 H. Daguerre[*], **S. O. Demir**[*], U. Culha, F. Marionnet, M. Gauthier, M. Sitti, A. Bolopion, *"A Localization Method for Untethered Small-Scale Robots Using Electrical Impedance Tomography"*, **IEEE/ASME Transactions on Mechatronics**, 2022, doi: 10.1109/TMECH.2022.3142924.
([*] Co-first authors)

Other publications, which I contributed and not part of this thesis, are listed below.

- Z. Zheng[*], J. Han[*], Q. Shi[*], **S. O. Demir**, W. Jiang, M. Sitti, *"Single-step precision programming of decoupled multiresponsive soft millirobots"*, **Proceedings of National Academy of Sciences**, 2024, doi: 10.1073/pnas.2320386121
- C. B. Dayan, D. Son, A. Aghakhani, Y. Wu, **S. O. Demir**, M. Sitti, *"Machine Learning-Based Shear Optimal Adhesive Microstructures with Experimental Validation"*, **Small**, 2023, doi: 10.1002/smll.202304437.
- U. Bozuyuk, E. Yildiz, M. Han, **S. O. Demir**, M. Sitti, *"Size-Dependent Locomotion Ability of Surface Microrollers on Physiologically Relevant Microtopographical Surfaces"*, **Small**, 2023, doi: 10.1002/smll.202303396.
- Z. Zheng, J. Han, **S. O. Demir**, H. Wang, W. Jiang, H. Liu, M. Sitti *"Electrodeposited Superhydrophilic-Superhydrophobic Composites for Untethered Multi-Stimuli-Responsive Soft Millirobots*", **Advanced Science**, 2023, 10, doi: 10.1002/advs.202302409.

- Z. Zheng, H. Wang, **S. O. Demir**, Q. Huang, T. Fukuda, M. Sitti, *"Programmable aniso-electrodeposited modular hydrogel microrobots"*, **Science Advances**, 2022, 8, doi: 10.1126/sciadv.ade6135.
- M. E. Tiryaki, **S. O. Demir**, M. Sitti, *"Deep Learning-based 3D Magnetic Microrobot Tracking using 2D MR Images"*, **IEEE Robotics and Automation Letters**, 2022, 7, 6982-6989, doi: 10.1109/LRA.2022.3179509
- U. Bozuyuk, E. Suadiye, A. Aghakhani, N. O. Dogan, J. Lazovic, M. E. Tiryaki, M. Schneider, A. C. Karacakol, **S. O. Demir**, G. Richter, M. Sitti, *"High-Performance Magnetic FePt (L10) Surface Microrollers Towards Medical Imaging-Guided Endovascular Delivery Applications"*, **Advanced Functional Materials**, 2022, 8, doi: 10.1002/adfm.202109741.
- F. N. P. Basualdo, G. Gardi, W. Wang, **S. O. Demir**, A. Bolopion, M. Gauthier, P. Lambert, M. Sitti, *"Control and Transport of Passive Particles Using Self-Organized Spinning Micro-Disks"*, **IEEE Robotics and Automation Letters**, 2022, 7, 2156-2161, doi: 10.1109/LRA.2022.3143306.
- G. Dogan, **S. O. Demir**, R. Gutzler, H. Gruhn, C. B. Dayan, U. T. Sanli, C. Silber, U. Culha, M. Sitti, G. Schütz, C. Grevent, K. Keskinbora, *"Bayesian Machine Learning for Efficient Minimization of Defecs in ALD Passivation Layers"*, **ACS Applied Materials & Interfaces**, 2021, 13, 54503-54515, doi: 10.1021/acsami.1c14586

2. Methods

2.1. Magnetic Soft Millirobot Design

In this study, the performance of the proposed learning approaches was evaluated using a magnetic soft millirobot design previously introduced by Hu *et al.*[24]. The chosen design (Figure 1) was selected for its ability to exhibit multimodal locomotion, such as rolling, walking, and crawling (Figure 2), by controlling the applied magnetic field acting on the robot (Figure 3). Moreover, this design has been employed in various prior studies, demonstrating its adaptability to diverse environmental conditions and tasks through manual control[15,16,74].

The robot was fabricated by mixing Ecoflex 00-10 (Smoot-On Inc.) and neodymium-iron-boron (NdFeB) magnetic particles with around $5\mu m$ diameter (MQP-15-7, Magnequench) with a 1:1 body mass ratio. After curing the pre-polymer mixture on a methyl methacrylate plate, the robots were cut using a high-resolution laser cutter (LPKF Protolaser U4) with dimensions of length $L = 3.7mm$, width $w = 1.5mm$, and height $h = 185\mu m$.To magnetize the robots, they were folded around a cylindrical

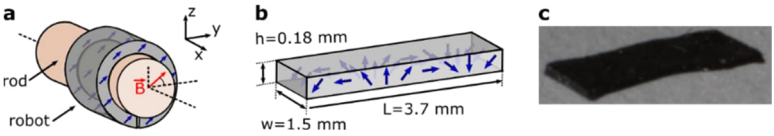


Figure 1 Fabrication of the magnetic soft millirobot. (a) The sheet-shaped soft millirobot was cut from the cured polymer mixture and magnetized inside a homogeneous magnetic field with a magnitude of $|B|$=1.8T (red arrow) after folding around a cylindrical rod. (b) The unfolded robot maintained a periodic magnetization profile (blue arrows) along its body. (c) Image of the fabricated and magnetized soft millirobot.

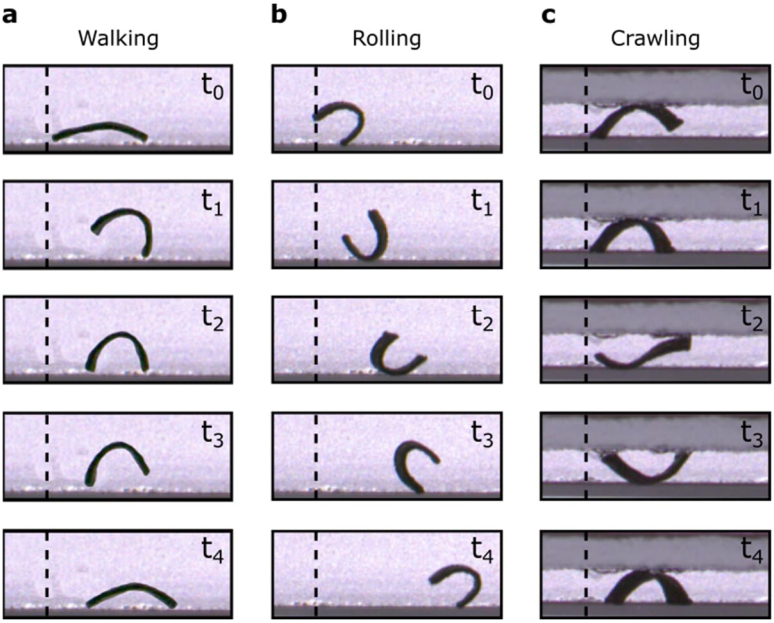


Figure 2 Locomotion modes of the magnetic soft millirobot. Physical performance of the magnetic soft millirobot during (a) walking, (b) rolling, and (c) crawling locomotion modes generated by manually designed actuation signals[24].

rod with a circumference equal to L and placed inside a uniform magnetic field with a magnitude of $1.8T$. The magnetic field was oriented at 45° counterclockwise from the y-axis. After separation from the rod, the magnetic particles retained their magnetization orientation, forming a circular profile along the robot body (Figure 1b). To actuate the robot, the homogeneous magnetic field in the environment and the resultant magnetic torque acting on the robot were continuously modulated (Figure 3).

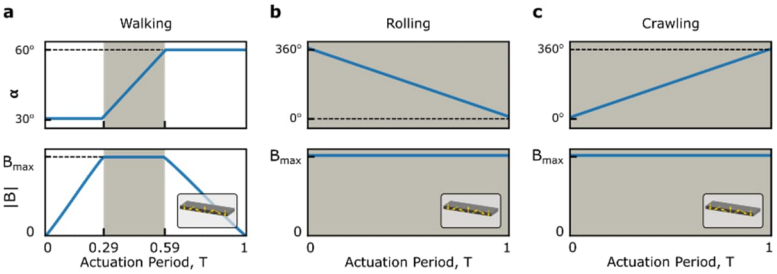


Figure 3 Manually designed magnetic actuation signals to generate predefined locomotion modes, (a) walking, (b) rolling, and (c) crawling[24]**.** The top and bottom rows show the direction and magnitude of the applied magnetic field in degrees and mT, respectively. The yellow arrows represent the magnetic profile of the robots.

To validate the repeatability of the reported results and the accuracy of the proposed mathematical model, three robots were fabricated and tested using the parameter sets reported by Hu *et al.*[24]. The results presented in Figure 4 revealed the following,

- Locomotion performance showed clear inconsistency due to the variations in the fabrication and environmental factors, even though the same materials, methods, controller parameters, and walking surfaces were used.
- Using a model-based controller was not possible due to non-monotonic behavior with increasing *f*, unlike the model prediction.

These results align with the objectives of this dissertation, which aim to develop a data-efficient controller learning system capable of robustly handling variations arising from the material properties, fabrication

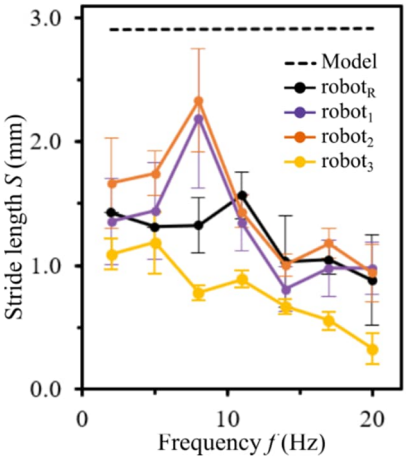


Figure 4 Predicted and observed walking performance of the magnetic soft millirobot in terms of stride length *S*. Previously reported model and experimental results (*robot*$_R$) were compared to the three replica robots (*robot*$_{1,2,3}$) tested with the same controller parameters[24]. Each data point for *robot*$_{1,2,3}$ represents the mean of 10 experiments and the error bars show the standard deviation.

processes, and the task environment of small-scale, medical-oriented, untethered soft robots.

2.2. Magnetic Actuation and Feedback System

Two distinct electromagnetic coil setups were designed for the studies reported in this dissertation (Figure 5). The first coil setup consisted of three orthogonal pairs of custom-made electromagnets (Figure 5a). It could generate a 3D uniform magnetic field within a 4×4×4cm^3 workspace with a maximum field strength of 15mT. The magnetic field was indirectly controlled by regulating the voltage values through six motor driver units (SyRen25) and an Arduino microcontroller. The

Arduino was responsible for processing analog readings from the current sensors, regulating the voltage outputs of the motor drivers, and handling all necessary communications. The robot's behavior was tracked by two high-speed cameras (Basler aCa2040-90uc) (Figure 5b). The first camera, running at 120 frames per second (fps), was placed orthogonally to the robot's motion plane (yz-plane) and was used to detect and evaluate the robot's locomotion mode. The second camera, operating at 60 fps, provided an isometric view of the test area and measured the robot's displacement following the perspective correction. At the end of every experiment, the system automatically commanded the robot to return to its starting position to minimize any human intervention and to avoid damage to the robot while handling.

The second coil setup was built following the Helmholtz coil design to enlarge the workspace with a uniform magnetic field to $13.1\times8.5\times4.5cm^3$

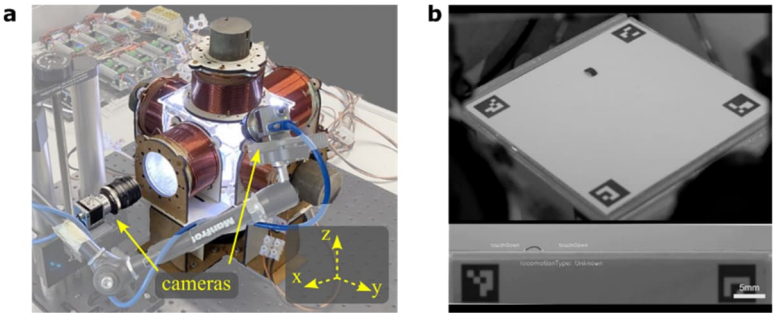


Figure 5 Electromagnetic coil setup. (a) Magnetic actuation setup with six electromagnetic coils could generate a 3D uniform magnetic field within a $4\times4\times4cm^3$ workspace with a maximum value of 15mT. (b) Two high-speed cameras with isometric and front views observed and evaluated the robot's motion in real-time.

(Figure 6a). Moreover, this design provided clear side, front, and top views while allowing new extra imaging modalities, such as C-arm and multispectral optoacoustic tomography device (MSOT 512-element transducer, iThera Medical), for future studies. Unlike the first setup, the magnetic field was directly controlled by modulating the currents on the coils through six motor driver units (Maxon ESCON 70/10). To enhance the achievable actuation frequency, an FPGA module (NI PXIe-7847R) was used as the interface to control the motor drivers, receive current readings, and communicate with the master PC. The robot's motion was tracked with two high-speed cameras (Basler aCa2040-90uc) running at 120 fps (Figure 6d). The first camera with a top view measured the robot's displacement, while the second one identified the robot's locomotion mode. The top view of the test area, captured without obstruction or perspective correction, expanded the range of experimental scenarios. Moreover, to enlarge the workspace having a homogeneous 3D magnetic field and to test the robot's performance for longer runs without reaching the workspace's limits, we integrated a motorized linear stage with 150 mm stroke (Thorlabs LTS150C) to the y-axis of the experimental setup. The linear stage continuously adjusted its position based on the displacement information received from the imaging system to keep the robot centered in the magnetic field.

The entire software architecture, such as learning algorithms, actuation signal generation, and image processing codes, ran on a master PC. Communication between different elements of the experimental system (e.g., image capture and electric current control) was executed on ROS

architecture, which allowed the system to be scalable for further extensions.

Moreover, the automated experimental platforms minimized the need for human intervention during physical experiments, reducing disturbances introduced by human interactions with the soft millirobot and the test environment. This approach maintained the repeatability of the physical learning experiments while avoiding human-induced alterations to the soft millirobot's behavior.

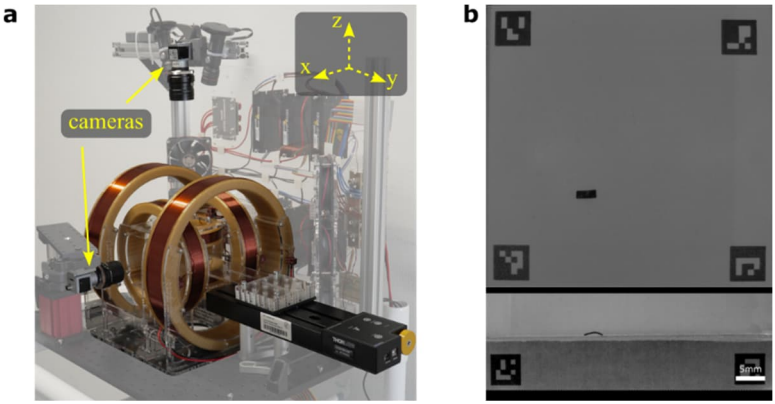


Figure 6 Electromagnetic (Helmholtz) coil setup. (a) Helmholtz coil setup with six electromagnetic coils could generate a 3D uniform magnetic field within a 13.1×8.5×4.5cm^3 workspace with a maximum value of 12 mT. (b) Two high-speed cameras with top and front views observed and evaluated the robot's motion in real-time. The linear stage attached to the xy-plane enlarged the test area with the homogeneous magnetic field on the y-axis.

2.3. Gaussian Processes

The performance of the magnetic soft millirobot, even when the actuation signal is defined by a single variable (e.g. Figure 4), cannot be modeled or predicted accurately due to its inherent complexity. Furthermore, soft millirobots tend to exhibit variations in performance over time, primarily due to material degradation. These factors limit the size of the physical dataset that can be generated for analysis. To overcome the sparsity in the dataset, to include uncertainties coming from the physical experiments, and to make probabilistic predictions, GP was used to map the input parameter set θ to scalar reward values $R(\theta)$, i.e., the stride length of the magnetic soft millirobot in the event of locomotion optimization.

$$R(\theta) \sim GP\big(\mu(\theta), k(\theta, \theta')\big), \tag{1}$$

where $R(\theta)$ is the reward function mapping the input parameter θ to scalar reward values, $\mu(\theta)$ denotes the prior mean for the input parameter θ, and $k(\theta, \theta')$ is the kernel function defining the covariance between $R(\theta)$ and $R(\theta')$ for $\theta, \theta' \in \Theta$, where Θ defines the complete search space of input parameters. For the cases where the performance metric can only be measured with noise, then the observed reward value $\tilde{R}(\theta)$ is defined as follows,

$$\tilde{R}(\theta) = R(\theta) + n, \tag{2}$$

where n stands for the zero-mean Gaussian noise with variance σ_n^2 for each measurement. At each iteration of the optimization run, the GP model is updated with $\tilde{R}(\theta)$ as shown in Figure 7. Then, the expected

value of $R(\theta)$ can be predicted with variance for any given θ using the test data $D = \{\theta_i, \tilde{R}(\theta_i)\}_{i=1}^{N}$, where N is the size of the dataset D, as follows:

$$\mu_{post}(\theta) = \mu(\theta) + k^T(\theta)K^{-1}y, \tag{3}$$

$$\sigma_{post}^2(\theta) = k(\theta,\theta) - k^T(\theta)K^{-1}k(\theta), \tag{4}$$

$$R_{post}(\theta)|D \sim \mathcal{N}\left(\mu_{post}(\theta), \sigma_{post}^2(\theta)\right), \tag{5}$$

where $k(\theta)$, $y \in \mathbb{R}^N$, and $K \in \mathbb{R}^{N\times N}$ denote $[k(\theta)]_i = k(\theta,\theta_i)$, $y_i = \tilde{R}(\theta_i) - \mu(\theta_i)$, and $K_{i,j} = k(\theta_i,\theta_j) + \delta_{i,j}\sigma_n^2$ with Kronecker delta $\delta_{i,j}$, respectively.

Since the kernel encodes the prior assumptions on the approximated reward function, such as the smoothness and periodicity of the signal, its definition and hyperparameter settings directly affect the prediction accuracy of the GP. Based on the results presented by Von Rohr *et al.*[49], squared exponential was selected as the kernel function, which is defined for multi-dimensional cases as follows:

$$k_{SE}(\theta,\theta') = \sigma_f^2\, exp\left(-\sum_{d=1}^{d_c}\frac{(\theta_d-\theta_d')^2}{2l_{c,d}^2}\right), \tag{6}$$

where $l_c \in \mathbb{R}_c^d$ is the length scale defining the rate of change in the approximated function for each parameter space dimension and σ_f^2 is the signal variance describing the uncertainty in the predictions for unobserved parameter sets[75].

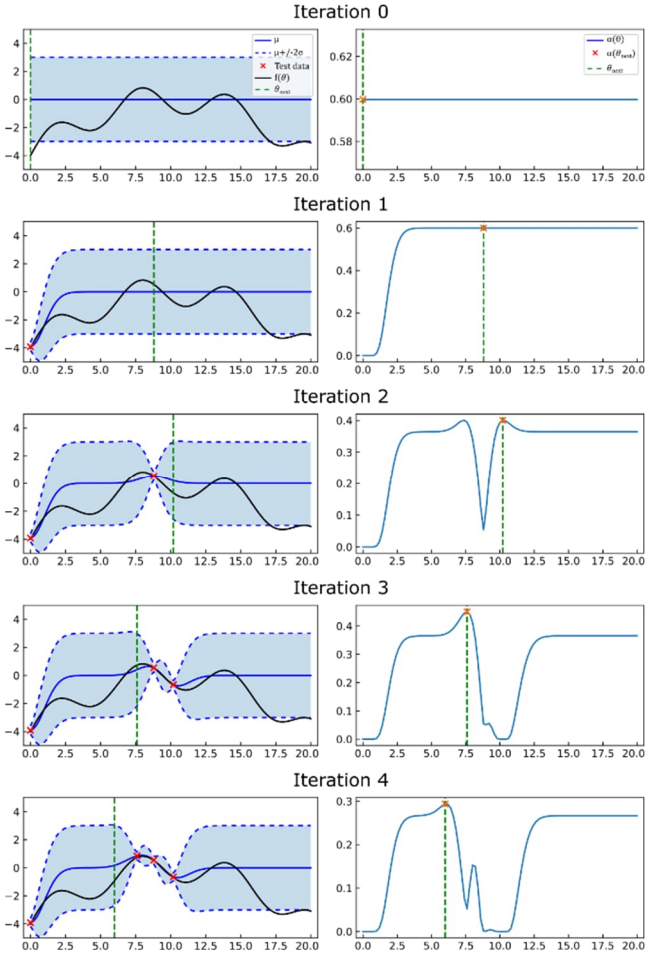


Figure 7 Sample learning run to maximize a function $f(\theta)$ using Bayesian optimization (BO) with Gaussian processes (GPs). The left column shows the posterior GP model as new data points are observed. The right column shows the value of the acquisition function at the end of each iteration. The next parameter set (θ_{next}) chosen by BO to be used in the next iteration is shown by orange mark on the acquisition function plots.

As illustrated in Figure 8, hyperparameter selection has a vital role in the prediction accuracy of GP. One crucial hyperparameter is the length scale. While long length scales are used to model slowly-varying functions, short length scales are used to model quickly-varying functions, enabling the GP to capture rapid variations within the data (Figure 8a). Another key hyperparameter is the signal variance (σ_f^2), which influences the exploration-exploitation balance during the optimization run. A higher value of σ_f^2 encourages exploration by increasing the prediction uncertainty, which can be useful for discovering optimal solutions. Conversely, a lower value promotes exploitation (Figure 8b). Noise in the observation (σ_n^2) is another important hyperparameter controlling the impact of a new data point on the posterior mean function (Figure 8c).

In addition to the kernel selection and hyperparameter definition, the prior mean ($\mu(\theta)$) has an important effect on the performance of GP. While the prior mean is often assumed to be constant and set to zero (i.e., $\mu = 0$) in the context of maximization problems[76], its selection significantly impacts optimizing performance by controlling the exploration-exploitation balance. To illustrate, a constant zero mean may cause the optimizer to get stuck on local optima for a maximization problem with large reward values by making it believe it has already found the best-performing result. On the other hand, a constant non-zero mean may lead the optimizer to keep exploring new regions in the search space even after finding the global optimum by making it assume that it evaluated poor-performing parameter sets.

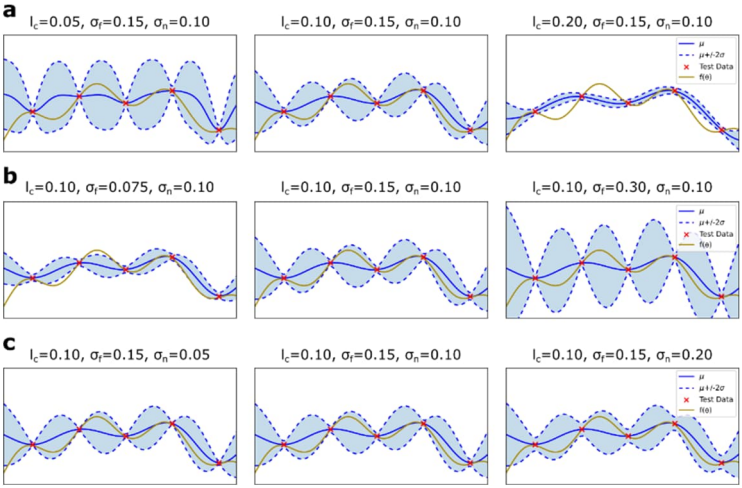


Figure 8 Effect of the hyperparameters on the prediction accuracy of Gaussian processes, (a) length scale (l_c), (b) signal variance (σ_f), and (c) observation noise (σ_n).

Alternative to the constant prior mean approach (Figure 9a,c), the transfer of information from previous learning runs is proposed in this thesis. In this approach, the prior mean was set equal to the posterior mean of a previously trained GP model (Figure 9b,d). It enabled the transfer of the topology of the target function across different test scenarios, as long as variations between the robots and the environments did not significantly alter the function shape.

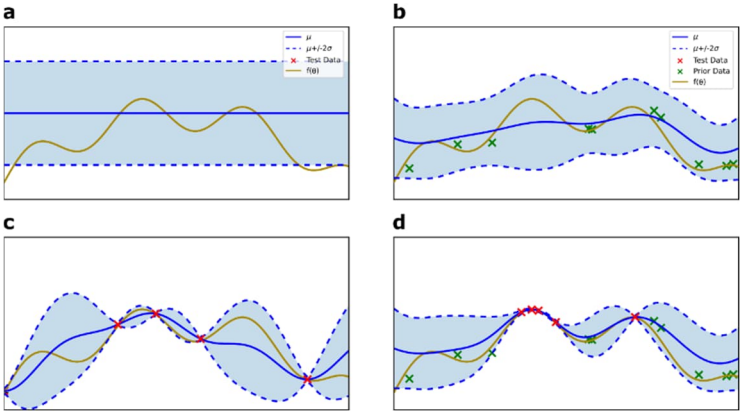


Figure 9 Gaussian process (GP) model representation of a function $f(\theta)$. (a), (c) show the GP model prediction at the beginning and after completing 5 iterations of learning run found by setting the prior mean function to constant mean. (b), (d) show the GP model prediction at the beginning and after completing 5 iterations of learning run found by setting the prior mean function to previously learned posterior mean defined by the training data points shown in green markers.

2.4. Bayesian Optimization

BO was chosen as the optimization method for the studies conducted in the scope of this thesis because it allows the maximization of a performance function, i.e., the stride length of the magnetic soft millirobot in the event of locomotion optimization, using a small number of physical experiments[42,43]. At each iteration of the learning run, BO selects the next parameter set (θ_{next}) maximizing the acquisition function ($\alpha_{acq}(\theta)$) as follows:

$$\theta_{next} = \underset{\theta \in \Theta}{\operatorname{argmax}}\, \alpha_{acq}(\theta), \tag{7}$$

The acquisition function was selected as the expected improvement (EI) due to its better performance compared to its alternatives in similar test scenarios[49]. Additionally, the exploration-exploitation balance capability made EI a particularly effective choice. EI is defined as,

$$\alpha_{acq}(\theta) = \mathbb{E}\left[\max\left(0, \left(R(\theta) - \tilde{R}(\theta^*)\right)\right)\right], \quad (8)$$

where $\tilde{R}(\theta^*)$ is the highest observed reward function value[77]. The analytical solution for Eq. (8) is given as

$$\alpha_{acq}(\theta) = \left(\mu(\theta) - \tilde{R}(\theta^*) - \xi\right)\Phi(Z) + \sigma(\theta)\phi(Z), \quad (9)$$

where Φ and ϕ are the Gaussian cumulative density and probability density functions, respectively[78]. Z is defined as $Z = Z(\theta) = (\mu(\theta) - \tilde{R}(\theta^*) - \xi)/\sigma(\theta)$, with $\mu(\theta)$ and $\sigma(\theta)$ are calculated by Eq.s (3) and (4). The two terms in Eq. (9) represent the exploitation and exploration weights of the BO, respectively. Their balance is controlled by setting the hyperparameter ξ. As ξ gets higher, BO tends to choose the parameter set in unobserved regions of the search space. BO focuses more on exploitation by testing parameters close to already explored regions as ξ gets lower.

A sample run of GP-BO maximizing a function $f(\theta)$ is given in Figure 7 to show the posterior GP models and calculated acquisition function values at each iteration. The pseudocode showing the detailed flow of the learning run using GP-BO is given in Algorithm 1.

Algorithm 1 Pseudo code for the Bayesian optimization with Gaussian processes learning run.

Inputs: Search space, Θ
Prior mean function, $\mu(\theta)$
Output: Best performing parameter set, θ^*

$resultArray \leftarrow$ Initialize an empty array to store $R(\theta)$ and θ
$GP \leftarrow$ Initialize GP with $\mu(\theta)$
for $iteration < maxIteration$ **do**
$\mu_{post}, \sigma_{post} = GP(\theta)$
$\alpha_{acq} = EI(\mu_{post}, \sigma_{post})$
$\theta_{next} = \underset{\theta \in \Theta}{\operatorname{argmax}}\, \alpha_{acq}(\theta)$
$\tilde{R}(\theta_{next}) \leftarrow$ Test θ_{next} by running experiment
$GP \leftarrow$ Update GP with $\tilde{R}(\theta_{next})$
$resultArray \leftarrow$ Add $[\theta_{next}, \tilde{R}]$ to $resultArray$
end
$\theta^* = \underset{\theta \in \Theta}{\operatorname{argmax}}(resultArray)$
return θ^*

2.5. Magnetic Soft Millirobot Simulation

Even though automated physical experiments can generate the required data to optimize the robot's locomotion performance in different conditions, as the search space enlarges and environmental topologies substantially vary, running physical experiments becomes impractical [26,79]. Therefore, a data-driven simulation environment was designed based on the open-source software *Voxelyze*, as it could efficiently simulate heterogeneous 3D rigid and soft bodies [80].

The dynamic behavior of the magnetic and nonmagnetic rigid and soft materials is modeled as a mass-spring-damper system (Figure 10). Multi-body interaction is modeled by adapting the contact mechanics of *Voxelyze*, defining the interaction between the robot and the floor. To model the magnetic actuation, first, the magnetic torque acting on each voxel due to the external homogeneous field is calculated and then added into dynamic equations. Magnetic torque τ_t acting on a voxel of the robot at time step t is calculated as follows.

$$\tau_t = m_t \times B_t, \tag{10}$$

$$m_t = M_r d_v^3 R_t, \tag{11}$$

where B_t, M_r and d_v denote the homogeneous magnetic field at time t, magnetic remanence, and voxel size, respectively. R_t is the rotational matrix defining the magnetic orientation of the voxel at time step t (Figure 10d).

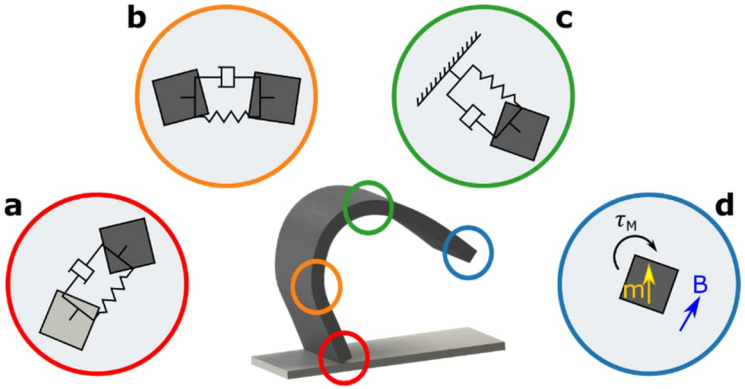


Figure 10 Simulation parameters of the data-driven magnetic soft millirobot simulation. Schematic representation of the magnetic soft millirobot simulation by mass-spring-damper model between (a) colliding voxels, (b) connected voxels, (c) voxel and the surrounding, and (d) magnetic torque acting on a voxel due to external magnetic field B.

For all the simulations, the voxel size (d_v) was set to $185\mu m$. Density (ρ), Young's modulus (E), and magnetic remanence (M_r) values for the magnetic soft millirobot were taken from [24] as $1.86 g/cm^3$, $8.45 kPa$, and $62 kA/m$, respectively. The Poisson's ratio is assumed to be 0.49.

Simulated results for walking, rolling, and crawling locomotions of the magnetic soft millirobot actuated by the same control signals in Figure 2 are shown in Figure 11.

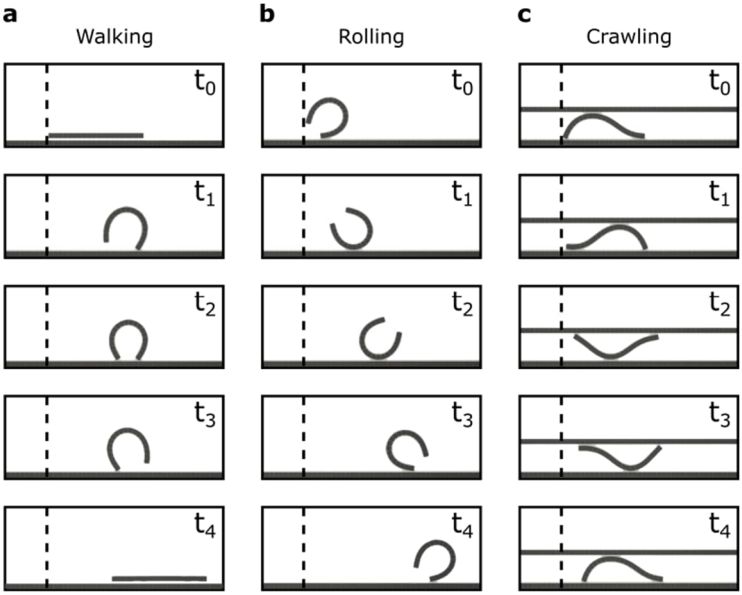


Figure 11 Simulated locomotion modes of the magnetic soft millirobot. Simulated performance of the magnetic soft millirobot during (a) walking, (b) rolling, and (c) crawling locomotion modes generated by manually designed actuation signals[24].

2.6. Robot Tracking by Electrical Impedance Tomography

EIT is proposed as an alternative solution to the camera-based localization of the magnetic soft millirobot, as it can be used in nontransparent working environments.

EIT estimates the internal structure of the measured object or the environment by processing the emitted and measured electrical signals, i.e., voltages or currents, applied through multiple electrodes attached to the surface. The forward problem focuses on modeling these electrical measurements at any position on the domain's boundary using the internal conductivity distribution σ and the stimulating current injection. It can be solved by numerical methods, such as FEM. Conversely, the inverse problem aims to reconstruct the conductivity distribution via the electrical measurements and needs to be solved to detect the robot's position (Figure 12). However, the inverse problem is particularly complex due to nonuniqueness and the instability of the solution[81]. Thus, the solver developed by Adler *et al.*[81] employing a spatiotemporal resolution approach is used with the open-source software EIDORS[82].

To stimulate and collect the necessary electrical measurements from the environment, a circular test area surrounded by eight evenly distributed electrodes was designed (Figure 13b). Although the number of electrodes could vary, it was determined by the maximum number of analog input channels available in the data acquisition system (PXIe-7847R, National Instruments). The electrodes were excited and measured following a full-scan stimulation pattern involving a single current source and a sink

(Figure 13e). As a result, 28 different stimulation combination was generated in the eight-electrode EIT system, and they were given to the inverse problem solver as input data to reconstruct the conductivity map of the environment.

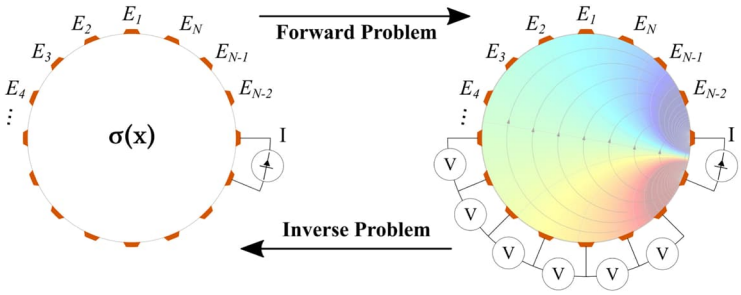


Figure 12 Forward and inverse problem definitions for EIT. An electrical stimulation was applied through the boundary of the measured domain, and the resulting voltages were measured. The forward problem provides the measurements from the knowledge of the medium properties, while the inverse problem estimates the medium properties from the measurements.

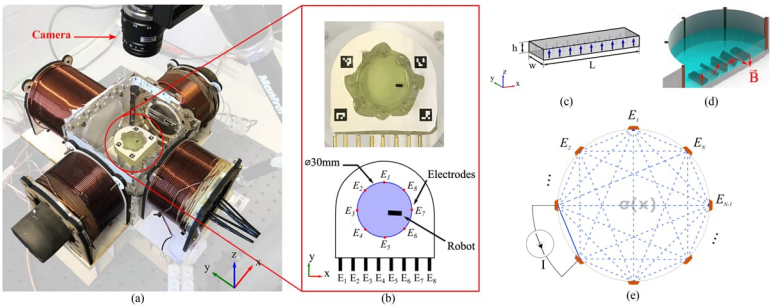


Figure 13 Experimental setup and the electrical stimulation pattern used for the EIT experiments. (a) Image of the magnetic actuation system (Figure 5a) with the test area surrounded by the eight electrodes of the EIT system and a high-speed camera to generate the ground truth data. (b) Top-view image and drawing of the test area showing the electrode positions and a sample robot in the test fluid. (c) Uniform magnetization profile of the magnetic small-scale robot along its body (blue arrows). (d) Sample movement of the magnetic small-scale robot by applying a rotating magnetic field B shown by red arrows. (e) Illustration of the full-scan EIT stimulation scheme. All electrode pairs were successively selected as driving electrodes.

3. Conclusion and Outlook

3.1. Conclusion

This dissertation represents a comprehensive exploration of data-efficient adaptive locomotion for soft millirobots demonstrated on a sheet-shaped elastomeric magnetic soft millirobot with the reported publications. The first paper in Appendix A, *"Learning of Sub-optimal Gait Controllers for Magnetic Walking Soft Millirobots"*, showed the performance variations in magnetic soft millirobots by repeated tests. It proposed using GP-BO to adapt to the variabilities inherent in soft robots, even in the absence of model-based control.

Next, the second paper in Appendix B, *"Task Space Adaptations via the Learning of Magnetic Soft Millirobots"*, demonstrated the controller parameters learning for the walking gait of the magnetic soft millirobot by GP-BO. To this end, a benchmark dataset consisting of 3750 experimental results was generated through exhaustive grid search, and the effectiveness of the prior mean transfer in GP-BO was shown by running 10^4 augmented tests. The results highlighted that transferring prior mean information improved the learning performance by reducing the required experiments to find parameter sets generating walking locomotion. The performance of the proposed transfer learning approach was tested further in various task spaces with changing surface adhesion, surface roughness, and medium viscosity. Apart from the proposed learning approach, the electromagnetic coil setup built in the scope of this study has allowed the running of automated experiments requiring

minimized human intervention, making running repeated experiments with lower performance variations possible.

The third paper in Appendix C, *"Learning Soft Millirobot Multimodal Locomotion with Sim-to-Real Transfer"*, introduced sim-to-real transfer learning to improve the optimization performance in larger search spaces. In this regard, a data-driven soft millirobot simulation was designed. Using the previously presented transfer learning approach to bridge the gap between simulation and real-world performance, the efficacy of *a priori* knowledge generation in the simulation and performing sim-to-real transfer learning was demonstrated in various environments and robots. Moreover, an autonomous environment identification method was proposed by matching experimental results to simulated data. Its performance was demonstrated with the autonomous locomotion adaptation of the robot in previously not encountered conditions.

The fourth paper in Appendix D, *"A Localization Method for Untethered Small-Scale Robots Using Electrical Impedance Tomography"*, demonstrated an alternative solution to track the position of the magnetic soft millirobot using EIT, which was shown to operate without interfering with the magnetic actuation system.

To summarize, this thesis presents a holistic approach to advancing the field of soft millirobots, encompassing adaptive control, simulation, and alternative tracking methods. The findings and proposed methodologies contribute to the data-efficient, adaptive, and versatile locomotion capabilities of soft millirobots.

3.2. Future Work

The research presented in this dissertation provides a solid foundation for future explorations and advancements in the field of soft millirobots. While ongoing research into small-scale fabrication techniques may address reproducibility concerns[83–85], the inherent variability and the performance decrease over prolonged use due to material degradations in soft millirobots underscore the continued need for adaptive control.

The learning method and simulation presented in this thesis are demonstrated on a specific sheet-shaped magnetic soft millirobot, but they can be applied to different soft robot shapes and materials[25] and various tasks in more complex environments, such as climbing, path following, and velocity control in 3D confined spaces[15]. The adaptability of these approaches opens up exciting possibilities toward the real-world applications of these robots, especially with the sim-to-real transfer learning method. Moreover, the developed simulation environment enables the exploration of alternative control and adaptation approaches. Furthermore, the successful application of sim-to-real transfer learning introduces the potential to plan robotic operations beforehand, such as in medical procedures.

The effectiveness of sim-to-real transfer learning and adaptive control methods based on BO has been validated in various test cases; however, these methods may face challenges in dynamic environments with momentary changes owing to the episodic nature of BO. One potential solution would be using continuous control algorithms, such as deep reinforcement learning (DRL)[55,86,87]. However, these algorithms,

especially the neural network-based ones, require a larger training datasets. Therefore, as the next step, the simulation environment will be updated to integrate adhesion mechanics and fluidic interactions. This will enable modeling the robot's dynamics in scenarios involving sticky surfaces, such as biological tissues covered by mucus, and diverse fluid environments, such as blood. Moreover, simulation speed will be improved to explore planning and new continuous control algorithms. This can be achieved by limiting the update in the simulation to a local frame where the robot is located, thereby omitting unnecessary calculations in the rest of the simulated environment. The possibility of replacing the simulation environment with a deep neural network model will also be explored[88].

In addition to adaptive control, the GP-BO-based learning method can be further used to design the robots, i.e., their morphological and magnetic properties, and to learn actuation signals for a given task without requiring physical experiments[55,87,89]. Furthermore, the domain identification methods introduced in this research can be used for enhancing localization accuracy and mapping complex environments.

As the field of soft millirobots continues to evolve, these future research directions will play a pivotal role in enhancing their adaptability and expanding their applicability across various domains.

Bibliography

1. Majidi, C. Soft Robotics: A Perspective - Current Trends and Prospects for the Future. *Soft Robot* 1, 5–11 (2014).

2. Polygerinos, P. *et al.* Soft Robotics: Review of Fluid-Driven Intrinsically Soft Devices; Manufacturing, Sensing, Control, and Applications in Human-Robot Interaction. *Adv Eng Mater* 19, 1700016 (2017).

3. Haddadin, S., De Luca, A. & Albu-Schäffer, A. Robot collisions: A survey on detection, isolation, and identification. *IEEE Transactions on Robotics* 33, 1292–1312 (2017).

4. Hughes, J. *et al.* Soft manipulators and grippers: A review. *Frontiers Robotics AI* 3, 1–12 (2016).

5. Gu, G. *et al.* A soft neuroprosthetic hand providing simultaneous myoelectric control and tactile feedback. *Nature Biomedical Engineering 2021 7:4* 7, 589–598 (2021).

6. Yin, J. *et al.* Wearable Soft Technologies for Haptic Sensing and Feedback. *Adv Funct Mater* 31, 2007428 (2021).

7. Li, M., Pal, A., Aghakhani, A., Pena-Francesch, A. & Sitti, M. Soft actuators for real-world applications. *Nature Reviews Materials 2021 7:3* 7, 235–249 (2021).

8. Reeder, J. T. *et al.* Soft, bioresorbable coolers for reversible conduction block of peripheral nerves. *Science (1979)* 377, 109–115 (2022).

9. Cianchetti, M., Laschi, C., Menciassi, A. & Dario, P. Biomedical applications of soft robotics. *Nature Reviews Materials 2018 3:6* 3, 143–153 (2018).

10. Hines, L., Petersen, K., Lum, G. Z. & Sitti, M. Soft Actuators for Small-Scale Robotics. *Advanced Materials* 29, (2017).

11. Wang, B., Kostarelos, K., Nelson, B. J. & Zhang, L. Trends in Micro-/Nanorobotics: Materials Development, Actuation, Localization, and System Integration for Biomedical Applications. *Advanced Materials* 33, (2021).

12. Sitti, M. Miniature soft robots — road to the clinic. *Nature Reviews Materials 2018 3:6* 3, 74–75 (2018).

13. Wang, T. *et al.* Adaptive wireless millirobotic locomotion into distal vasculature. *Nat Commun* 13, 1–17 (2022).

14. Soon, R. H. *et al.* Pangolin-inspired untethered magnetic robot for on-demand biomedical heating applications. *Nat Commun* 14, 1–15 (2023).

15. Wu, Y., Dong, X., Kim, J. K., Wang, C. & Sitti, M. Wireless soft millirobots for climbing three-dimensional surfaces in confined spaces. *Sci Adv* 8, (2022).

16. Ren, Z. *et al.* Soft-bodied adaptive multimodal locomotion strategies in fluid-filled confined spaces. *Sci Adv* 7, (2021).

17. Zheng, Z. *et al.* Electrodeposited Superhydrophilic-Superhydrophobic Composites for Untethered Multi-Stimuli-Responsive Soft Millirobots. *Advanced Science* 10, (2023).

18. Pilz Da Cunha, M., Debije, M. G. & Schenning, A. P. H. J. Bioinspired light-driven soft robots based on liquid crystal polymers. *Chem Soc Rev* 49, 6568–6578 (2020).

19. Li, G. *et al.* Self-powered soft robot in the Mariana Trench. *Nature* 591, 66–71 (2021).

20. Son, D., Gilbert, H. & Sitti, M. Magnetically Actuated Soft Capsule Endoscope for Fine-Needle Biopsy. *Soft Robot* 7, 10–21 (2020).

21. Sitti, M. & Wiersma, D. S. Pros and Cons: Magnetic versus Optical Microrobots. *Advanced Materials* 32, (2020).

22. Demir, S. O. *et al.* Task space adaptation via the learning of gait controllers of magnetic soft millirobots. *International Journal of Robotics Research* 40, 1331–1351 (2021).

23. Daguerre, H. *et al.* A Localization Method for Untethered Small-Scale Robots Using Electrical Impedance Tomography. *IEEE/ASME Transactions on Mechatronics* 27, 3506–3516 (2022).

24. Hu, W., Lum, G. Z., Mastrangeli, M. & Sitti, M. Small-scale soft-bodied robot with multimodal locomotion. *Nature* 554, 81–85 (2018).

25. Zheng, Z. *et al.* Programmable aniso-electrodeposited modular hydrogel microrobots. *Sci Adv* 8, (2022).

26. Culha, U., Demir, S. O., Trimpe, S. & Sitti, M. Learning of Sub-optimal Gait Controllers for Magnetic Walking Soft Millirobots. in *Robotics: Science and Systems* (2020). doi:10.15607/RSS.2020.XVI.070.

27. Rus, D. & Tolley, M. T. Design, fabrication and control of soft robots. *Nature* 521, 467–475 (2015).

28. Della Santina, C., Bicchi, A. & Rus, D. On an Improved State Parametrization for Soft Robots with Piecewise Constant Curvature and Its Use in Model Based Control. *IEEE Robot Autom Lett* 5, 1001–1008 (2020).

29. Webster, R. J. & Jones, B. A. Design and kinematic modeling of constant curvature continuum robots: A review. *International Journal of Robotics Research* 29, 1661–1683 (2010).

30. Cianchetti, M., Renda, F., Giorelli, M., Laschi, C. & Calisti, M. Dynamic Model of a Multibending Soft Robot Arm Driven by Cables. *IEEE Transactions on Robotics* 30, 1109–1122 (2014).

31. Renda, F., Boyer, F., Dias, J. & Seneviratne, L. Discrete Cosserat Approach for Multisection Soft Manipulator Dynamics. *IEEE Transactions on Robotics* 34, 1518–1533 (2018).

32. Chenevier, J., González, D., Aguado, J. V., Chinesta, F. & Cueto, E. Reduced-order modeling of soft robots. *PLoS One* 13, (2018).

33. Goury, O. & Duriez, C. Fast, Generic, and Reliable Control and Simulation of Soft Robots Using Model Order Reduction. *IEEE Transactions on Robotics* 34, 1565–1576 (2018).

34. Huang, W., Huang, X., Majidi, C. & Jawed, M. K. Dynamic simulation of articulated soft robots. *Nature Communications 2020 11:1* 11, 1–9 (2020).

35. George Thuruthel, T., Ansari, Y., Falotico, E. & Laschi, C. Control Strategies for Soft Robotic Manipulators: A Survey. *Soft Robot* **5**, 149–163 (2018).

36. Rich, S. I., Wood, R. J. & Majidi, C. Untethered soft robotics. *Nature Electronics 2018 1:2* 1, 102–112 (2018).

37. Kim, Y., Parada, G. A., Liu, S. & Zhao, X. *Ferromagnetic Soft Continuum Robots*. *Sci. Robot* vol. 4 https://www.science.org (2019).

38. Chin, K., Hellebrekers, T. & Majidi, C. Machine Learning for Soft Robotic Sensing and Control. *Advanced Intelligent Systems* 2, 1900171 (2020).

39. Hyatt, P., Wingate, D. & Killpack, M. D. Model-based control of soft actuators using learned non-linear discrete-time models. *Frontiers Robotics AI* **6**, (2019).

40. George Thuruthel, T., Shih, B., Laschi, C. & Thomas Tolley, M. *Soft Robot Perception Using Embedded Soft Sensors and Recurrent Neural Networks*. *Sci. Robot* vol. 4 https://www.science.org (2019).

41. Bern, J. M., Schnider, Y., Banzet, P., Kumar, N. & Coros, S. Soft Robot Control with a Learned Differentiable Model. *2020 3rd IEEE International Conference on Soft Robotics, RoboSoft 2020* 417–423 (2020) doi:10.1109/ROBOSOFT48309.2020.9116011.

42. Ghahramani, Z. *Probabilistic Machine Learning and Artificial Intelligence*. *Nature on* vol. 27 (2015).

43. Shahriari, B., Swersky, K., Wang, Z., Adams, R. P. & De Freitas, N. Taking the human out of the loop: A review of Bayesian optimization. *Proceedings of the IEEE* vol. 104 148–175 Preprint at https://doi.org/10.1109/JPROC.2015.2494218 (2016).

44. Rasmussen, C. Edward. & Williams, C. K. I. *Gaussian Processes for Machine Learning*. (MIT Press, 2006).

45. Calandra, R., Seyfarth, A., Peters, J. & Deisenroth, M. P. Bayesian optimization for learning gaits under uncertainty: An experimental comparison on a dynamic bipedal walker. *Ann Math Artif Intell* 76, 5–23 (2016).

46. Liao, T. *et al.* Data-efficient learning of morphology and controller for a microrobot. *Proc IEEE Int Conf Robot Autom* 2019-May, 2488–2494 (2019).

47. Marco, A., Hennig, P., Bohg, J., Schaal, S. & Trimpe, S. Automatic LQR tuning based on Gaussian process global

optimization. *Proc IEEE Int Conf Robot Autom* 2016-June, 270–277 (2016).

48. Yang, B. *et al.* Learning flexible and reusable locomotion primitives for a microrobot. *IEEE Robot Autom Lett* 3, 1904–1911 (2018).

49. Von Rohr, A., Trimpe, S., Marco, A., Fischer, P. & Palagi, S. Gait Learning for Soft Microrobots Controlled by Light Fields. in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* 6199–6206 (2018). doi:10.1109/IROS.2018.8594092.

50. Kaufmann, E. *et al.* Champion-level drone racing using deep reinforcement learning. *Nature 2023 620:7976* 620, 982–987 (2023).

51. Song, Y., Romero, A., Müller, M., Koltun, V. & Scaramuzza, D. Reaching the limit in autonomous racing: Optimal control versus reinforcement learning. *Sci Robot* 8, eadg1462 (2023).

52. Bongard, J., Zykov, V. & Lipson, H. Resilient machines through continuous self-modeling. *Science (1979)* 314, 1118–1121 (2006).

53. Hwangbo, J. *et al.* Learning agile and dynamic motor skills for legged robots. *Sci Robot* 4, (2019).

54. Lee, Y. H. *et al.* Whole-Body Control and Angular Momentum Regulation using Torque Sensors for Quadrupedal Robots. *J Intell Robot Syst* 102, 1–15 (2021).

55. Choi, S. *et al.* Learning quadrupedal locomotion on deformable terrain. *Sci Robot* 8, (2023).

56. Yang, Y. *et al.* Data Efficient Reinforcement Learning for Legged Robots. *Proc Mach Learn Res* 100, 1–10 (2019).

57. Mahler, J. *et al*. Learning ambidextrous robot grasping policies. *Sci Robot* 4, (2019).

58. Lázaro-Gredilla, M., Lin, D., Swaroop Guntupalli, J. & George, D. Beyond imitation: Zero-shot task transfer on robots by learning concepts as cognitive programs. *Sci Robot* 4, (2019).

59. Fazeli, N. *et al.* See, feel, act: Hierarchical learning for complex manipulation skills with multisensory fusion. *Sci Robot* 4, (2019).

60. Ficuciello, F., Migliozzi, A., Laudante, G., Falco, P. & Siciliano, B. Vision-based grasp learning of an anthropomorphic hand-arm system in a synergy-based control framework. *Sci Robot* 4, (2019).

61. Hiller, J. & Lipson, H. Dynamic Simulation of Soft Multimaterial 3D-Printed Objects. *https://home.liebertpub.com/soro* 1, 88–101 (2014).

62. Shah, D. S. *et al.* A soft robot that adapts to environments through shape change. *Nat Mach Intell* 3, 51–59 (2020).

63. Tiryaki, M. E., Demir, S. O. & Sitti, M. Deep Learning-based 3D Magnetic Microrobot Tracking using 2D MR Images. *IEEE Robot Autom Lett* **7**, 6982–6989 (2022).

64. Pane, S., Iacovacci, V., Sinibaldi, E. & Menciassi, A. Real-time imaging and tracking of microrobots in tissues using ultrasound phase analysis. *Appl Phys Lett* 118, 26 (2021).

65. Yan, X. *et al.* Multifunctional biohybrid magnetite microrobots for imaging-guided therapy. *Sci Robot* 2, (2017).

66. Iacovacci, V. *et al.* High-Resolution SPECT Imaging of Stimuli-Responsive Soft Microrobots. *Small* 15, (2019).

67. Reale, R., De Ninno, A., Businaro, L., Bisegna, P. & Caselli, F. High-throughput electrical position detection of single flowing

particles/cells with non-spherical shape. *Lab Chip* 19, 1818–1827 (2019).

68. Masner, A., Blasina, F. & Simini, F. Electrical impedance tomography for neonatal ventilation assessment: A narrative review. in *Journal of Physics: Conference Series* vol. 1272 (Institute of Physics Publishing, 2019).

69. Rapin, M. *et al.* Wearable Sensors for Frequency-Multiplexed EIT and Multilead ECG Data Acquisition. *IEEE Trans Biomed Eng* 66, 810–820 (2019).

70. Shi, X. *et al.* High-precision electrical impedance tomography data acquisition system for brain imaging. *IEEE Sens J* 18, 5974–5984 (2018).

71. Darnajou, M. *et al.* High Speed EIT With Multifrequency Excitation Using FPGA and Response Analysis Using FDM. *IEEE Sens J* 20, 8698–8710 (2020).

72. Zhang, Y., Xiao, R. & Harrison, C. Advancing hand gesture recognition with high resolution electrical impedance tomography. in *UIST 2016 - Proceedings of the 29th Annual Symposium on User Interface Software and Technology* 843–850 (Association for Computing Machinery, Inc, 2016). doi:10.1145/2984511.2984574.

73. Demir, S. O., Tiryaki, M. E., Karacakol, A. C. & Sitti, M. Learning Soft Millirobot Multimodal Locomotion with Sim-to-Real Transfer. *Advanced Science* (2024) doi:10.1002/advs.202308881.

74. Wang, C., Wu, Y., Dong, X., Armacki, M. & Sitti, M. In situ sensing physiological properties of biological tissues using wireless miniature soft robots. *Sci Adv* 9, (2023).

75. Duvenaud, D., Nickisch, H. & Rasmussen, C. E. Additive Gaussian Processes. in *Proceedings of the 24th International Conference on Neural Information Processing Systems* 226–234 (Curran Associates Inc., Granada, Spain, 2011). doi:10.5555/2986459.2986485.

76. Chen, Z. & Wang, B. How priors of initial hyperparameters affect Gaussian process regression models. *Neurocomputing* 275, 1702–1710 (2018).

77. Jones, D. R., Schonlau, M. & Welch, W. J. Efficient Global Optimization of Expensive Black-Box Functions. *Journal of Global Optimization* 13, 455–492 (1998).

78. Brochu, E., Cora, V. M. & de Freitas, N. A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning. (2010).

79. Ibarz, J. *et al.* How to train your robot with deep reinforcement learning: lessons we have learned. *International Journal of Robotics Research* 40, 698–721 (2021).

80. Hiller, J. & Lipson, H. Dynamic Simulation of Soft Multimaterial 3D-Printed Objects. *Soft Robot* 1, 88–101 (2014).

81. Silvera-Tawil, D., Rye, D., Soleimani, M. & Velonaki, M. Electrical impedance tomography for artificial sensitive robotic skin: A review. *IEEE Sens J* 15, 2001–2016 (2015).

82. Adler, A. & Lionheart, W. R. B. Uses and abuses of EIDORS: an extensible software base for EIT. *Physiol Meas* 27, (2006).

83. Kim, Y., Yuk, H., Zhao, R., Chester, S. A. & Zhao, X. Printing ferromagnetic domains for untethered fast-transforming soft materials. *Nature* 558, 274–279 (2018).

84. Xu, T., Zhang, J., Salehizadeh, M., Onaizah, O. & Diller, E. Millimeter-scale flexible robots with programmable three-dimensional magnetization and motions. *Sci Robot* 4, (2019).

85. Alapan, Y., Karacakol, A. C., Guzelhan, S. N., Isik, I. & Sitti, M. Reprogrammable shape morphing of magnetic soft machines. *Sci Adv* 6, 6414–6432 (2020).

86. Heess, N. *et al.* Emergence of Locomotion Behaviours in Rich Environments. *arXiv preprint arXiv:1707.02286* (2017) doi:10.48550/arXiv.1707.02286.

87. Yao, J. *et al.* Adaptive Actuation of Magnetic Soft Robots Using Deep Reinforcement Learning. *Advanced Intelligent Systems* 5, (2023).

88. Tariverdi, A. *et al.* A Recurrent Neural-Network-Based Real-Time Dynamic Model for Soft Continuum Manipulators. *Front Robot AI* 8, 631303 (2021).

89. Schaff, C., Sedal, A. & Walter, M. R. Soft Robots Learn to Crawl: Jointly Optimizing Design and Control with Sim-to-Real Transfer. in *Robotics: Science and Systems* (2022).

Publication 1

Learning of Sub-optimal Gait Controllers for Magnetic Walking Soft Millirobots

Abstract

Untethered small-scale soft robots have promising applications in minimally invasive surgery, targeted drug delivery, and bioengineering applications as they can access confined spaces in the human body. However, due to highly nonlinear soft continuum deformation kinematics, inherent stochastic variability during fabrication at the small scale, and lack of accurate models, the conventional control methods cannot be easily applied. Adaptivity of robot control is additionally crucial for medical operations, as operation environments show large variability, and robot materials may degrade or change over time, which would have deteriorating effects on the robot motion and task performance. Therefore, we propose using a probabilistic learning approach for millimeter-scale magnetic walking soft robots using Bayesian optimization (BO) and Gaussian processes (GPs). Our approach provides a data-efficient learning scheme to find controller parameters while optimizing the stride length performance of the walking soft millirobot robot within a small number of physical experiments. We demonstrate adaptation to fabrication variabilities in three different robots and to walking surfaces with different roughness. We also show an improvement in the learning performance by transferring the learning results of one robot to the others as prior information.

I. Introduction

Soft-bodied robots are composed of functional soft materials exhibiting shape-programmable properties that allow passive/active structural compliance and large degrees of freedom, which are hard to achieve using

conventional rigid materials [16]. The research on soft robots is getting more attention owing to easier access to novel fabrication methods and functional materials, and potential high-impact medical and other applications [25]. Biologically inspired soft robots can be used to study their soft-bodied biological counterparts [17], and open new application areas in multi-terrain locomotion [4], adaptive manipulation [15, 28], and human-assistive wearable systems [41]. Soft robots also enable safe human-robot physical interaction due to their high compliance and limited output force, which normally require additional computational effort in conventional robotic systems [12]. Small-scale (i.e., millimeter) untethered soft robots have further potential usage in medicine owing to their ability to access to enclosed small spaces non-invasively [26, 35] and the embodiment of functionalized materials enabling targeted drug delivery and bio-sensing [5].

Despite their potential, the virtual infinite degrees of freedom, the lack of accurate models, fabrication variations, and non-linear behavior (e.g., hysteresis) render the application of conventional control methods challenging for soft robots [32]. So far, constant curvature (CC) models utilizing bending beam theories have been widely-used to approximately represent the deformation of continuum robots [42]. Alternatively, analytical and geometrically exact models have been suggested for continuum robots that are represented as simplified rods [29]. Finite element methods (FEM) provide numerical solutions to soft robot kinematics by utilizing a chain of rigid elements connected with tunable spring-damper mechanisms [21]. These kinematic models allow the implementation of static and dynamic controllers for continuum robots

on a larger scale [9]. However, these controllers typically depend on the continuous sensing of body deformations from embedded sensors and computationally heavy model solutions, which are conditions that may not be met for untethered soft robots at the small scales [30]. The dynamic task environment, complex deformation kinematics, fabrication-dependent performance variations, and actuation/sensing limitations have further impacts on the soft mobile robots targeting medical applications, which make adaptive and data-efficient control methods attractive for these robots [36].

In the case of uncertainty and lack of a parametric model that represents the system, data-driven control [13] and reinforcement learning [38, 19] provide promising alternatives over model-based designs in small-scale soft robotic systems. However, the need for data efficiency, i.e., the ability to learn from only a few experimental trials, presents a core challenge for such methods [6]. Conversely, Bayesian optimization (BO) [10, 34] allows for the maximization of a performance function using a small number of physical experiments. BO typically employs Gaussian processes (GPs) [27] as a probabilistic model of the latent objective function. While no explicit dynamics model is needed, GPs allow for incorporating information as probabilistic priors, thus reducing data requirements. There are emerging examples that demonstrate the application of this approach to optimize the locomotion performance of robots on different length scales [3, 44, 22]. Despite its potential to address the control challenge for untethered soft robots, there are only a few examples that apply this method such as in the gait exploration of a tensegrity system [31], and the optimization of an undulating motion of a

microrobot [40]. So far, a data-efficient procedure that adapts the learned controllers to different robots and environmental conditions for untethered small-scale soft robots has not been demonstrated.

In this paper, we propose a learning procedure to find the controller parameters of magnetically actuated, untethered, soft millirobots (see Fig. 1) that generate optimum walking gaits within a small number of physical experiments. We specifically focus on these types of robots due to their biocompatible use of external magnetic actuation that supports multi-functionality in future medical tasks [14, 23] and the high-resolution magnetization methods that allow more complex deformation capabilities at the small scale [24, 7]. We produce three replicas of a previously reported soft millirobot demonstrating a hand-tuned walking gait aiming for medical applications [14] and test the repeatability of their results. We begin with finding the optimum walking gait controllers of our robots using BO with GPs; initially without any prior information about the correlation between the controller parameters and the robot performance. Later, we explore the controller parameter space of one robot and present a straightforward way in the context of BO to transfer prior information from the first robot to all three robots while finding their optimum walking gait controllers in a data-efficient fashion. We report the optimum controllers and walking gait performances in terms of achieved stride lengths for all three robots and compare the two learning approaches (i.e., with and without using prior information). We also transfer this information to adapt to the changes in the task environment by finding the controller parameters for walking on rough surfaces.

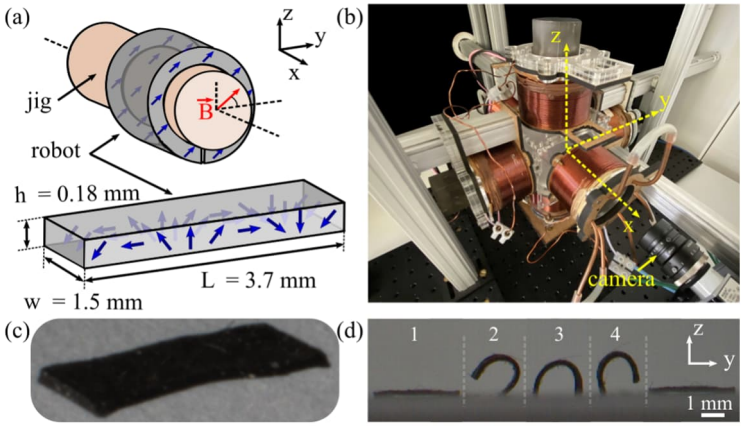


Figure 1. (a) The magneto-elastomer robot is rolled around a jig and magnetized with $|\vec{B}| = 1.8T$ field (red arrow) with a 45° angle with respect to the y-axis. The unfolded robot maintains a circular magnetization profile along its body (blue arrows). (b) Photo of the experimental setup with 6 electromagnetic coils and a high-speed camera. (c) Image of the fabricated and magnetized soft millirobot. (d) Projected planar images showing the four consecutive states of the robot walking gait: (1) relaxed, (2) front-stance, (3) double-stance, and (4) back-stance. These images are placed with a separation on the y-axis for visual clarity, i.e., the robot does not jump in between states during the experiments. Numbers represent the four states.

The organization of this paper is as follows. We describe the robot design and its walking gaits in Section II and introduce our learning approach in Section III. Section IV presents the experiments on learning of the optimum gait controllers without using the prior information, the generation of the prior controller information from one robot, and the optimization of the walking gaits by transferring the learned controllers to three robots and different locomotion surfaces. We discuss the experimental results in Section V and conclude our work in Section VI.

II. Robot Design and Gait Definition

We followed the methods and materials reported in [14] and fabricated three magnetic soft millirobots with a 1:1 body mass ratio of Ecoflex 00-10 (Smooth-On Inc.) and neodymium-iron-boron (NdFeB) magnetic microparticles with around $5\mu m$ diameter (MQP-15-7, Magnequench). We placed this pre-polymer mixture on a methyl methacrylate plate and cut the robots out of the cast using a high-resolution laser cutter (LPKF Protolaser U4) after the polymer is cured. Our robots had the final dimensions of length $L = 3.7mm$, width $w = 1.5mm$, and height $h = 185\mu m$ as shown in Fig. 1-a. We separately folded the robots around a circular jig with a circumference equal to L and magnetized them within a magnetic field with a magnitude of $1.8T$ and orientation of $45°$ measured counterclockwise from the y-axis. Once the robots are unfolded from the jig, the magnetic particles maintained their magnetization orientation forming a circular profile along the longitudinal axis of the robot body (Fig. 1-a). We used these robots (i.e., robots 1, 2, and 3) with the same nominal material properties and dimensions for our experiments (see Fig. 1-c for a sample robot image).

The walking gait of our robot is composed of four consecutive quasi-static states that are inspired by the planar quadrupedal bounding [1] and a caterpillar's inching motion [39]. These states are depicted as (1) relaxed, (2) front-stance, (3) double-stance, and (4) back-stance as shown in Fig. 1-d. We placed our magnetized robot along the y-axis of

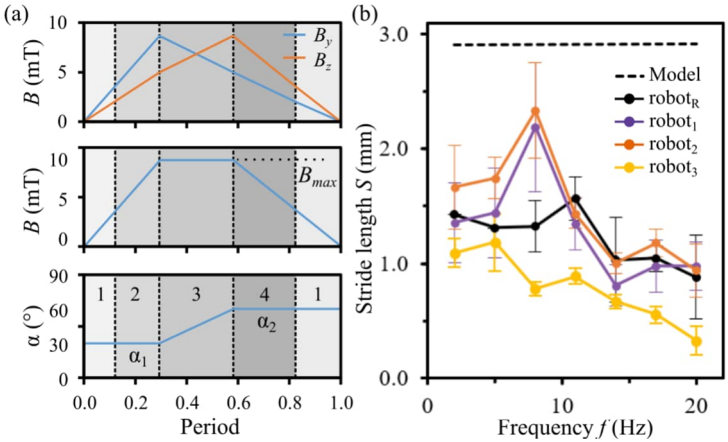


Figure 2. (a) Walking gait control parameters during a single period of a sample motion (a single period of $1/f = 90ms$ for $f = 11Hz$ is normalized to 0-1 on the abscissa). The magnetic field B is controlled on the y-z plane and shown with its y and z components (top) whose magnitude (middle) reaches B_{max} and orientation (bottom) changes from α_1 to α_2. Dashed vertical lines represent the (1) relaxed, (2) front-stance, (3) double-stance, and (4) back-stance states of the walking gait. (b) The stride length S performance of the previously reported robot ($robot_R$) vs. the performance of three replica robots using the same controller parameters ($robot_{1,2,3}$). Each data point for each robot ($robot_{1,2,3}$) represents the mean of 10 experiments and the error bars show the standard deviation. The performance of $robot_R$ and the horizontal dashed line which represents the model prediction are adapted from [14].

the magnetic coil setup consisting of three orthogonal pairs of custom-made electromagnets (Fig. 1-b) that generated a 3-D uniform magnetic field within a 4×4×4cm^3 space. We modulated the magnetic field on the y-z plane that coincided with the center of the test environment. We controlled four parameters to generate the walking gait: the maximum magnetic field magnitude (B_{max}), the frequency of the actuation cycle

(f), and two magnetic field orientation angles (α_1 and α_2) measured counterclockwise from the y-axis. The plots in Fig. 2-a show the change of the control parameters during a single period of the motion for $B_{max} = 10mT$, $f = 11Hz$, $\alpha_1 = 30°$ and $\alpha_2 = 60°$, which are hand-tuned parameters reported in [14]. At the beginning of a single gait period, the robot started at a relaxed state for $0 \leq B \leq 4mT$. The robot tilted forward when $\alpha = \alpha_1$ and B increased from $4mT$ to $B_{max} = 10mT$. While B remained constant at B_{max}, the orientation of the magnetic field changed from α_1 to α_2 causing the robot to initially switch to the double-stance state and then to the back-stance state when $\alpha = \alpha_2$. Then, B decreased while keeping the orientation of the magnetic field constant, and the robot gradually switched back to the relaxed state. For $B < 4mT$, the robot assumed the relaxed state, and a single period of walking actuation ended when $B = 0mT$. We reset B at the end of every gait cycle to avoid jerky motion when α changed from α_1 to α_2. In our experiments, the relaxed state was never skipped but its duration changed according to f. The consecutive images from a single walking gait period are shown in Fig. 1-d. We tracked the robot gait using a high-speed camera (Basler aCa2040-90uc, 60 frames per second (fps), 1pixel~$27\mu m$ resolution) that is orthogonally placed to the axis of robot motion (Fig. 1-b). In every experiment, we calculated the stride length (S) of the robot by tracking the average distance covered by its center of mass in 10 consecutive steps.

To test the repeatability of the previously reported results in [14], we experimented with our fabricated robots using their suggested controller parameter sets. Fig. 2-b shows the stride length performances of our

robots ($robot_{1,2,3}$) and compares them with the reported robot ($robot_R$) performance (i.e., we calculated the stride length of $robot_R$ from the values reported in [14]) for $B_{max} = 10mT$, $\alpha_1 = 30°$, $\alpha_2 = 60°$, and $2 \leq f \leq 20Hz$. Our preliminary results revealed that:

- In this scale, the gait performance showed clear inconsistency due to the variability during fabrication and environmental factors even though the same materials, methods, controller parameters, and walking surfaces are used in the fabrication and experimentation of the millimeter-scale soft robots.
- Unlike the model prediction, the robot performance showed non-monotonic behavior along with increasing f, which rendered the design of a model-based gait controller unreliable.
- In addition to the virtual infinite degrees of freedom inherited by the soft materials, the controller parameters existed in a continuous space, making the hand-tuning of these parameters within physical experiments impractical.

These observations found the goals of our paper in which we address the necessity for a data-efficient controller learning system that is robust to the variabilities caused by the material, fabrication, and the task environment of the miniature scale, medical-oriented, untethered soft robots.

III. Learning Approach

We aim to optimize the walking gait controller parameters to maximize the stride length S of the robot. Therefore, we define the reward function as

$$S: \Theta \rightarrow \mathbb{R}, \tag{1}$$

which maps the parameter set $\theta = [B_{max}, f, \alpha_1, \alpha_2]$ to scalar reward values (i.e., the experimental stride length performance of a robot). According to the definition of the reward function, we formulate the parameter learning as the (global) optimization problem

$$\theta^* = \underset{\theta \in \Theta}{\operatorname{argmax}} S(\theta), \tag{2}$$

where Θ denotes the complete search space, θ is the parameter set, and $S(\theta)$ is the experimentally observed stride length performance of the robot for a given θ.

We define the range of the controller parameters based on the findings in [14] and the physical limitations of our magnetic actuation setup. Accordingly, B_{max} is defined between $5mT$ and $12mT$, and the walking frequency, f, ranges from $f_{min} = 0.5Hz$ to $f_{max} = 20Hz$. We limit α_1 and α_2 to $[10, 50]°$ and $[40,80]°$ respectively and select values that satisfy $\alpha_2 > \alpha_1$ to generate the walking gait in Fig. 1-d. We use a step size of $1mT$ for B, $1°$ for each α, and a variable step size of $0.25Hz$ for $f < 2Hz$ and $2Hz$ for $f \geq 2Hz$, which yield a total number of 203520 possible parameter sets in Θ.

A. Gaussian Processes (GPs)

The magnetic soft millirobots in our paper did not have an accurate kinematic or dynamics model (see Fig. 2-b). Therefore, it is necessary to approximate the reward function based on the data collected from physical experiments rather than numerical analysis. However, the physical data has inherent uncertainty due to the noise in the measurements and the variations during the experiments. To include these uncertainties in the model, overcome the sparsity in the data, and make probabilistic predictions at unobserved locations, we model the reward function $S(\theta)$ using GPs following the previous study in [40]:

$$S(\theta) \sim GP\big(\mu(\theta), k(\theta, \theta')\big), \tag{3}$$

However, as $S(\theta)$ can only be measured with noise, we define $\tilde{S}$ as

$$\tilde{S}(\theta_i) = S(\theta_i) + n_i, \tag{4}$$

where n_i is zero-mean Gaussian noise with variance σ_n^2 for each measurement i.

A GP is a non-parametric model defined by its prior mean $\mu(\theta)$ and the covariance function $cov\big(S(\theta), S(\theta')\big) = k(\theta, \theta')$, where k is the kernel. During one run of BO, the GP model is sequentially updated with $\tilde{S}(\theta)$ observed from experiments. We define one "learning run" as a run of BO until the desired stopping criterion is matched (e.g., a fixed number of experiments is reached).

From the experimental data $D = \{\theta_i, \tilde{S}(\theta_i)\}_{i=1}^{N}$, the stride length of the robot for an unobserved θ can be predicted using the posterior mean and variance as follows.

$$\mu_{post}(\theta) = \mu(\theta) + k^T(\theta)K^{-1}y, \tag{5}$$

$$\sigma^2_{post}(\theta) = k(\theta,\theta) - k^T(\theta)K^{-1}k(\theta), \tag{6}$$

$$S_{post}(\theta)|D \sim N(\mu_{post}(\theta), \sigma^2_{post}(\theta), \tag{7}$$

where $k(\theta)$, $y \in \mathbb{R}^N$ with $k(\theta_i) = k(\theta,\theta_i)$, $y_i = \tilde{S}(\theta_i) - \mu(\theta_i)$, and $K \in \mathbb{R}^{N\times N}$ with $K_{i,j} = k(\theta_i,\theta_j) + \delta_{i,j}\sigma_n^2$, where $\delta_{i,j}$ is the Kronecker delta and σ_n^2 is the noise in the collected data set.

We select the squared exponential as the kernel function in the GPs, which is for the 1D case:

$$k_{SE}(\theta,\theta') = \sigma_f^2 \exp(-(\theta-\theta')^2/2l_c^2), \tag{8}$$

where l_c is the length scale that defines the rate of variation in the modeled function for each dimension of the parameter space. Long length scales are used to model slowly-varying functions and short length scales are used to model quickly-varying functions. The signal variance σ_f^2 describes the width of distribution, e.g., high σ_f^2 means higher uncertainty in the predictions of the unobserved θ. Hyperparameters of the GPs can be listed as the noise in the collected data σ_n^2, length scale l_c for each dimension of the parameter space $\mathbb{R}^{d_c}$, and signal variance σ_f^2. To determine the value of σ_n^2, we use the maximum variance found in the

experimental results shown in Fig. 2-b. We set the length scale l_c to one-fourth of the total range of each corresponding parameter. We also set the signal variance σ_f^2 to half of the body length of the robot so that the highest possible reward value (i.e., $L = 3.7mm$) remained inside the 95% confidence interval of the prior.

B. Bayesian Optimization (BO)

We use BO to select the parameter set θ_{next} to be tested in the next step of the learning run using the acquisition function $\alpha_{acq}(\theta)$.

$$\theta_{next} = \underset{\theta \in \Theta}{\operatorname{argmax}}\, \alpha_{acq}(\theta). \tag{9}$$

In this work, we choose the expected improvement (EI) as the acquisition function $\alpha_{acq}(\theta)$ due to its better performance compared to its alternatives as demonstrated in [40]. EI seeks the parameter set for the next step where the expected improvement in reward function is the highest compared to the previously collected data:

$$\alpha_{acq}(\theta) = \mathbb{E}\big[\max\big(0, (S(\theta), S(\theta^*) - \xi)\big)\big], \tag{10}$$

$$\alpha_{acq}(\theta) = (\mu(\theta) - S(\theta^*) - \xi)\Phi(Z) + \sigma(\theta)\phi(Z), \tag{11}$$

where $S(\theta^*)$ is the highest reward function value collected so far, Φ and ϕ are the Gaussian cumulative density and probability density functions, respectively [2]. The term Z is described as $Z = Z(\theta) = (\mu(\theta) - S(\theta^*) - \xi)/\sigma(\theta)$, with $\mu(\theta)$ and $\sigma(\theta)$ computed from Eq. (5) and Eq. (6). In Eq. (11), the two terms define the exploitation and the exploration weights of the BO, respectively. The balance between

these two terms is controlled by the hyperparameter ξ. As ξ gets higher, BO focuses more on exploration and seeks the next parameter set in regions with high prediction uncertainty. Conversely, BO focuses more on exploitation and selects the next parameter set within a close range to already explored regions. We choose $\xi = 0.1$ to balance the exploration and exploitation weights.

C. Transfer of the Prior Mean

In addition to the kernel (see Section III-A), the prior mean $\mu(\theta)$ must be chosen at the beginning of a BO run. Often, $\mu = 0$ is the default choice for an uninformed prior. For the millirobot learning problem herein, we suggest and investigate the transfer of information from previous learning runs by setting the prior mean to the posterior mean of a previously trained GP model, such as from a different robot. In this way, we can approximately transfer the topology of the target function between robots, which is reasonable as long as the differences between the robots and the environment do not significantly alter the function shape. In this work, we adopt and compare both approaches of an uninformed prior ($\mu(\theta) = 0$) in Section IV-A, and the transfer of the posterior mean from robot 1's previous run to all three robots in Section IV-B.

IV. Experiments

We modulated the currents running through the electromagnetic coils and the resulting magnetic field by controlling six motor driver units (SyRen25) using an Arduino microcontroller running at 1.2 kHz. We regularly calibrated the magnetic actuation matrix inside the workspace,

i.e., the mapping between the applied electric current and the generated magnetic field, to maintain reliable and repeatable experiments. The learning process ran on a PC that additionally handled image processing and hardware communication tasks. One step of the learning run involved five steps:

1) BO selected a new parameter set θ that maximized the acquisition function based on the GP model,
2) The microcontroller regulated the magnetic field based on the selected θ and initiated the physical experiment,
3) The camera recorded the robot motion and measured the average stride length performance S,
4) The learning system updated the GP model using the newly collected data from the experiment,
5) The robot returned to its initial position for the next step.

A. Optimization of the Walking Gait without the Prior

To test our controller learning approach without prior information, i.e., $\mu(\theta) = 0$, we experimented with all three robots in the same environmental conditions and limited the number of steps for each learning run to 20 experiments. We initialized the BO with the best controller parameter set reported in [14]. We performed three independent learning runs (i.e., 60 experiments in total) for every robot with the same initial state, whose results are shown in Fig. 3. Each data point represents the robot's stride length performance S resulting from a different controller parameter set chosen by the BO at a given step of a

learning run. As BO actively chose sample locations (e.g., to explore unknown regions described in Section III), the variation in these data points was the desired behavior of the explorative learning algorithm. See Supplementary Video 1 for the gait performances of four different controller parameter sets for robot 1.

We chose the optimum controller parameter sets (θ^*) from these learning runs and repeated the walking gait five times to collect statistical information about the stride length performances. The rows designated with "no prior" in Table I shows the values for θ^* and the resulting S for each robot. It can be seen that the walking gait performances of the robots were significantly improved compared to the robot reported in [14]. Also, the standard deviation within these repeated experiments agreed with the previously reported values, platform. In this optimization approach, we achieved 86.6%, 94.7%, and 60.5% increase in S for robots 1 to 3 respectively (i.e., compared to the S of the previous robot shown in the last row of Table I). The difference between the optimum controller parameter values in Table I demonstrates the influence of the fabrication variabilities on the robot design and performance. Separate from the optimum stride length performance of each robot, we evaluated the overall performance of the learning runs based on the achieved stride length average of all of the tested controller parameters. This performance metric P_{LR} shows the overall quality of the learning run's parameter selection in terms of the average of all the S and the standard deviation in 60 experiments (i.e., avg(S) ± std). In these experiments, the learning run for robot 1 yielded $P_{LR_1} = 1.07 \pm 0.80mm$, $P_{LR_2} = 1.21 \pm 0.88mm$ for

robot 2, and $P_{LR_3} = 0.75 \pm 0.64mm$ for robot 3. Even though there were multiple individual θs within these runs (e.g., the optimums reported in Table I) that outperformed the previous study, the large standard deviation shows that the BO selected parameters that generated a wide range of performances.

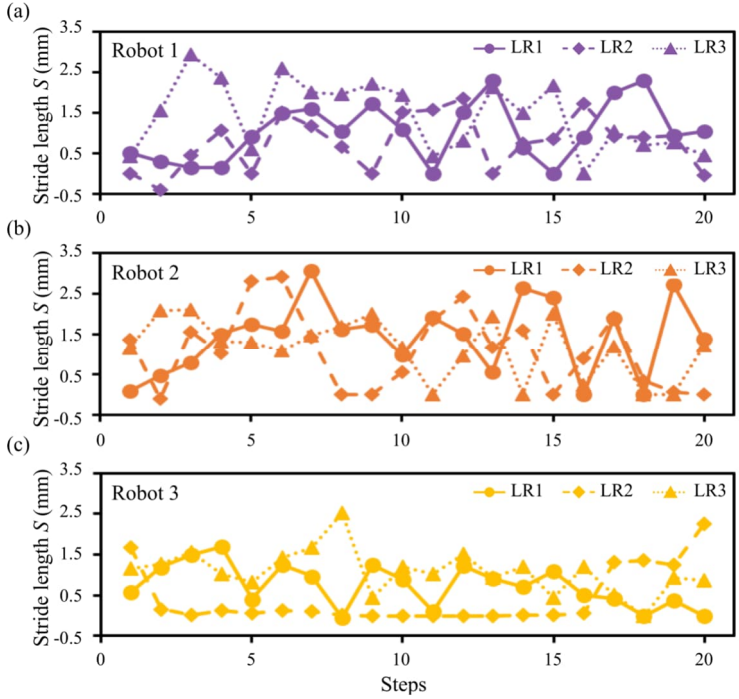


Figure 3. The learning of the controller parameters without utilizing the prior information within 20 physical experiments in 3 independent learning runs (shown as LR1-3) for (a) robot 1, (b) robot 2, and (c) robot 3.

B. Optimization of the Walking Gait with the Prior

1) Generation of the Prior Information

To generate useful prior information, we constructed the posterior mean $\mu_{post}(\theta)$ for robot 1 using the BO and GP described in Section III. Initially, we adopted nine different controller parameters from [14] and collected the stride length information from repeated experiments in our setup. Fig. 4-a and 4-c show the two-dimensional projection of the approximation of S function generated by the GP model (utilizing the same hyperparameters in Section III) based on these experiments. After the initial approximation, we used the BO to select new parameter sets from the unexplored parts of the 4-D search space and collected the experimental stride length performance information. We explored 123 different parameter sets in total by selectively isolating the search space dimensions. Initially, we fixed $\alpha_1 = 30°$ and $\alpha_2 = 60°$ and explored 18 different parameter values for B_{max} and f. Then, we fixed $B_{max} = 10mT$ and $f = 2Hz$ and explored 38 values for α_1 and α_2. We performed 17 additional tests for α_1 and α_2 for $B_{max} = 10mT$ and $f = 8Hz$. Finally, we explored the complete search space for four parameters with 50 more tests. For all of these tests, we stopped the exploration when the BO converged in the sense of repetitively selecting similar θs. Fig. 4 shows the two-dimensional projections of the GP-based probabilistic approximation of the performance function before (Fig. 4-a,c) and after (Fig. 4-b,d) all of the physical experiments dictated by our BO. These results show that our BO approach revealed parts of the parameter space that were not effectively explored using the hand-tuning in [14]. We used this posterior information of robot 1 and transferred it to all robots

(i.e., robots 1, 2, and 3) as prior information of the stride length function approximation in the remaining part of the optimization experiments.

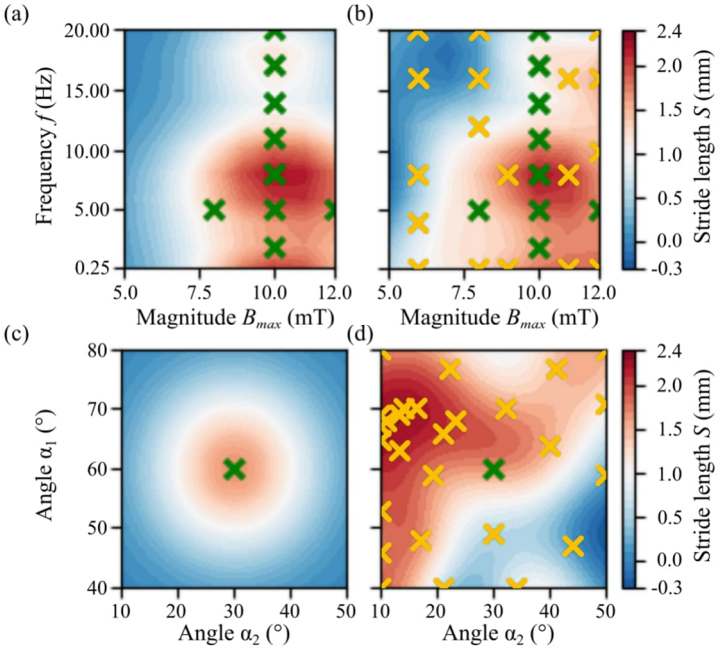


Figure 4. Approximation of the stride length performance as a function of four control parameters using GPs. The upper row shows S projected on B-f plane for $\alpha_1 = 30°$ and $\alpha_2 = 60°$, and the lower row shows S projected on the α_1-α_2 plane for $B = 10mT$ and $f = 2Hz$. (a,c) Initial approximation of S applying the 9 hand-tuned controller parameters reported in [14] on robot 1 in our experiments. Each green cross mark represents 10 trials for the chosen parameter set. (b,d) The final probabilistic approximation of S after running the prior information generation step. Experiments with the parameters selected by our BO are represented with yellow cross marks.

2) Transfer of Learning Between Different Robots

Similar to the experiments in Section IV-A, we performed three independent learning runs each consisting of 20 experiments for every robot. Unlike the previous learning approach, the GP model in every learning run started with the prior mean set to the posterior mean information of robot 1 that was generated in Section IV-B1. Fig. 5 shows the walking gait performance results of three robots in these learning runs. The optimum controller parameter sets (θ^*) and the resulting stride length performances S from these learning runs are reported in Table I on the rows designated with "prior". Compared to the robot in [14], we achieved optimized walking gaits with an increased performance of 70.7%, 73.9%, and 113.3% for robot 1 to 3, respectively. See Supplementary Video 2 for a comparison of the walking gaits of three robots with the optimum of the parameters found in the experiments.

The utilization of the transferred prior information can be seen as a clear improvement in the overall learning run performance P_{LR}. In these experiments, the learning runs for robot 1 yielded $P_{LR_1} = 1.45 \pm 0.43mm$, $P_{LR_2} = 1.56 \pm 0.42mm$ for robot 2, and $P_{LR_3} = 1.43 \pm 0.70mm$ for robot 3. The improved averages compared to the results in Section IV-A show that once the prior information is transferred, the BO selected parameters that yielded better performing stride lengths in the same number of limited physical experiments. Likewise, the lower deviation in the averages implies that the performance range of the selected parameters was consistent.

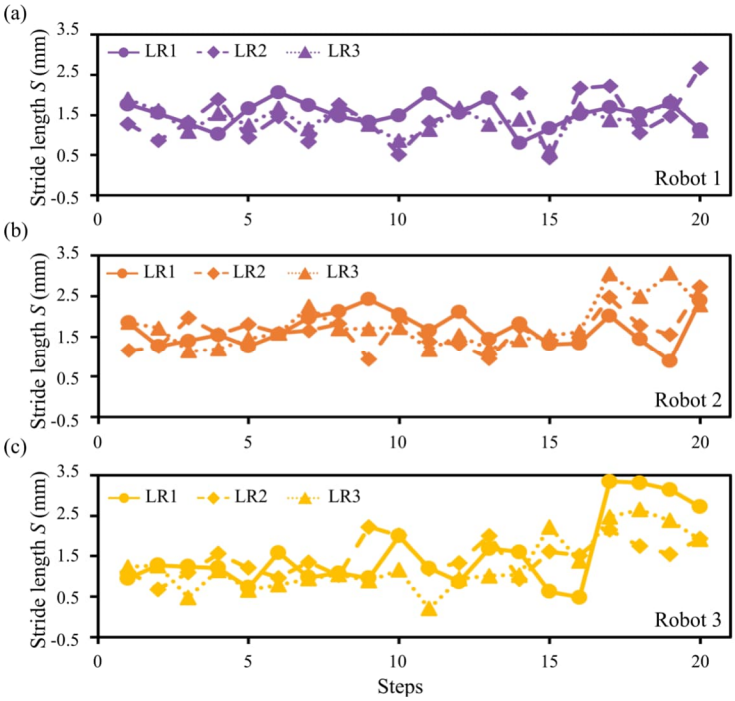


Figure 5. The learning of the controller parameters by utilizing the prior information presented in Section IV-B1 within 20 physical experiments in 3 independent learning runs for (a) robot 1, (b) robot 2, (c) robot 3.

C. Adaptation to Different Surfaces

Similar to the variances during the fabrication, the changes in the task environment also have a significant impact on the untethered soft small-scale robots. To demonstrate the adaptation capability of our learning approach to different surfaces, we experimented with robot 1 on a surface coated with 60-grit sandpaper (Klingspor, KL385-JF), which had a higher

Table I Selected gait controller parameters

Robot	Type	Controller Parameters				Stride length S (mm) (avg±std)
		B_{max} (mT)	f (Hz)	α_1 $(°)$	α_2 $(°)$	
Robot 1	no prior	9	10	21	61	2.25±0.19
	Prior	12	8	20	65	2.68±0.34
Robot 2	no prior	11	8	27	65	3.06±0.38
	Prior	9	10	32	73	2.73±0.24
Robot 3	no prior	10	10	19	80	2.52±0.27
	Prior	12	18	10	80	3.35±0.08
Robot in [14]		10	11	30	60	1.57±0.38

roughness compared to the plexiglass surface used in our previous experiments (Fig. 6-a). The surface profile examination in Fig. 6-b and 6-c (Keyence VK-X260K) shows that two surfaces had significant differences between their roughness Sq (root mean square height): sandpaper $Sq = 85.01\mu m$ compared to plexiglass surface $Sq = 0.38\mu m$. Fig. 6-c shows that the terrain of the rough surface had features almost three times taller than the height of our robot.

Initially, we used the optimum control parameters of robot 1 in Section IV-B2 on the rough surface and observed that the walking gait performance dropped from $S = 2.65 \pm 0.34mm$ to $S = 0.93 \pm 0.26mm$. Then, we optimized the robot on the rough surface using our learning approach utilizing the prior mean information generated in Section IV-B1. Within a single learning run of 20 experiments, we found a parameter set that increased the stride length performance to $S = 1.15 \pm 0.15mm$, yielding a 24.7% optimization

(Fig. 6-d). The optimum walking gait controller parameter sets for both surfaces are reported in Table II. These results show that the learning system adapted the controller parameters for the new terrain features to maintain a successful walking gait. During the optimization process, we observed that the robot typically got stuck inside the cavities that were larger than its height. To overcome this problem, BO optimized the parameters such that the robot moved slower with the lower f, and tilted back and forward with larger α_2 and smaller α_1 to release its "legs" from the cavities. See Supplementary Video 3 for a comparison between walking gaits on two surfaces.

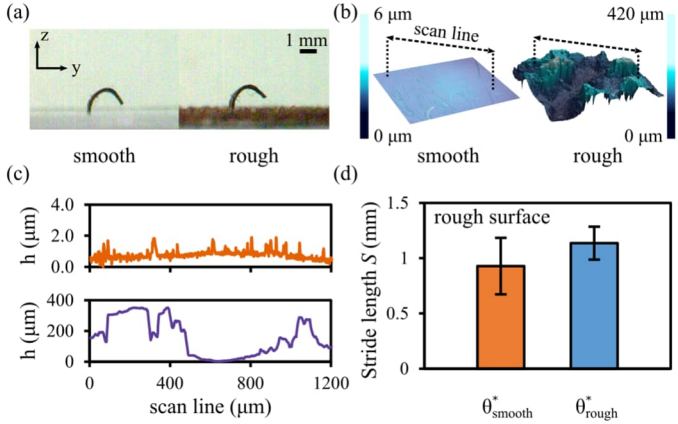


Figure 6. Adaptation to different surface roughnesses. (a) Colored images showing the robot walking on different surfaces: smooth plexiglass (left) and rough sandpaper (right). (b) Profilometer analysis showing the roughness difference between two surfaces. (c) Linear profiling along the scan line axis marked with dashed lines in (b) reveals the average height difference between surfaces (note the two orders-of-magnitude difference). (d) The average stride length performances of robot 1 on the rough surface using the parameters optimized for the smooth surface (left) vs. parameters using the prior information for the rough surface (right).

Table II Gait controller parameters for different surfaces

Surface	**Controller Parameters**				**Stride length** S (mm) (avg±std)
	B_{max} (mT)	f (Hz)	α_1 $(°)$	α_2 $(°)$	
Smooth	12	8	20	65	2.68±0.34
Rough	11	2	10	76	1.15±0.15

V. Discussions

When the stride length performance results in Table I are compared, it can be seen that some of the controller parameters selected without the prior information outperformed the parameters selected with the prior information. Regardless of the prior information, as BO is a probabilistic optimization algorithm and promotes some exploration, these results were expected. Nonetheless, all of these optimized parameters significantly outperformed the hand-tuned values in [14], highlighting one of the major contributions of our work. As a second contribution, we showed that transferring the posterior mean of one robot as the prior mean for the learning experiments of other robots lead to benefits in terms of improved average performance of learning runs P_{LR} as shown in Fig. 7. For robot 1, the average in the P_{LR} increased by 35.5%, 29.3% for robot 2, and 91% for robot 3. We note, however, that even though our method showed positive influence for the considered robot cases, further investigation on the most appropriate means of transfer for the considered problem is interesting future work.

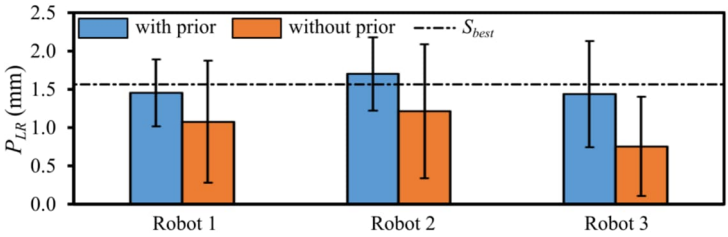


Figure 7. Improvement in the controller learning procedure represented with the overall stride length average P_{LR} and its standard deviation from the learning runs with and without using prior information. The dashed line (S_{best}) refers to the best stride length performance of the robot in [14].

Our results can reveal design guidelines to improve the kinematic models of the small-scale robots while utilizing the constant curvature (CC) approximations [42], analytical models [29], and FEM methods [21]. Additionally, recent studies suggesting fabrication methods with higher magnetization resolution on a smaller scale [43, 18] may address the fabrication variability problem owing to their automated procedures. However, especially for robots designed for non-invasive medical operations, the interaction with the dynamic task environment may still have degrading effects on the robot's soft body and change its performance unpredictably. In the absence of an adaptive online controller with a high-bandwidth feedback system, a data-efficient controller learning system may adapt the previously optimum controller parameters to the changes in the robot. For example, such an adaptive learning system may be applied for endoscopic soft robots within or outside the gastrointestinal (GI) tract [37, 11] using a small number of trials. Contributing to this idea, our paper demonstrated the data-efficient

learning of controller parameters and adaptation to different task environments without depending on the robot models whose results are shown in Table II.

In our experiments, we noticed that the duty factor of the double-stance state reduced with the increased actuation frequencies, which is commonly observed in legged locomotion in nature [1]. The highest stride length performances for all three robots were lower than the body-length (i.e., $L = 3.7mm$) of the robot, which also suggests that robots were following the walking gait state sequence by avoiding ballistic flight as in running. However, our approach can be extended to investigate the switch between dynamic gaits and the change of controller parameters accordingly.

In this paper, we focused on finding the optimal walking gait parameters inside this Θ using only physical experiments with BO and GPs. The systematic comparison of our experimental approach to alternative methods supported with simulations, such as intelligent trial and error [8], evolution algorithms [20], or policy gradients [33] is also an interesting future work.

VI. Conclusions

The results in this paper show the potential of a control learning system that can learn the new robot parameters quickly, and adapt to variabilities in the absence of a model-based control for soft robots. Our experimental results suggest that the boundaries for the parameter search space may be widened further to explore richer behaviors in future studies. This study

can be further extended to involve the design parameters, such as the magnetic particle density in our robots, and guide the task-oriented design strategies for medical-oriented robots. Our long term vision is to build a completely autonomous system that can actuate, track, evaluate, and optimize a complex soft robot with minimum human involvement.

Acknowledgments

U.C. thanks the Alexander von Humboldt Foundation for the Humboldt Postdoctoral Research Fellowship and the Federal Ministry for Education and Research. S.O.D. thanks the Ministry of National Education of the Republic of Turkiye for the Doctoral Scholarship. We thank Wenqi Hu for his contributions in developing the magnetic coil setup. This work was funded in part by the Cyber Valley Initiative, Grassroots Initiative, the Max Planck Society, and the European Research Council (ERC) Advanced Grant “SoMMoR” Project with Grant No: 834531.

References

[1] R. McN. Alexander. The Gaits of bipedal and quadrupedal animals. The International Journal of Robotics Research, 3(2):49–59, 1984.

[2] Eric Brochu, Vlad M. Cora, and Nando de Freitas. A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. arXiv:1612.06830 [cs/LG], 2010.

[3] Roberto Calandra, Andre´ Seyfarth, Jan Peters, and Marc Peter Deisenroth. Bayesian optimization for learning gaits under uncertainty. Annals of Mathematics and Artificial Intelligence, 76(1-2):5–23, 2016.

[4] Marcello Calisti, Giacomo Picardi, and Cecilia Laschi. Fundamentals of soft robot locomotion. Journal of The Royal Society Interface, 14(130):20170101, 2017.

[5] Hakan Ceylan, Immihan C Yasa, Ugur Kilic, Wenqi Hu, and Metin Sitti. Translational prospects of untethered medical microrobots. Progress in Biomedical Engineering, 1(1):012002, 2019.

[6] Konstantinos Chatzilygeroudis, Vassilis Vassiliades, Freek Stulp, Sylvain Calinon, and Jean-Baptiste Mouret. A survey on policy search algorithms for learning robot controllers in a handful of trials. IEEE Transactions on Robotics, 36(2):328 – 347, 2019.

[7] Jizhai Cui, Tian-Yun Huang, Zhaochu Luo, Paolo Testa, Hongri Gu, Xiang-Zhong Chen, Bradley J Nelson, and Laura J Heyderman. Nanomagnetic encoding of shapemorphing micromachines. Nature, 575(7781):164–168, 2019.

[8] Antoine Cully, Jeff Clune, Danesh Tarapore, and Jean-Baptiste Mouret. Robots that can adapt like animals. Nature, 521(7553):503–507, 2015.

[9] Thomas George Thuruthel, Yasmin Ansari, Egidio Falotico, and Cecilia Laschi. Control strategies for soft robotic manipulators: A survey. Soft Robotics, 5(2):149–163, 2018.

[10] Zoubin Ghahramani. Probabilistic machine learning and artificial intelligence. Nature, 521(7553):452–459, 2015.

[11] Evin Gultepe, Jatinder S Randhawa, Sachin Kadam, Sumitaka Yamanaka, Florin M Selaru, Eun J Shin, Anthony N Kalloo, and David H Gracias. Biopsy with thermally-responsive untethered microtools. Advanced materials, 25(4):514–519, 2013.

[12] Sami Haddadin, Alessandro De Luca, and Alin Albuschäffer. Robot collisions: A survey on detection, isolation, and identification. IEEE Transactions on Robotics, 33(6):1292–1312, 2017.

[13] Zhong-Sheng Hou and Zhuo Wang. From model-based control to data-driven control: Survey, classification and perspective. Information Sciences, 235:3–35, 2013.

[14] Wenqi Hu, Guo Zhan Lum, Massimo Mastrangeli, and Metin Sitti. Small-scale soft-bodied robot with multimodal locomotion. Nature, 554(7690):81–85, 2018.

[15] Josie Hughes, Utku Culha, Fabio Giardina, Fabian Guenther, Andre Rosendo, and Fumiya Iida. Soft manipulators and grippers: a review. Frontiers in Robotics and AI, 3:69, 2016.

[16] Filip Ilievski, Aaron D Mazzeo, Robert F Shepherd, Xin Chen, and George M Whitesides. Soft robotics for chemists. Angewandte Chemie International Edition, 50 (8):1890–1895, 2011.

[17] Sangbae Kim, Cecilia Laschi, and Barry Trimmer. Soft robotics: a bioinspired evolution in robotics. Trends in Biotechnology, 31(5):287–294, 2013.

[18] Yoonho Kim, Hyunwoo Yuk, Ruike Zhao, Shawn A Chester, and Xuanhe Zhao. Printing ferromagnetic domains for untethered fast-transforming soft materials. Nature, 558(7709):274–279, 2018.

[19] Jens Kober, J Andrew Bagnell, and Jan Peters. Reinforcement learning in robotics: A survey. The International Journal of Robotics Research, 32(11):1238–1274, 2013.

[20] Sam Kriegman, Douglas Blackiston, Michael Levin, and Josh Bongard. A scalable pipeline for designing reconfigurable organisms. Proceedings of the National Academy of Sciences, 117(4):1853–1859, 2020.

[21] Frederick Largilliere, Valerian Verona, Eulalie Coevoet, Mario Sanz-Lopez, Jeremie Dequidt, and Christian Duriez. Real-time control of soft-robots using asynchronous finite element modeling. In 2015 IEEE International Conference on Robotics and Automation (ICRA), pages 2550–2555, 2015.

[22] Thomas Liao, Grant Wang, Brian Yang, Rene Lee, Kristofer Pister, Sergey Levine, and Roberto Calandra. Data-efficient learning of morphology and

controller for a microrobot. In 2019 IEEE International Conference on Robotics and Automation (ICRA), pages 2488–2494, 2019.

[23] Haojian Lu, Mei Zhang, Yuanyuan Yang, Qiang Huang, Toshio Fukuda, Zuankai Wang, and Yajing Shen. A bioinspired multilegged soft millirobot that functions in both dry and wet conditions. Nature Communications, 9 (1):1–7, 2018.

[24] Guo Zhan Lum, Zhou Ye, Xiaoguang Dong, Hamid Marvi, Onder Erin, Wenqi Hu, and Metin Sitti. Shape-programmable magnetic soft matter. Proceedings of the National Academy of Sciences, 113(41):E6007–E6015, 2016.

[25] Carmel Majidi. Soft robotics: a perspective—current trends and prospects for the future. Soft Robotics, 1(1): 5–11, 2014.

[26] Bradley J Nelson, Ioannis K Kaliakatsos, and Jake J Abbott. Microrobots for minimally invasive medicine. Annual Review of Biomedical Engineering, 12:55–85, 2010.

[27] Carl Edward Rasmussen. Gaussian processes in machine learning. In Summer School on Machine Learning, pages 63–71. Springer, Berlin, Heidelberg, 2003.

[28] Ziyu Ren, Wenqi Hu, Xiaoguang Dong, and Metin Sitti. Multi-functional soft-bodied jellyfish-like swimming. Nature Communications, 10(1):1–12, 2019.

[29] Federico Renda, Michele Giorelli, Marcello Calisti, Matteo Cianchetti, and Cecilia Laschi. Dynamic model of a multibending soft robot arm driven by cables. IEEE Transactions on Robotics, 30(5):1109–1122, 2014.

[30] Steven I Rich, Robert J Wood, and Carmel Majidi. Untethered soft robotics. Nature Electronics, 1(2):102, 2018.

[31] John Rieffel and Jean-Baptiste Mouret. Adaptive and resilient soft tensegrity robots. Soft Robotics, 5(3):318–329, 2018.

[32] Daniela Rus and Michael T Tolley. Design, fabrication and control of soft robots. Nature, 521(7553):467–475, 2015.

[33] Frank Sehnke, Christian Osendorfer, Thomas Rückstieß, Alex Graves, Jan Peters, and Jürgen Schmidhuber. Parameter-exploring policy gradients. Neural Networks, 23(4):551–559, 2010.

[34] Bobak Shahriari, Kevin Swersky, Ziyu Wang, Ryan P Adams, and Nando De Freitas. Taking the human out of the loop: A review of Bayesian optimization. Proceedings of the IEEE, 104(1):148–175, 2015. [35] Metin Sitti. Miniature soft robots—road to the clinic. Nature Reviews Materials, 3(6):74–75, 2018.

[36] Metin Sitti, Hakan Ceylan, Wenqi Hu, Joshua Giltinan, Mehmet Turan, Sehyuk Yim, and Eric D Diller. Biomedical applications of untethered mobile milli/microrobots. Proceedings of the IEEE, 103(2):205–224, 2015.

[37] Donghoon Son, Hunter Gilbert, and Metin Sitti. Magnetically actuated soft capsule endoscope for fine-needle biopsy. Soft Robotics, 7(1):10–21, 2019.

[38] Richard S Sutton and Andrew G Barto. Reinforcement Learning: An Introduction. MIT Press, 2018.

[39] Barry A. Trimmer and Huai Ti Lin. Bone-free: Soft mechanics for adaptive locomotion. Integrative and Comparative Biology, 54(6):1122–1135, 2014.

[40] Alexander von Rohr, Sebastian Trimpe, Alonso Marco, Peer Fischer, and Stefano Palagi. Gait learning for soft microrobots controlled by light fields. In 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pages 6199–6206, 2018.

[41] Conor Walsh. Human-in-the-loop development of soft wearable robots. Nature Reviews Materials, 3(6):78–80, 2018.

[42] Robert J Webster III and Bryan A Jones. Design and kinematic modeling of constant curvature continuum robots: A review. The International Journal of Robotics Research, 29(13):1661–1683, 2010.

[43] Tianqi Xu, Jiachen Zhang, Mohammad Salehizadeh, Onaizah Onaizah, and Eric Diller. Millimeter-scale flexible robots with programmable three-dimensional magnetization and motions. Science Robotics, 4(29):eaav4494, 2019.

[44] Brian Yang, Grant Wang, Roberto Calandra, Daniel Contreras, Sergey Levine, and Kristofer Pister. Learning flexible and reusable locomotion primitives for a microrobot. IEEE Robotics and Automation Letters, 3 (3):1904–1911, 2018.

Publication 2

Task Space Adaptation via The Learning of Gait Controllers of Magnetic Soft Millirobots

Abstract

Untethered small-scale soft robots have promising applications in minimally invasive surgery, targeted drug delivery, and bioengineering applications as they can directly and non-invasively access confined and hard-to-reach spaces in the human body. For such potential biomedical applications, the adaptivity of the robot control is essential to ensure the continuity of the operations, as task environment conditions show dynamic variations that can alter the robot motion and task performance. The applicability of the conventional modeling and control methods are further limited for soft robots at the small scale due to their kinematics with virtually infinite degrees of freedom, inherent stochastic variability during fabrication, and changing dynamics during real-world interactions. To address the controller adaptation challenge to dynamically changing task environments, we propose using a probabilistic learning approach for a millimeter-scale magnetic walking soft robot using Bayesian optimization (BO) and Gaussian processes (GPs). Our approach provides a data-efficient learning scheme by finding the gait controller parameters while optimizing the stride length of the walking soft millirobot using a small number of physical experiments. To demonstrate the controller adaptation, we test the walking gait of the robot on task environments with different surface adhesion and roughness, and medium viscosity, which aim to represent the possible conditions for future robotic tasks inside the human body. We further utilize the transfer of the learned GP parameters among different task spaces and robots and compare their efficacy on the improvement of data-efficient controller learning.

1. Introduction

Soft robots are composed of highly deformable soft materials exhibiting programmable shape change, mechanical compliance and high degrees of freedom, which are hard to achieve using rigid materials (Majidi 2014). The easier access to novel fabrication methods further allow the engineering of stimuli-responsive soft materials that enable new functionalities for soft robots in multiple length scales (Shen et al. 2020). Biology remains to be a source of inspiration for the design, control, and behavior of soft robots (Laschi et al. 2016), and provides templates for new application areas in multi-terrain locomotion (Calisti et al. 2017), adaptive manipulation (Hughes et al. 2016), sensing (Iida and Nurzaman 2016), human-assistive wearable systems (Walsh 2018), and biomedicine (Cianchetti et al. 2018). Soft robots also enable safe human-robot physical interaction due to their physical compliance and the mechanical dampening of excess forces (Polygerinos et al. 2017), which otherwise require additional computational effort in conventional robotic systems (Haddadin et al. 2017). Small-scale (i.e., $\leq$ 1 cm) untethered soft robots have further potential application areas in medicine owing to their ability to access enclosed small spaces non-invasively (Sitti 2018) and the embodiment of functionalized materials enabling targeted drug delivery, diagnostics, and surgery (Cianchetti et al. 2018).

Despite their exciting potential and new capabilities, soft robots face challenges that arise from the nature of their soft materials, such as having virtually infinite degrees of freedom, being prone to fabrication-dependent performance variabilities that cause more significant effects at

the smaller scale, and nonlinear material behavior (e.g., hysteresis, creep). Moreover, physical interactions of soft-bodied robots with their operation environment, such as solid or fluid operation medium, are very hard to model due to complex fluid-structure interactions, soft body dynamics, and contact mechanics. The combination of these aspects renders the application of conventional modeling and control methods challenging for soft robots, especially for untethered systems at the small scales (Rus and Tolley 2015). One of the most widely-used methods is the employment of the constant curvature (CC) models that utilize the well-established beam theories to model the kinematics and dynamics of axisymmetrically bending soft robotic systems (Webster III and Jones 2010; Della Santina et al. 2020). Alternatively, analytical approaches using Cosserat rod models (Renda et al. 2018) and geometrically exact models have been suggested for continuum robots (Grazioso et al. 2019). Simulation techniques build upon these modeling methods as in the finite element methods (FEM), which construct continuum robot structures using a chain of rigid elements connected with tunable spring-damper mechanisms (Chenevier et al. 2018; Goury and Duriez 2018). Numerical approaches using voxel-based representations (Hiller and Lipson 2014) and discrete differential geometries (DDG) (Huang et al. 2020) improve the computation time of soft robotic simulations at the expense of nonlinear dynamics precision. These models and simulation tools typically allow the implementation of static and dynamic controllers for continuum robots on a larger scale (Thuruthel et al., 2018). However, the physical application of these closed-loop controllers depends on the continuous sensing of body deformations from embedded sensors and

highly responsive actuators, and computationally-heavy model solutions, which are conditions that may not be met for untethered soft robots at the small scales (Rich et al. 2018). Therefore, the soft robotic platforms that successfully employ the analytical models at the small scales still depend on either open-loop (Lu et al. 2018; Wu et al. 2019; Gu et al. 2020; Ren et al. 2019) or manually applied (Kim et al. 2019) controllers. Especially for those robots targeting medical applications, their dynamically changing and deformable task environments, fabrication-based variations, and material degradation over prolonged use significantly alter their robotic function performances and pose challenges for the conventional control strategies (Sitti 2018). The combination of these challenges makes the machine learning-based, adaptive, and data-efficient control methods more desirable for untethered small-scale soft robots.

Data-driven machine learning methods may provide alternative solutions for the design and control of soft robots in the lack of existing analytical or numerical models that describe their underlying kinematics, dynamics, and functions (Chin et al., 2020). One common approach is to learn these models by gathering data from robot experiments and training a neural network (NN) architecture (Hyatt et al. 2019; Thuruthel et al. 2019; Bern et al. 2020) or using regression (Fang et al. 2019; Holsten et al. 2019). However, the need for data efficiency, i.e., the ability to learn from only a few experimental trials, presents a core challenge for such methods (Chatzilygeroudis et al. 2019). Conversely, Bayesian optimization (BO) (Ghahramani 2015; Shahriari et al. 2015) allows for the maximization of a performance function using a small number of physical experiments.

BO typically employs Gaussian processes (GPs) (Rasmussen and Williams 2006) as a probabilistic model of the latent objective function. While no explicit dynamics model is needed, GPs allow for incorporating information as probabilistic priors, thus reducing the experimental data requirements. There are emerging examples that demonstrate the application of this approach to optimize the locomotion performance of robots on different length scales (Calandra et al. 2016; Yang et al. 2018; Liao et al. 2019; Marco et al. 2020). Despite its potential, there are only a few examples that apply this method to address the controller challenge for untethered soft robots, such as in the gait exploration of a tensegrity system (Rieffel and Mouret 2018), and the optimization of an undulating motion of a microrobot (von Rohr et al. 2018).

For cases where the training and testing domains show differences in terms of features or data distribution, transfer learning (TL) methods may provide further improvements in data-efficient learning and adaptation to new test cases (Pan and Yang 2009). Within the BO applications that employ GPs, the prior knowledge can be transferred as GP priors (Raina et al. 2006) and hyperparameters (Perrone et al. 2019) from the trained domains to provide predictive information about the unknown features and distributions in the new test domains. In robotics, TL is typically employed as the transfer of the models of kinematics and dynamics between simulated and physical platforms of conventional rigid robotic systems, such as manipulators (Devin et al. 2017; Makondo et al. 2018), humanoids (Delhaisse et al. 2017), and quadrotor platforms (Helwa and Schoellig 2017). However, the application of TL on soft robotics systems is still in its early infancy (Schramm et al. 2020).

In our recent work in (Culha et al. 2020), we demonstrated the controller learning of walking soft millirobots using BO and GPs, and showed the improvement of learning efficiency in means of transferring prior mean information between robots as a TL application example. We followed the initial example by von Rohr et al. in (von Rohr et al. 2018), which designed a learning scheme by comparing different GP priors and BO settings on generating a semi-synthetic data set that represents the estimated gait controller space and used this estimation to optimize the one-dimensional crawling gait of a light-driven soft microrobot. In our work in (Culha et al. 2020), we adopted the magnetic soft millirobots from (Hu et al. 2018) that lacked sufficient predictive kinematic models, and was therefore controlled with an open-loop system whose multi-dimensional parameters were manually-tuned. We showed that these robots suffered performance inconsistencies due to the fabrication reproducibility issues, material degradation over prolonged experiments, and environmental disturbances, which limited the derivation of a deterministic kinematic model and the application of relevant model-based controllers. Therefore, we applied BO and GPs to directly learn the controller parameters while optimizing the stride length performance of these robots and employed TL methods to improve learning efficiency using a small number of physical experiments.

In this study, we extend our previous work in (Culha et al. 2020) and provide an in-depth analysis on using BO, GPs and TL methods to directly and efficiently learn the controller parameters of the magnetic soft millirobots' walking gait on task spaces emulating bio-medical application environments (Figure 1). First, we introduce our new

automated and closed-loop experimental platform that can run the robot learning experiments repeatedly and reliably to eliminate the influence of any human intervention, which caused further material degradation and consequent performance inconsistencies in (Hu et al. 2018; Culha et al. 2020). We start with using an exhaustive search on the two-dimensional (2D) gait controller parameter space of the millirobot and generating benchmark data sets that show the stride length performances of three different robots on three different walking surfaces. We use this benchmark data to learn the optimum gait controllers using BO and GPs, and then to compare the influence of four different TL methods on the improvement of learning efficiency. We choose the best performing TL method from these experiments and use it with the BO and GPs to learn the walking gait controller parameters on a wide range of task spaces. We test our robots on task spaces with different surface roughness and friction, and liquid medium viscosity to emulate the conditions inside the human body for future target operations. Our results reveal that the direct controller learning with BO and GPs allows adaptation to different task spaces for small-scale untethered soft robots that are prone to fabrication-, material-, and interaction-dependent performance variabilities. We also show that the effective use of TL methods improves this adaptation by exploring a larger set of successful walking gait controllers within a limited number of physical experiments despite the significantly changing task space conditions. The methodology we present in this study can be used for controlling future small-scale soft robot applications for medical operations that require a data-efficient controller learning system

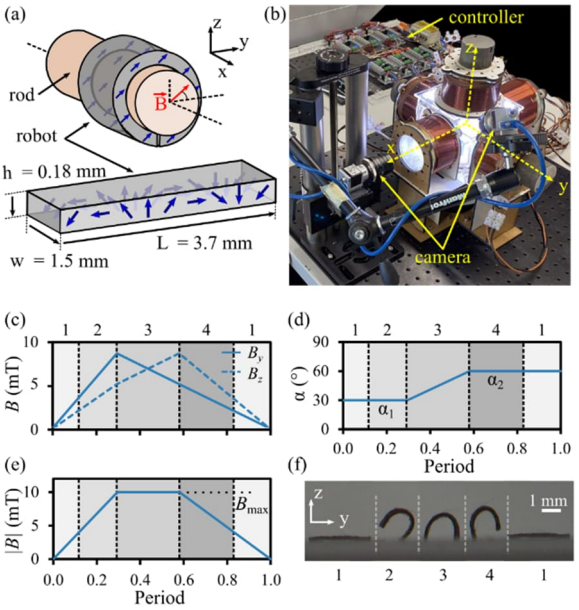


Figure 1. (a) Fabrication process of the magneto-elastomeric soft millirobot: The robot, which is composed of non-magnetized ferromagnetic microparticles homogeneously distributed inside a silicone elastomer sheet, is rolled around a cylindrical rod and magnetized with $|B| = 1.8T$ field (red arrow) with a $45°$ angle with respect to the y -axis (see (Hu et al. 2018) for details on the fabrication method). The unfolded robot maintains a periodic magnetization profile along its body (blue arrows). (b) Photo of the magnetic actuation and imaging experimental setup with 6 electromagnetic coils and two high-speed cameras that allow real-time evaluation of the robot's walking gait performance. Walking gait control parameters during a single period of a motion (a sample case of $1/f = 90ms$ and $f = 11Hz$ is normalized to 0 - 1 on the abscissa) (c) The magnetic field B is controlled on the y-z plane and shown with y and z components. (d) The magnetic field orientation changes from α_1 to α_2 and (e) magnitude reaches B_{max}. (f) In a single actuation period of the walking gait, the magnetic soft millirobot follows four consecutive gait states shown with the photos: (1) relaxed, (2) front-stance, (3) double-stance, (4) back-stance, and (1) relaxed again (shown with the numbers above and below the figures).

and quick adaptation to the changing task environments. The major contributions of our work are:

1. The demonstration of a data-driven optimization tool (i.e., BO) that can efficiently learn the gait controllers of a small-scale untethered robot whose performance is prone to fabrication-, material-, and physical interaction-based variabilities;
2. The testing of the walking gait on three different task spaces that emulate the dynamic environments inside the human body, and the adaptation of the robot controller parameters to these environments in a small number of experiments for future robot applications;
3. The implementation of an automated experimental platform that runs and evaluates the physical learning experiments repeatably and reliably without human intervention and simulated environments;
4. The comparison and evaluation of four different TL methods within the context of GP hyperparameters on the learning efficiency of the BO on the small-scale soft robots;
5. The generation of five benchmark datasets consisting of the exhaustively parsed controller parameter space involving 3750 different physical experiments for three different robots and three different walking surfaces that would allow further comparison between different optimization methods.

The organization of this paper is as follows. We describe the design of the robotic system, the walking gait, and the properties of the task environments in our experiments in Section 2. Section 3 describes the learning approach with the details on the BO, GP, and TL methods. In

Section 4, we present the experiments on generating the benchmark data sets, comparing the TL methods, and learning of the walking gait in different task environments. We discuss the experimental results and conclude our work in Section 5.

2. Experimental Robot System

2.1. Robot Design and Fabrication

We follow the methods and materials reported in (Hu et al. 2018) and our previous work (Culha et al. 2020), and fabricate three magnetic soft millirobots with a 1:1 body mass ratio of Ecoflex 00-10 (Smooth-On Inc.) and neodymium-iron-boron (NdFeB) ferromagnetic microparticles with around $5\mu m$ diameter (MQP-15-7, Magnequench). We place this pre-polymer mixture on a methyl methacrylate plate and cut the robots out of the cast using a high-resolution laser cutter (LPKF Protolaser U4) after the polymer is cured. Our robots have the final dimensions of length $L = 3.7mm$, width $w = 1.5mm$, and height $h = 185\mu m$ as shown in Figure 1(a). We separately fold the robots around a cylindrical rod with a circumference equal to L and magnetize them within a magnetic field with a magnitude of $1.8T$ and orientation of $45°$ measured counterclockwise from the y-axis. Once the robots are unfolded from the rod, the magnetic particles maintain their magnetization orientation, forming a circular profile along the longitudinal axis of the robot body. We use these robots (i.e., robots 1, 2, and 3), which have the same nominal material properties and dimensions, in the rest of our experiments.

2.2. Walking Gait Definition

The walking gait of our robot is composed of four consecutive quasi-static states that are inspired by the planar quadrupedal bounding (Alexander 1984) and a caterpillar inching motion (Trimmer and Lin 2014). These states are depicted as (1) relaxed, (2) front-stance, (3) double-stance, and (4) back-stance as shown in Figures 1(c-f). We control four parameters to generate the walking gait: the maximum magnetic field magnitude (B_{max}), the frequency of the actuation cycle (f), and two magnetic field orientation angles (α_1 and α_2) measured counterclockwise from the y-axis. The plots in Figures 1(c-e) show the change of the control parameters during a single period of the motion for $B_{max} = 10mT$, $f = 11Hz$, $\alpha_1 = 30°$ and $\alpha_2 = 60°$, which are adopted from the hand-tuned parameters reported in (Hu et al. 2018). At the beginning of a single gait period, the robot starts at a relaxed state for $0 \leq |B| \leq 4mT$. The robot tilts forward when $\alpha = \alpha_1$ and $|B|$ increases from $4mT$ to $B_{max} = 10mT$. While $|B|$ remains constant at B_{max}, the orientation of the magnetic field changes from α_1 to α_2 causing the robot to initially switch to the double-stance state and then to the back-stance state when $\alpha = \alpha_2$. Then, $|B|$ decreases while keeping the orientation of the magnetic field constant, and the robot gradually switches back to the relaxed state. For $|B| < 4mT$, the robot assumes the relaxed state, and a single period of walking actuation ends when $|B| = 0mT$. We reset B at the end of every gait cycle to avoid jerky motion when α changes from α_1 to α_2. In our experiments, the relaxed state is never skipped but its

duration changes according to f. The consecutive images from a single walking gait period are shown in Figure 1(f).

2.3. Actuation and Feedback Setup

We place our magnetized soft robot along the y-axis of the magnetic coil setup consisting of three orthogonal pairs of custom-made electromagnets (see Figure 1(b)) that can generate a 3D uniform magnetic field within a 4×4×4cm^3 workspace with a maximum value of $15mT$. We modulate the magnetic field on the y-z plane that coincides with the center of the test environment by controlling the electric currents running through the electromagnetic coils with six motor driver units (SyRen25) and an Arduino microcontroller, which runs at $1.2kHz$ operation frequency to compute the analog readings received from the current sensors, control the motor drivers and perform all necessary communications with the master PC for each single control cycle of 0.83 ms. We regularly calibrate the magnetic actuation matrix inside the workspace, i.e., the mapping between the applied electric current and the generated magnetic field, to maintain reliable and repeatable experiments.

We track the robot gait using two high-speed cameras (Basler aCa2040-90uc, shown in Figure 1(b)). The first camera running at 120 frames per second (fps) is placed orthogonal to the axis of robot motion (i.e., y-z plane of the controller). A tracking algorithm, whose pseudo-code is given in Appendix A, uses this camera to detect and evaluate the robot motion to identify if the robot is moving according to the walking gait definition given in Section 2.2. The second camera running at 60 fps has an isometric view of the test scene and is used to measure the distance

travelled by the robot following the perspective correction of the captured image. In every experiment, we calculate the stride length of the robot by tracking the average distance covered by its center of mass in 5 consecutive steps. At the end of every experiment, the robot is moved back to its original starting position automatically with the tracking and the actuation commands. See Extension 1 for the gait detection, position tracking, and repositioning for robot-3 walking on paper.

The learning process and image processing run on a master PC, and all the communication tasks between different elements of the robotic system (e.g., image capture and electric current control) are executed on Robot Operating System (ROS) architecture, which allows our system to be scalable for further extensions. The automated experimental platform implemented in our work allows the physical experiments to be executed with minimum human intervention, therefore reducing the human-based disturbances on the robot and the test surfaces. Without these interactions that can cause significant alterations on the soft millirobots, the physical learning experiments can be maintained repeatably and reliably.

2.4. Task Environments

In this study, we use a wide range of different task environments to test the efficacy of our adaptive learning strategy in comparison to the limited surface experiments in (Culha et al. 2020). Our goal is to emulate the in-air and liquid-immersed surface walking environments that a magnetic soft millirobot might experience during future medical operations inside the human body. To capture some of the characteristic properties of the target tissues and body fluids, we fabricate different task spaces and vary

their: (1) surface adhesion, (2) surface roughness, and (3) the liquid medium viscosity properties. For each of these properties, we experimentally identify the range of values that allow successful walking gaits and systematically test the robots in these specific ranges.

We fabricate a set of flat substrates with different surface adhesion strengths by using different materials. This set of substrates consists of paper, polystyrene (PS), and modified polydimethylsiloxane (PDMS). PDMS substrates are prepared by mixing Sylgard184 (Dow Corning) with its curing agent to a 10:1 ratio, degassing, and curing at $90°C$ for 1 hour. PDMS is modified by adding ethoxylated polyethylenimine (80% solution, Sigma Aldrich) prior to mixing and curing to increase its adhesive properties (Jeong et al. 2016). 0 (PDMS-0), 1 (PDMS-1), and 2 (PDMS-2) μL of polyethylenimine solution are added per 1g of silicone elastomer base. The adhesion between the walking surfaces and the robot are measured in a TA Instruments Discovery HR-2 rheometer with a custom-built adhesion setup (at loading and unloading speeds of $50 \mu m/s$). The resulting set of flat substrates resulted in walking surfaces with adhesion strengths between 1 to 10kPa as shown in Figure 2(a).

Walking surfaces with changing roughness are prepared by replicating the surface texture of different grits of sandpaper. First, we fabricate the negative molds of the original surfaces by pressing a glass plate with a layer of uncured vinylsiloxane polymer (Flexitime® medium flow, Heraeus Kulzer GmbH) onto the original surfaces. After curing the molds for 5 minutes at room temperature, we remove them and produce positive replicas of the original surfaces using clear casting epoxy (EpoxAcast™

690, Smooth-On Inc., 10:3 ratio by weight) onto the mold, with another glass plate pressed on top. At the end of 24 hours of curing time, we remove the positive replicas from the molds. According to surface profile examination of the replicated surfaces (Keyence VKX260K), surface roughness values R_q (root mean square height) changed between 7.5 and 77.2μm as shown in Figure 2(b).

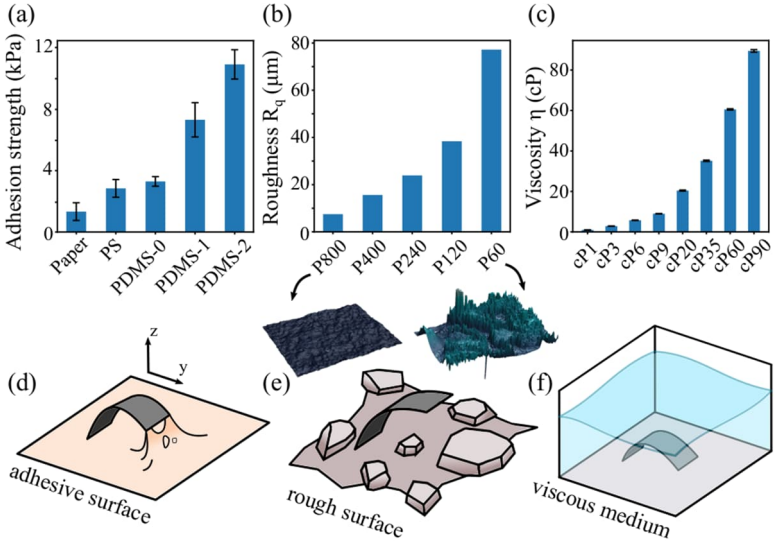


Figure 2. The type and range of task space properties investigated for the robot's walking gait. (a) Adhesion strength of different surfaces ranges between 1 to 10 kPa. (b) Roughness values of different surfaces that are named after the grit scale of their template sandpapers. The inset figures from the profilometer scans represent the two extremes of the roughness range (i.e., $R_{q,P800} = 7.5\mu m$ to $R_{q,P60} = 77.2\mu m$). (c) Viscosity values of the test fluids range between 1 to 90 cP where the robot is submerged while performing the walking gait. Fluids are named after their cP values. The sketches (d-f) aim to visualize the task space conditions during walking.

Last, we submerge the robots in different Newtonian fluids while walking on a flat surface to investigate the effect of bulk liquid medium viscosity on the walking performance. The Newtonian fluids are prepared by mixing different ratios of water and glycerol, and their viscosity is measured in a TA Instruments Discovery HR-2 rheometer. We analyze task environments with medium viscosity ranging from 1 to 90 cP as shown in Figure 2(c).

3. Learning Approach

We adapt the learning approach from our previous work (Culha et al. 2020) that aims to optimize the walking gait controller parameters to maximize the stride length S of the robot. Here we define the reward function as

$$S: \Theta \rightarrow \mathbb{R}, \tag{1}$$

which maps the parameter set $\theta = [B_{max}, f, \alpha_1, \alpha_2]$ to scalar reward values (i.e., the stride length performance of a robot). According to the definition of the reward function, we formulate the parameter learning as the (global) optimization problem

$$\theta^* = \underset{\theta \in \Theta}{\operatorname{argmax}} S(\theta), \tag{2}$$

where Θ denotes the complete search space, θ is the parameter set, and $S(\theta)$ is the average stride length performance of the robot for a given θ.

We define the range of the controller parameters based on the findings in (Hu et al. 2018) and the physical limitations of our magnetic actuation

setup. Accordingly, $B_{\max}$ defined between $5mT$ and $12mT$, and the walking frequency, f, ranges from $f_{min} = 0.5Hz$ to $f_{max} = 20Hz$. We limit α_1 and α_2 to $[0,70]°$ and $[30,90]°$, respectively and select values that satisfy $\alpha_2 > \alpha_1$ to generate the walking gait in Figures 1(c-f). We use a step size of $1mT$ for $|B|$, $5°$ for each α, and a variable step size of $0.5Hz$ for $f < 2Hz$ and $2Hz$ for $f \geq 2Hz$, which yield a total number of 15600 possible parameter sets in Θ.

3.1. Gaussian Processes

The magnetic soft millirobots in our paper did not have accurate models for kinematics or dynamics (i.e., we demonstrated the model inaccuracy of the original work of (Hu et al. 2018) in our previous study in (Culha et al. 2020)), therefore, it is necessary to approximate the reward function based on the data collected from physical experiments instead of model-based approaches. However, the physical data has inherent uncertainty due to the noise in the measurements and the variations during the experiments. To include these uncertainties in the model, overcome the sparsity in the data, and make probabilistic predictions at unobserved locations, we represent the reward function $S(\theta)$ using GPs following the study in (von Rohr et al. 2018):

$$S(\theta) \sim GP\big(\mu(\theta), k(\theta, \theta')\big), \tag{3}$$

where $\mu(\theta)$ is the prior mean and $k(\theta, \theta')$ is the kernel function defining the covariance between $S(\theta)$ and $S(\theta')$ for $\theta, \theta' \in \Theta$. However, as $S(\theta)$ can only be measured with noise, we define the observed stride length $\tilde{S}$ as

$$\tilde{S}(\theta) = S(\theta) + n, \tag{4}$$

Where n is zero-mean Gaussian noise with variance σ_n^2 for each measurement.

During one run of BO, the GP model is sequentially updated with $\tilde{S}(\theta)$ observed from experiments. We define one "learning run" as a run of BO until the desired stopping criterion is satisfied (e.g., a fixed number of experiments is reached).

From the experimental data $D = \{\theta_i, \tilde{S}(\theta_i)\}_{i=1}^{N}$, where N denotes the number of experiments in D, the stride length of the robot for an unobserved θ can be predicted using the posterior mean and variance as

$$\mu_{post}(\theta) = \mu(\theta) + k^T(\theta)K^{-1}y, \tag{5}$$

$$\sigma_{post}^2(\theta) = k(\theta,\theta) - k^T(\theta)K^{-1}k(\theta), \tag{6}$$

$$S_{post}(\theta)|D \sim \mathcal{N}\left(\mu_{post}(\theta), \sigma_{post}^2(\theta)\right), \tag{7}$$

where $k(\theta)$, $y \in \mathbb{R}^N$ with $[k(\theta)]_i = k(\theta,\theta_i)$, $y_i = \tilde{S}(\theta_i) - \mu(\theta_i)$, and $K \in \mathbb{R}^{N\times N}$ with $K_{i,j} = k(\theta_i,\theta_j) + \delta_{i,j}\sigma_n^2$, where $\delta_{i,j}$ is the Kronecker delta and σ_n^2 is the noise in the collected data set.

We select the squared exponential as the kernel function in the GPs, which is defined in (Duvenaud et al. 2011) for multi-dimensional cases as

$$k_{SE}(\theta, \theta') = \sigma_f^2 \exp\left(-\sum_{d=1}^{d_c} \frac{\left(\theta^d - \theta'^d\right)^2}{2l_c^{d\,2}}\right), \quad (8)$$

where $l_c \in \mathbb{R}^{d_c}$ is the length scales that defines the rate of variation in the modeled function for each dimension of the parameter space. Long length scales are used to model slowly-varying functions and short length scales are used to model quickly-varying functions. The signal variance σ_f^2 describes the width of distribution, e.g., high σ_f^2 means higher uncertainty in the predictions of the unobserved θ. We implement the GP model in our experiments using the libraries provided by GPy (GPy 2012).

3.2. Bayesian Optimization

We use BO to select the parameter set θ_{next} to be tested in the next step of the learning run using the acquisition function $\alpha_{acq}(\theta)$ as

$$\theta_{next} = \underset{\theta \in \Theta}{\operatorname{argmax}}\, \alpha_{acq}(\theta), \quad (9)$$

In this study, we choose the expected improvement (EI) as the acquisition function $\alpha_{acq}(\theta)$ due to its better performance compared to its alternatives as demonstrated in (von Rohr et al. 2018). EI seeks the parameter set for the next step where the expected improvement in reward function is the highest compared to the previously collected data and is defined in (Jones et al. 1998) as

$$\alpha_{acq}(\theta) = \mathbb{E}\left[\max\left(0, \left(S(\theta), \tilde{S}(\theta)\right)\right)\right], \quad (10)$$

where $\tilde{S}(\theta^*)$ is the highest reward function value collected so far. Analytical solution of Equation (10) is given in (Brochu et al. 2010) as

$$\alpha_{acq}(\theta) = \left(\mu(\theta) - \tilde{S}(\theta^*) - \xi\right)\Phi(Z) + \sigma(\theta)\phi(Z), \qquad (11)$$

where Φ and ϕ are the Gaussian cumulative density and probability density functions, respectively. The term Z is described as $Z = Z(\theta) = (\mu(\theta) - \tilde{S}(\theta^*) - \xi)/\sigma(\theta)$, with $\mu(\theta)$ and $\sigma(\theta)$ computed from Equations (5) and (6). The two terms in Equation (11) define the exploitation and the exploration weights of the BO, respectively. The balance between these two terms is controlled by the hyperparameter ξ. As ξ gets higher, BO focuses more on exploration and seeks the next parameter set in regions with high prediction uncertainty. On the contrary, BO focuses more on exploitation and selects the next parameter set within a close range to already explored regions. As the goal of our study is to adapt to task spaces by increasing the likelihood of finding more controller parameter sets that yield successful walking gaits under uncertainty, we choose $\xi = 0.1$ to increase the exploration tendency of the BO in our experiments.

3.3. Transfer Learning

In this study, we compare four different methods of TL on our walking gait experiments: (1) transfer of all GP hyperparameters: σ_n^2, σ_f^2, and length scales l_c for each dimension of the parameter space $\mathbb{R}^{d_c}$, i.e., l_c^u for $u \in \{B_{max}, f, \alpha_1, \alpha_2\}$, (2) transfer of only the length scales l_c, (3) transfer of prior mean information $\mu(\theta)$, and (4) the hybrid combination of length scales l_c and $\mu(\theta)$.

3.3.1. Transfer of GP Hyperparameters:

The choice of the types and values of GP hyperparameters influence the regression of the GP process (Chen and Wang 2018) and their transfer from prior models can change the dynamics of the learning process (Patacchiola et al. 2020; Wang et al. 2020). The hyperparameters we choose to investigate as a part of the GPs in this study can be listed as the noise in the collected data σ_n^2, the signal variance σ_f^2, and the length scales l_c. We start the BO learning by initializing the σ_n^2 to the maximum variance found in the repeated experimental results in our previous work in (Culha et al. 2020), and setting the signal variance σ_f^2 to square of half of the body length of the robot (i.e., $L = 3.7mm$) so that the highest possible reward value remained inside the 95% confidence interval of the prior. We also set the length scale values l_c^u to one-fourth of the total range of each corresponding parameter. After starting the BO runs with these initial values, we use the log marginal likelihood estimation derived from (Rasmussen and Williams 2006) as

$$\begin{aligned}\log p\left(\tilde{S}\middle|\theta, l_c^u, \sigma_n^2, \sigma_f^2\right) \\ = -\frac{1}{2}\left(\tilde{S}-\mu(\theta)\right)^T K^{-1}\left(\tilde{S}-\mu(\theta)\right) \\ -\frac{1}{2}\log|K| - \frac{N}{2}\log 2\pi,\end{aligned} \tag{12}$$

to simultaneously optimize the GP hyperparameters based on the collected data during the learning runs. We use these optimized estimation of the selected hyperparameters as the one of the TL methods in the following experiments in Section 4.3.

3.3.2. Transfer of Mean Prior Information:

In addition to the kernel, the prior mean $\mu(\theta)$ must be chosen at the beginning of a BO run as well. Often, constant zero mean (i.e., $\mu = 0$) is the default choice as an uninformed prior mean function for maximization problems (Chen and Wang 2018). For the millirobot learning problem herein, we investigate the transfer of information from previous learning runs by setting the prior mean to the posterior mean of a previously trained GP model, such as from a different robot. In this way, we can approximately transfer the topology of the target function between different test scenarios, which is reasonable as long as the differences between the robots and the environments do not significantly alter the function shape.

3.3.3. Hybrid Transfer:

Previous methods can be combined and both the optimized estimation of the GP hyperparameters and the mean prior information can be transferred between the BO experiments. In this study, we also investigate the combination of the estimated length scales l_c and the prior mean information $\mu(\theta)$ and their transfer between the test cases in Section 4.3.

4. Experimental Results

Our study aims to use BO and GPs to demonstrate adaptation to different task spaces while experimentally optimizing the stride length of the soft millirobots whose walking performances are prone to fabrication-, material-, and interaction-based reproducibility issues that cannot be

successfully predicted with kinematic models. In that sense, we focus more on exploring a variety of walking patterns under changing task space conditions rather than continuously optimizing a specific walking gait performance. Accordingly, we design the experiments to highlight the influence of BO and GPs and TL methods on increasing the average performance of finding successful walking gaits, i.e., gaits strictly following the consecutive states described in Section 2.2 that also yield sub-optimal stride length performances, during the limited number of learning runs, instead of only finding the optimum controller parameters.

We begin with using an exhaustive search approach to generate benchmark data sets for the walking gaits on five different test scenarios using our millirobots in Section 4.1. Here, we limit the controller parameter space to two dimensions and only explore the α_1 and α_2 parameters while experimenting with three robots on a flat paper surface and with one robot (i.e., robot 3) on two additional different walking surfaces. The results of the exhaustive search show the overall structure of the walking gait function based on the range of the two controller parameter values. This statistical information serves as benchmark data to compare the learning efficacy of the BO and different TL methods in Section 4.2 and Section 4.3. We choose the most effective TL method guided by the experiments on the benchmark data sets and apply it to a new set of task adaptation experiments in Section 4.4. Here we use a single robot on three different task spaces with a wide range of changing surface adhesion and roughness, and medium viscosity. In these experiments, we expand the controller space back to four dimensions

(i.e., B_{max}, f, α_1 and α_2), and optimize the walking gait with the BO and the chosen TL method within a limited number of learning runs.

4.1. Generation of the Walking Gait Benchmark Data Sets

In our previous work in (Culha et al. 2020), we observed that the soft millirobots we adopted from (Hu et al. 2018) experienced additional material degradation over long repeated experiments that altered their gait performances. While we investigate the influence of BO and TL methods on the improvement of learning efficiency in Section 4.2 and Section 4.3, we want to minimize this material degradation effect on the walking gait. That is why, here we use an exhaustive search approach and generate five different benchmark data sets that cover the walking gait function space necessary for the BO and TL methods investigations. To this end, we test three different robots (i.e., robot 1, 2, and 3) on a flat paper surface and a single robot (robot 3) on two additional surfaces: PDMS-0 and P800-grit sandpaper replica.

To explore the walking gait function space on these five test cases, we constrain the controller space into two dimensions by using a constant $B_{max} = 10mT$ and $f = 1Hz$, and changing the α_1 and α_2. We choose values from the set $[0,70]°$ for α_1 and $[30,90]°$ for α_2 with a step size of $5°$ that meet the condition $\alpha_2 > \alpha_1$, which consequently generate 150 different controller parameter pair values. For each of these pairs, we repeat the experiments 5 times; hence generating 750 physical experiments for each test case and report the results in Figures 3(a-e). The constrained dimensions reduce the necessary experiments from 390000 (i.e., 5 repetitions for each test case using 15600 parameter value sets) to

3750 for the exhaustive search, and significantly avoid the possible material-degradation over prolonged experiments.

In these experiments the robots do not necessarily follow the gait definition in Section 2.2, therefore we describe these resulting values as 'displacement measurements', S'. To filter out the non-walking gaits of the robot from this data set, which can be seen at the beginning of Extension 2, 3, and 4 for different task spaces, we use our gait tracking feedback system defined in Section 2.3 and evaluate every test result to penalize the α controller pairs that do not generate the desired walking gait. Accordingly, we obtain the average stride length performances, S, of the successful gaits in Figures 3(f-j). These results show that the stride length performances are limited up to $S \approx 2.4mm$ for the successful walking gaits, and the penalized motions with higher displacements ($S' > S$) do not comply with the walking gait definition. The statistical information collected from the physical experiments in these five test cases (shown in Figures 3(f-j)) constitutes our benchmark data sets that we use in the following BO and TL investigations. The optimum controller parameter sets found by the exhaustive search in the 2D function space are reported in Table 1. We use the mean and standard deviation values of these exhaustive search results to sample the stride length of the robots for the given α controller parameter sets while comparing the performances of different TL methods. The benchmark data is available online and can be accessed from this repository.

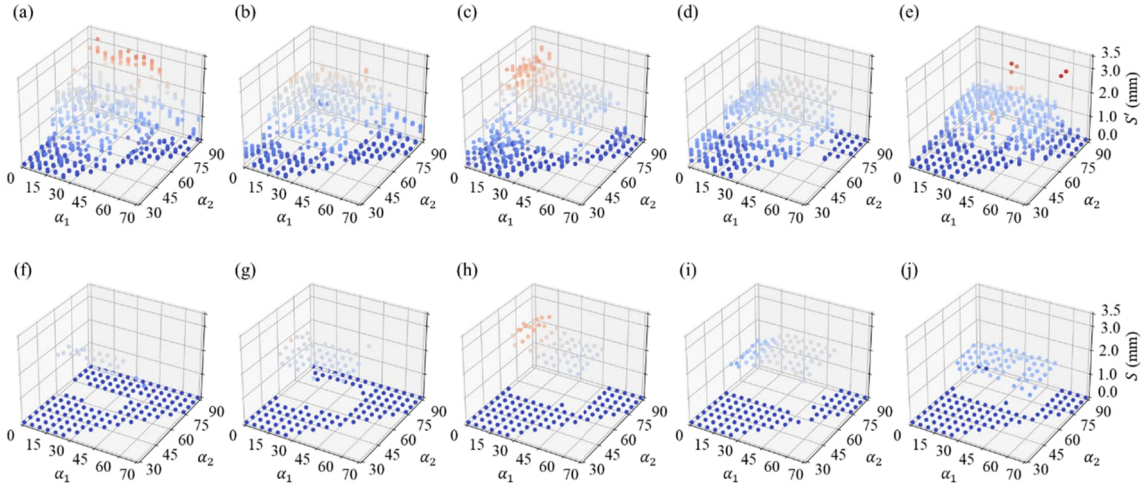


Figure 3. Experimental displacement measurements S' in five test case scenarios for (a) robot-1 on paper, (b) robot-2 on paper, (c) robot-3 on paper, and robot-3 on (d) PDMS-0, and (e) P800 surfaces. Each figure represents 5 repetition for 150 different controller value pairs, yielding 750 physical experiments. The displacement measurements are filtered out with the tracking algorithm to identify the desired walking gaits and the corresponding average stride length S values for the benchmark data sets (f) robot-1 on paper, (g) robot-2 on paper, (h) robot-3 on paper, and robot-3 on (i) PDMS-0, and (j) P800 surfaces.

Table 1 Best performing α controller parameter sets found by the exhaustive search and the corresponding stride length results $\overline{S_{exh}}$

Test Case	Controller Parameters*		Stride Length $\overline{S_{exh}}$ (mm) (avg ± std)
	$\alpha_1(°)$	$\alpha_2(°)$	
Robot-1 on Paper	15.0	70.0	1.75 ± 0.04
Robot-2 on Paper	0.0	70.0	1.91 ± 0.12
Robot-3 on Paper	5.0	70.0	2.43 ± 0.12
Robot-3 on PDMS-0	30.0	85.0	1.83 ± 0.09
Robot-3 on P800	35.0	70.0	1.77 ± 1.37

* $B_{max} = 10mT$ and $f = 1Hz$ in all cases.

4.2. Learning the Walking Gait with the 'Standard' BO

We initially test the walking gait learning with a BO approach on the benchmark data sets, where the prior mean information is set to zero (i.e., $\mu(\theta) = 0$) and the four hyperparameters are set to initial values described in Section 3.3.1 without inheriting any other prior information. We utilize this approach, which we refer to as 'standard BO' in the rest of this study, for the two controller parameters α_1 and α_2 and apply it separately to the benchmark data sets for the five different test cases (i.e., robot 1 to 3 on a flat paper surface, and robot 3 on the PDMS-0 and P800 surfaces). To be able to directly use the mean and standard deviation information from the exhaustive search results, we configure our BO to explore the same discrete controller parameter set space used in Section 4.1. We perform 100 independent learning runs with each involving 100 iteration steps for the five test cases. One iteration step of a learning run involves three steps:

1. BO selects a new parameter set θ that maximizes the acquisition function based on the GP model,
2. For the selected controller parameter pair, the corresponding stride length performance is sampled from the normal distribution defined by the mean and standard deviation values found in the relevant benchmark data set,
3. The learning system updates the GP model using this sampled data and prepares for the next iteration step of the learning run.

We report the median of the learning results with the upper and lower quartiles in Figure 4. These values represent the normalized gait performance $\hat{S} = \tilde{S}/\overline{S_{exh}}$, where $\overline{S_{exh}}$ is the mean of the stride length performances of the robots for the best α controller parameters reported in Table 1. For the first four of the five test cases, the standard BO approach finds the optimum gait controller parameters in less than an average of 25 iterations out of 100 independent learning runs. For the robot 3 walking on the P800 surface, the BO finds approximately

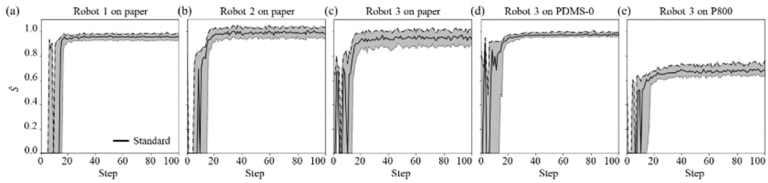


Figure 4. Performance of the standard BO for (a) Robot-1 on paper, (b) Robot-2 on paper, (c) Robot-3 on paper, (d) Robot-3 on PDMS-0, and (e) Robot-3 on P800. The stride length performances are normalized with respect to $\overline{S_{exh}}$ given in Table 1 for easier comparison between tests. Each figure shows the statistical results as median, and upper and lower quartiles from 100 independent BO runs with each consisting of 100 iterations.

74% ($\hat{S} \approx 0.74$) of the optimum stride length performance as shown in Figure 4(e). The results show that the standard BO starts the learning without any prior information and occasionally finds controller parameter sets that yield $\hat{S} = 0$ in the first 15 iteration steps. The non-monotonic optimization results in these initial steps is an expected result due to the statistical exploration nature of our BO approach. After 25 iteration steps the BO maintains the exploration in the close vicinity of the optimum controller parameters it finds so far for all the test cases (i.e., $\hat{S} \approx 1$ for robot 1-3 on paper and robot 3 on PDMS-0, and $\hat{S} \approx 0.74$ for robot 3 on P800). The variation of the generated walking gait performances is bounded by the standard deviations reported in Table 1. The exact normalized performance results for the standard BO approach can also be seen in Table 2.

4.3. Comparison of Transfer Learning Methods on the Benchmark Data Sets

In this study, we extend our previous investigation on the role of TL in learning efficiency (Culha et al. 2020) and compare four different methods while optimizing the gait controllers of our soft millirobots. Similar to Section 4.2, we apply our BO learning to the benchmark data sets generated in Section 4.1, where the controller parameter space is limited to two dimensions with α_1 and α_2. However, unlike the 'standard' BO, here we initialize the learning runs of the robots in all test cases with different types of prior information learned from the robot 3 walking on a flat paper surface. In the remaining parts of this section, we refer to this prior information source as “the source robot”. The four different TL

methods we compare are: (1) transfer of all GP hyperparameters (HP-4 Transfer), (2) transfer of only the length scales (HP-l_c Transfer), (3) transfer of prior mean information (Mean Transfer) and (4) the hybrid combination of length scales with mean information (Hybrid Transfer). All the experiments shown in this section are the result of 100 independent learning runs with each having 100 iterations for all the test cases. The values for the hyperparameters used in each of the TL methods can be found in Appendix B.

HP-4 Transfer: We initially transfer all of the four GP hyperparameters (i.e., noise in the collected data σ_n^2, signal variance σ_f^2, and length scales l_c^u for $u \in \{\alpha_1, \alpha_2\}$) that are optimized with the log marginal likelihood estimation in Equation 12 from the learning runs of the source robot to all other five test cases. Here, we initialize the BOs with this prior information and start the learning experiments. The normalized stride length performances, $\hat{S}$, of the BO learning with this TL method (depicted as 'HP-4') are shown in comparison with the 'standard' BO approach in the first column of Figure 5. These results show that the transfer of all of the four hyperparameters improves the learning performance in terms of decreasing the number of iteration steps to find the optimum controller parameters within these experiments. For robot 1 on paper and robot 3 on P800, the BO manages to find the performances achieved by the 'standard' approach in less than half of the iteration steps, which are shown Figures 5(a) and 5(q), respectively. For robot 2 and 3 on paper, and robot 3 on PDMS-0 (as seen in Figures 5(e,i,m)), the BO with this TL method does not reach the performances previously achieved by the 'standard' BO.

After finding the optimum performances, the BO maintains the exploration in their close vicinity for all the test cases.

HP-l_c Transfer: Second, we investigate the transfer of only the two length scale hyperparameters l_c^u for $u \in \{\alpha_1, \alpha_2\}$ optimized by the source robot. We run the BO learning experiments after all the length scale parameters are initialized with the transferred l_c values. The results in the second column of Figure 5 show that this TL method (depicted as 'HP-l_c') worsens the learning performance as it increases the number of iteration steps for the BO to find the optimum performance parameters for all the cases. However, in comparison to the previous 'HP-4' method, the transfer of only the length scale hyperparameters allows the BO to explore the gait performances achieved by the standard BO approach.

Mean Transfer: As the third method, we transfer the posterior mean information, $\mu(\theta)$, from the source robot as the prior information for the robots in the test cases. Here, we set all of the other GP hyperparameters to the initial values as described in Section 3.3.1 and restart the BO learning experiments. The results in the third column of Figure 5 show that when the BO learning starts with the prior mean information (depicted as 'Mean'), it achieves the average gait performances found by the standard BO ($\hat{S} \approx 1$) in fewer iteration steps for the first four test cases. While the median of the achieved $\hat{S}$ semains close to the standard BO's results for the robot 3 on P800, this method also allows the exploration of the controller parameter sets those yield performances close to the optimum results from the exhaustive search as shown in Figure 5(s).

Hybrid Transfer: Finally, we adopt a hybrid approach and transfer the posterior mean information, $\mu(\theta)$, together with the two length scale hyperparameters l_c^u for $u \in \{\alpha_1, \alpha_2\}$, whose results are shown in last column of Figure 5 (depicted as 'Hybrid'). The hybrid transfer improves the learning run performance by decreasing the number of iteration steps to find the optimum performing parameter sets found by the standard BO. Similar to the results in Figure 5(s), the hybrid transfer also allows robot 3 on P800 to investigate parameter regions that yield performances close to the optimum results from the exhaustive search.

The comparative performance results of the standard BO and the four TL methods for each of the five test cases are reported in Figure 6 and Table 2. Due to the statistical and explorative nature of the BO, the stride length performances do not monotonically increase at every consecutive iteration step (as also visible in Figure 5). That is why, to provide a clear comparison between our methods in Figure 6, we identify the iteration steps that show the 'best so far' performance during the learning run of each approach. The first row of Figures 6(a-e) compares these methods in terms of the normalized error of the achieved stride length performances during the BO learning runs ($\varepsilon = 1 - \hat{S}$). For the first four cases, it can be seen that except for the 'HP-4' method, the other three transfer methods and the standard BO manage to consistently explore the optimum stride length performances ($\varepsilon \approx 0$) as shown in Figures 6(a-d). The 'HP-4' method does not allow the BO to find the optimum controller parameters and the results remain approximately 20% below these optimum values for three of the test cases, which are also shown in Figures 6(b-d). The

'HP-l_c' method shows similar performances with the standard BO in terms of the initial and the final performance error. Both the 'Mean' and 'Hybrid' transfer methods allow the BO to generate comparable final performances, while the mean transfer method typically starts with significantly lower error compared to the hybrid approach. For the last test case (i.e., robot 3 on P800, Figure 6(e)) we can see that the standard BO approach only manages to explore gaits that are 74% ($\hat{S} \approx 0.74$) of the optimum gait performance. Every TL method increases the performance yield of the BO learning, with the 'Mean' transfer outperforming all other methods by exploring gaits that have $\hat{S} \approx 0.91$, or $\varepsilon \approx 0.09$, on the surface of P800-grit sandpaper in terms of the final performance. The comparison between the standard BO and the TL methods in terms of achieved $\hat{S}$ performances are given in detail in Table 2.

The second row of Figures 6(f-j) represents the learning efficiency performance comparison in terms of the iteration steps needed to explore the best performance by the standard BO and to achieve standard BO level performance by the four TL methods. We describe this exploration performance with 'convergence steps', which is calculated by finding the performance value that stays within the 5% band of the averaged remaining steps. Normally, a monotonic convergence is not expected from the statistical and explorative BO learning. However, viewing the learning runs with the 'best-so-far' evaluation method allows us to represent the required iteration steps to achieve comparable performance,

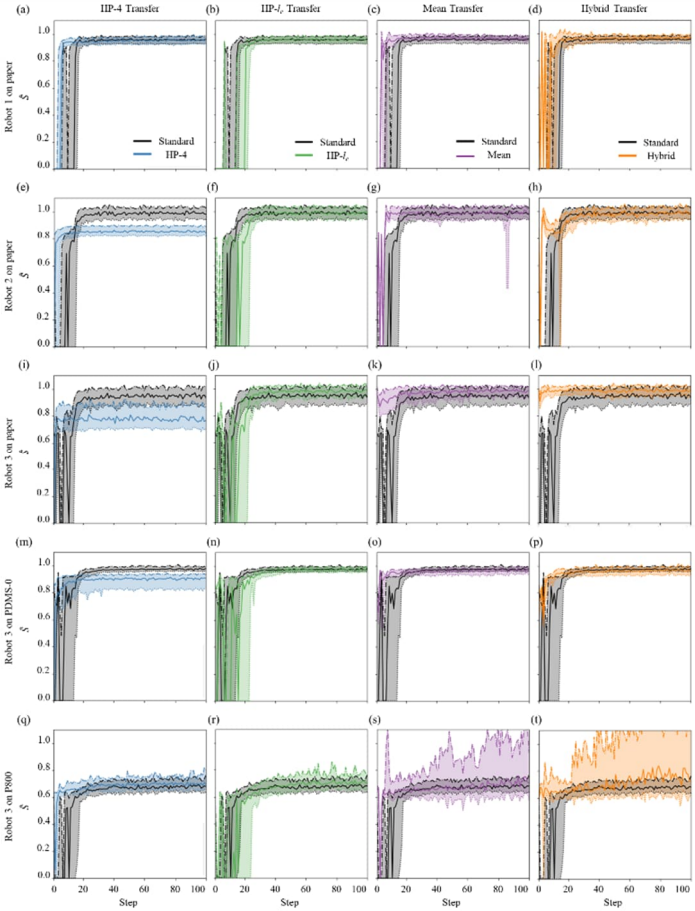


Figure 5. Comparison of each TL method are shown in separate columns in terms of normalized stride length performance $\hat{S}$ with the standard BO approach on five test cases (1st row) Robot-1 on paper, (2nd row) Robot-2 on paper, (3rd row) Robot-3 on paper, (4th row) Robot-3 on PDMS-0, and (5th row) Robot-3 on P800. The gray scale plots represent the standard-BO replicated from Figure 4. Straight lines represent the median and the shaded regions between dashed lines show the upper and lower quartiles. Each figure shows the statistical results from 100 independent BO runs with each consisting of 100 iterations.

and to capture the relative data-efficiency of the TL methods. Accordingly, we can see that the 'HP-4' method fails to achieve standard BO level performance for the three cases (Figures 6(g-i)) as its convergence steps are equal to the number of iteration steps in the experiments. The 'HP-l_c' method shows comparable results with the standard BO for all the cases (Figures 6(f-j)) in terms of the convergence steps. In comparison, both of the 'Mean' and 'Hybrid' TL methods attain to the standard BO level performance significantly faster (i.e., fewer convergence steps) in all the test cases. The details of the convergence steps are given in the last column of Table 2.

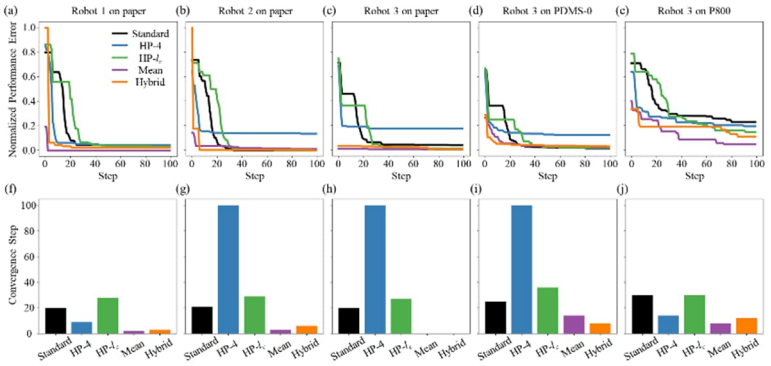


Figure 6. Performance comparison of standard BO and four TL methods on five test cases. (a-e) Normalized performance error ε with respect to the optimum gait performances from the exhaustive search results (Table 1). Each figure shows the 'best-so-far' performance results over 100 iterations. (f-j) Comparison of resulting convergence step of standard BO and TL methods. The convergence steps of the TL methods are calculated as the iteration step achieving the performance level of standard BO approach. Maximum iteration step of 100 is reported for the cases that cannot reach standard BO level.

Table 2 Comparison of the transfer learning methods

Robot Number	Test Surface	Type	Performance		Convergence Step
			* (%)	** (%)	
1	Paper	Standard	95.83	↔	20
		HP-4	95.35	-0.51	9
		HP-l_c	96.17	+0.35	28
		Mean	97.20	+1.42	2
		Hybrid	97.34	+1.57	3
2	Paper	Standard	99.63	↔	21
		HP-4	86.32	-13.36	100
		HP-l_c	99.62	-0.01	29
		Mean	98.44	-1.19	3
		Hybrid	99.40	-0.23	6
3	Paper	Standard	95.600	↔	20
		HP-4	81.25	-15.02	100
		HP-l_c	98.03	+2.54	27
		Mean	98.81	+3.36	0
		Hybrid	98.23	+2.75	0
	PDMS-0	Standard	98.38	↔	25
		HP-4	86.90	-11.68	100
		HP-l_c	97.74	-0.66	36
		Mean	96.54	-1.87	14
		Hybrid	96.68	-1.73	8
	P800	Standard	73.84	↔	30
		HP-4	77.61	+5.11	14
		HP-l_c	80.00	+8.33	30
		Mean	91.12	+23.40	8
		Hybrid	84.47	+14.40	12

* Relative performance with respect to the optimum exhaustive search results.

** Relative performance with respect to the standard BO learning approach.

As 'Mean' TL method outperforms the standard BO and other three TL methods by finding better performing parameter sets in less number of iterations, we select it to use for task space adaptation experiments in Section 4.4.

4.4. Adaptation to Task Spaces

The task environment is more susceptible to dynamic changes than the robot morphology, especially for medical operations inside the human body. Therefore, a quick adaptation of the robot controller is important to maintain successful robot task handling. In the following experiments, we investigate the learning efficiency of our BO approach while focusing on the three physical properties that may dynamically change during the walking task of our soft millirobots in future in vivo operations, which are (1) surface adhesion, (2) surface roughness, and (3) medium viscosity. Here, we expand the controller parameter space exploration back to four dimensions by including the magnetic field magnitude B_{max} and the actuation frequency f. To efficiently learn the controller parameters in this higher dimensional search space, we utilize the prior mean transfer application (i.e., 'Mean' TL), which is shown to be the best performing TL method for our experimental scenario in Section 4.3. However, as the source robot in previous experiments explored only two dimensions of the controller space, we generate the posterior mean $\mu_{post}(\theta)$ required for the following experiments by performing a new set of physical experiments with the robot 3 on a flat paper surface. For these experiments, we run the standard BO for 156 iteration steps (i.e., 156 different controller parameter sets), which corresponds to 1% of the

complete controller parameter space with given step sizes in Section 3. We use the posterior mean optimized at the end of these learning runs as the prior mean information for all the task space experiments. We only test robot 3 in the following task space adaptation experiments. Again, we compare the learning efficiency of the standard BO with the prior mean transfer method on all the task spaces defined in Section 2.4. Here, the objective of these learning runs is to adapt to dynamic task spaces and learn the optimized controller parameters in as few experiments as possible especially for future medical operations. To find the number of learning steps sufficient enough for BO to find the desired walking gaits, we used the results in Figures 4, 5, and 6, which involve approximately 250000 data points. These results show that the BO finds the optimized controller parameters that generate desired walking gaits consistently in less than 20 steps for different robots and walking surfaces. Therefore, we limit the number of steps of a learning run to 20 experiments (i.e., iteration steps), and perform three independent learning runs with the same initial conditions, yielding 60 experiments in total. One step of the learning run involves five steps:

1. BO selects a new parameter set θ that maximizes the acquisition function based on the GP model,
2. The microcontroller initiates the physical experiment and regulates the magnetic field based on the selected θ,
3. The cameras record the robot's motion and measure the average stride length performance $\tilde{S}$ after running the gait tracking system,

4. The learning system updates the GP model using the newly collected data from the experiment,
5. The robot returns to its initial position for the next experiment.

Surface Adhesion: We initially test the robot on five surfaces with different adhesion strengths reported in Figure 2(a). Figures 7(a-d) shows the walking gait performances during the three independent learning runs on the two ends of the adhesion range: paper (1.34 kPa) and PDMS-2 (11.02 kPa). The figures on the left column (a,c) show the learning runs with the standard BO approach, and the figures on the right column (b,d) show the learning runs with the prior mean transfer method. The difference between these figures shows that the TL method improves the learning runs by finding more of the controller parameters that yield positive walking gait performances. Additionally, the BO with the TL manages to explore these parameters in the earlier steps of the learning runs compared to a standard BO approach. We represent the general influence of the mean transfer method on all the adhesion surfaces with the standard interquartile range (IQR) method in Figure 7(e). In this figure, the horizontal lines represent the median of the generated $\tilde{S}$ in 60 experiments for each surface. These lines are surrounded by boxes that show the upper and lower quartiles. The error bars show the extremes in terms of the highest and lowest $\tilde{S}$ performances and the circles represent the outlier $\tilde{S}$ performances. The results for the learning with standard BO and BO with the chosen TL method are given side by side for each test surface in Figure 7(e) and the exact values can be seen in Table 3. For example, we can see that the learning runs with the standard BO on PS,

PDMS-0, and PDMS-1 surfaces typically explore controller parameters that do not yield successful walking gaits. In comparison, the BO with the TL method explores numerous successful controller parameters on these surfaces. The increased median lines in Figure 7(e) provide evidence that the prior mean transfer improves the learning of the BO in terms of increasing the number of successful walking gait generating controller parameters that are explored in a limited number of experiments. See Extension 2 for a comparison between the walking gaits on paper and PDMS-2, and Extension 6 for the details of the independent learning runs for all the surface adhesion experiments.

Table 3 Comparison of the learning performances for changing surface adhesion

Test Surface	**Type**	**Performance (mm)**	
		Median	**IQR**
Paper	Standard GP-BO	0.00	1.59
	Mean Transfer	1.60	0.46
PS	Standard GP-BO	0.00	0.00
	Mean Transfer	0.00	0.61
PDMS-0	Standard GP-BO	0.00	0.00
	Mean Transfer	1.60	0.75
PDMS-1	Standard GP-BO	0.00	0.00
	Mean Transfer	0.81	1.45
PDMS-2	Standard GP-BO	0.00	0.67
	Mean Transfer	1.07	0.46

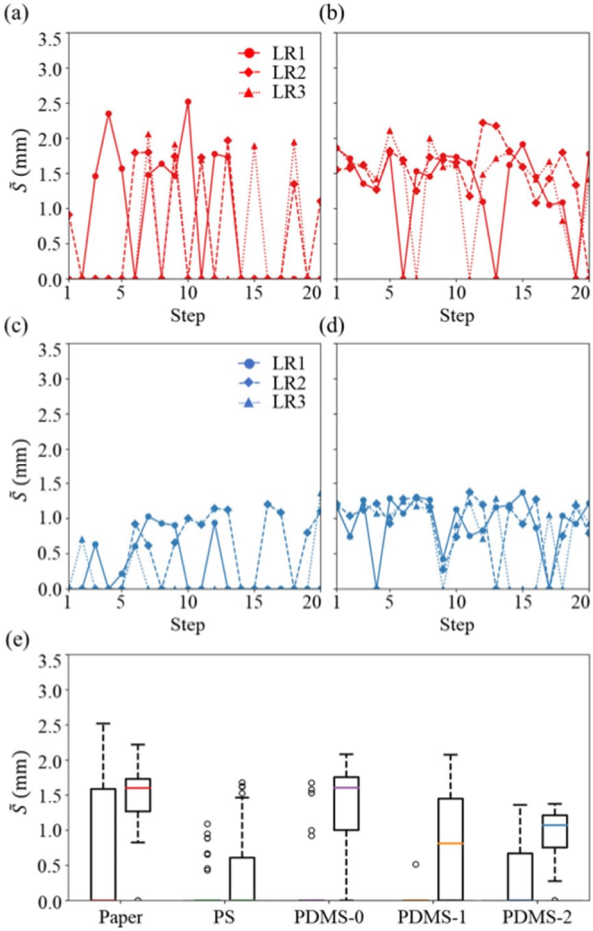


Figure 7. The learning of the controller parameters for the changing adhesion properties of the test surfaces (upper row paper, lower row PDMS-2) within 20 physical experiments in 3 independent learning runs (depicted as LR 1-3). Learning runs with the standard BO in (a) and (c) are compared to the learning runs with the mean transfer in (b) and (d). Overall performances of the learning runs with the standard BO (left bars) and the mean transfer method (right bars) reported with box plots for all test cases (e).

Surface Roughness: Next, we test the same robot on five surfaces with increasing roughness properties that are reported in Figure 2(b). The walking gait performances achieved during the learning runs for the two extreme surfaces, P800-grit ($R_q = 7.488\mu m$) and P60-grit ($R_q = 77.195\mu m$) sandpaper replica are shown in Figures 8(a-d). Similar to the surface adhesion experiments, the mean transfer method improves the learning performance of the standard BO by increasing the number of explored parameter sets that generate non-zero walking gait performances. The comparative performances of the standard BO and the mean transfer method for all the roughness test surfaces are reported in Figure 8 (e) and Table 4. See Extension 3 for a comparison between the walking gaits on P800 and P60, and Extension 7 for the details of the independent learning runs for all the surface roughness experiments.

Table 4 Comparison of the learning performances for changing surface roughness values

Test Surface	Type	Performance (mm)	
		Median	IQR
P800	Standard GP-BO	0.00	1.11
	Mean Transfer	1.18	0.59
P400	Standard GP-BO	0.48	1.06
	Mean Transfer	1.07	0.79
P240	Standard GP-BO	0.00	0.00
	Mean Transfer	1.14	0.35
P120	Standard GP-BO	0.00	0.82
	Mean Transfer	0.37	1.14
P60	Standard GP-BO	0.00	0.00
	Mean Transfer	0.00	1.01

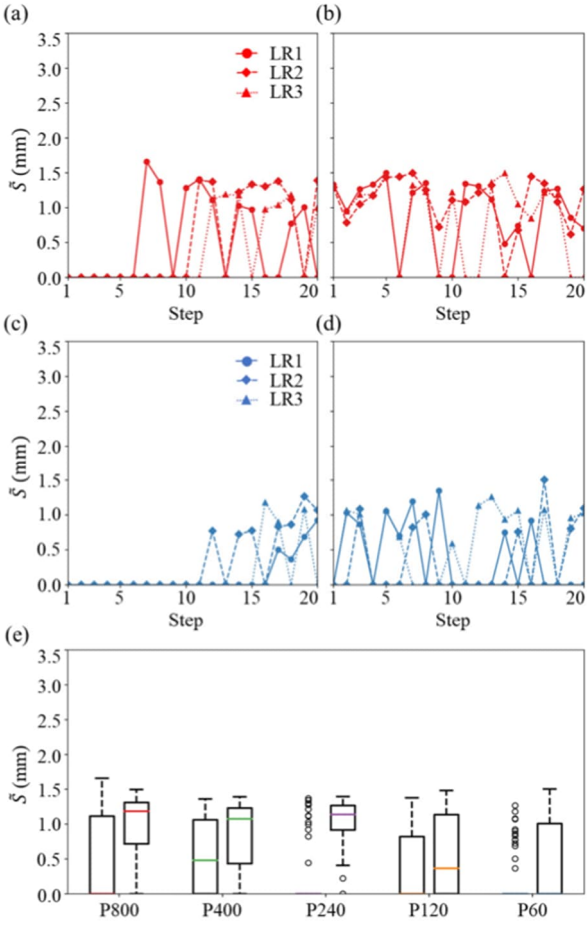


Figure 8. The learning of the controller parameters for the changing roughness values of the test surfaces (upper row P800, lower row P60) within 20 physical experiments in 3 independent learning runs (a, c) without utilizing the prior information and (b, d) with utilizing the prior information. Overall performances of the learning runs with the standard BO (left bars) and the mean transfer method (right bars) reported with box plots for all test cases (e).

Medium Viscosity: Finally, we test our robot walking in eight different media with changing viscosity as reported in Figure 2(c), and report the results for two extreme cases, cP1 and cP90 in Figures 9(a-d). It can be seen that the ability to explore a wide range of controller parameter sets that generate walking gaits decreases for both BO approaches for the high viscosity fluids (cP > 35). However, the mean transfer method still manages to increase the number of successful sets compared to the standard BO for all the media. The overall performances of the standard and mean transfer approaches are reported as box plots in Figure 9(e) and Table 5. These results are consistent with the other two test surfaces that the TL method allows the BO to explore more of the controller sets that generate walking gaits against the changing task space properties. See Extension 4 for a comparison between the walking inside cP1 and cP90, and Extension 8 for the details of the independent learning runs for all the medium viscosity experiments.

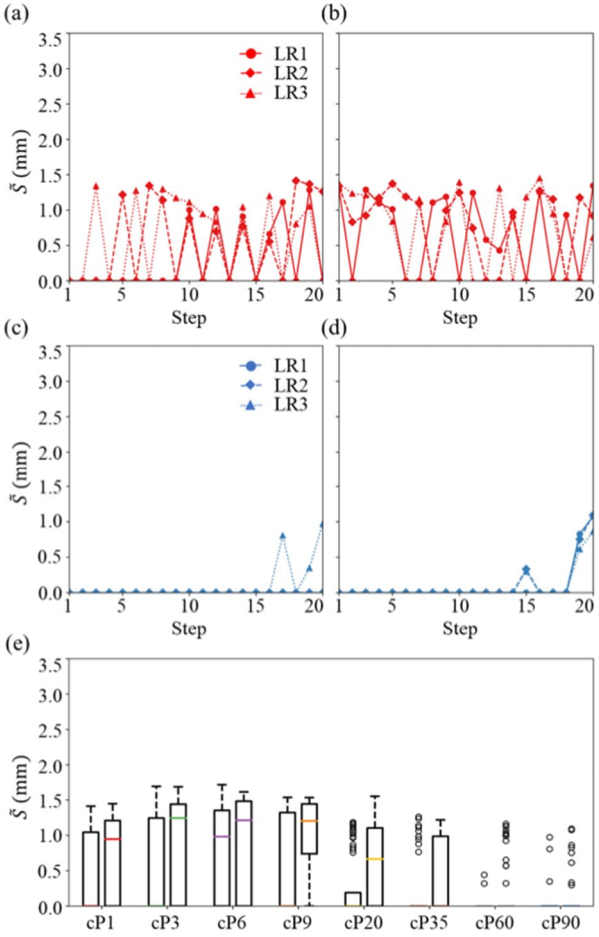


Figure 9. The learning of the controller parameters for the changing test medium viscosity (upper row cP1, lower row cP90) within 20 physical experiments in 3 independent learning runs (a, c) without the prior information and (b, d) with the prior information. Overall performances of the learning runs with the standard BO (left bars) and the mean transfer approach (right bars) reported with box plots for all test cases (e).

Table 5 Comparison of the learning performances for changing medium viscosities

Test Surface	Type	Performance (mm)	
		Median	IQR
cP1	Standard GP-BO	0.00	1.04
	Mean Transfer	0.95	1.21
cP3	Standard GP-BO	0.00	1.25
	Mean Transfer	1.25	1.44
cP6	Standard GP-BO	0.98	1.35
	Mean Transfer	1.21	1.48
cP9	Standard GP-BO	0.00	1.32
	Mean Transfer	1.21	0.70
cP20	Standard GP-BO	0.00	0.19
	Mean Transfer	0.67	1.10
cP35	Standard GP-BO	0.00	0.00
	Mean Transfer	0.00	0.99
cP60	Standard GP-BO	0.00	0.00
	Mean Transfer	0.00	0.00
cP90	Standard GP-BO	0.00	0.00
	Mean Transfer	0.00	0.00

5. Discussion

The displacement measurements from the exhaustive search experiments in Figures 3(a-c) show that even though three identical robots are tested with the same controller parameters on the same surface, they generate different walking gait performances. These initial results also confirm the observations related to performance repeatability in our previous work (Culha et al. 2020). Moreover, the influence of the task space on the robot performance can be seen clearly from Figures 3(h-j), where the adhesion and roughness differences between different surfaces are reflected. These

observations support the necessity of a data-efficient controller learning system that is robust to the robot performance variabilities caused by the material, fabrication, and the task environment of the small scale, medical-oriented, and untethered soft robots.

Our choice on the application of BO and GPs to directly learn the controller parameters of our soft millirobot is based on three aspects. First, the BO offers the efficient data-driven optimization of continuum and complex black-box functions that do not have a closed-form definition. This feature addresses our challenges with not having a deterministic model for the kinematics for our robot and requiring to achieve optimized walking gaits in a small number of experiments. Second, the investigated function can be represented with GPs within the BO, which allow robustness to noise and unknown disturbances. As our robot inherits fabrication-, material-, and interaction-based performance disturbances, GPs provide us with a walking gait function representation that are robust to these variances. Finally, BO is a global optimization tool that avoids getting stuck at local minima, which is important for exploring the parameter space of the investigated function. This feature does not only allow us to find the optimum controller parameters for our robot but it also enables the transfer learning between different robots and task spaces. Even though we do not claim that BO is the best optimization tool, these three aspects successfully address the controller and modeling challenges existing for our robotic system and makes the BO the choice of our application.

The experimental results in this study show that our approach of using BO with GPs and TL methods allowed a data-efficient (i.e., using as few experiments as possible) controller learning that achieves adaptation to different task spaces within a wide range (i.e., on the scale of an order of magnitude) of surface and medium properties. Our main goal is to allow the learning system to explore the controller parameter space to find more of the parameter sets that generate successful walking gaits in response to changing task environments. For this purpose we configured our BO to favor exploration more than exploitation. That is why we do not focus on finding the optimum walking gait controller parameters for each robot or task space in our experiments. Consequently, our current approach does not establish a straightforward correlation between the change of controller parameters with respect to changing robot and task conditions. The comparative results between the standard BO and the TL methods show that both approaches can find sub-optimum parameter sets owing to the statistical nature of the learning method, whose results are given in Appendix B for the task space adaptation experiments. However, we propose that TL methods may allow the system to explore a larger portion of the function space in fewer number of physical experiments, hence achieving data-efficiency in learning.

In terms of experimental learning efficiency, the transfer of the prior mean information outperformed the other TL methods in our experiments. The transfer of this information allowed the BO to start the parameter exploration in the function space within the regions of high-performance result expectations. Therefore, it took the BO much faster to explore the parameter spaces that generate optimum walking gaits (see Extension 5

for a sample comparison of parameter selection with the standard BO and mean transfer method). We see the same effect for the test case of robot 3 walking on P800 in Figure 5(s). Here, this TL method allowed the exploration of the regions with higher expected results and surpassed the exploration boundary of the standard BO. The larger variance in the stride length performances explored by this TL method is caused by this exploration tendency. In comparison, we see that the HP-4 method failed to explore the controller parameters that yield optimum gaits because of the transfer of the signal variance parameter σ_f^2. When this parameter was optimized σ_f^2 for a single robot (i.e., robot 3 on paper in our case), it resulted in a smaller value than the one used in standard BO and eventually it decreased the exploration weight in Equation 11. So, the BO that started with this transferred parameter value focused more on the exploitation of known regions as soon as it found a parameter set generating non-zero performance, and avoided exploring the unobserved regions for the remaining of the iteration steps in the learning runs. We see the influence of the hindered exploration in the first column of Figure 5. We see that the transfer of the length scale hyperparameters in HP-l_c was typically ineffective in our test cases because the length scale is bound to the robot's geometry and magnetic profile, and we used robots with the same geometries in the experiments. The hybrid transfer approach, which also included the transfer of length scale hypermarameters, showed similar performances with the prior mean transfer. These similar results also show the ineffectiveness of length scales hyperparameters and the dominance of prior mean information. Although the hybrid approach could be extended by also transferring the

variance terms σ_n^2 and σ_f^2 along with the prior mean information, we did not include these alternatives in the TL comparison experiments for two reasons. First, the negative impact of the signal variance σ_f^2 on the exploration capability was already shown, and second, σ_n^2, i.e., the noise in the collected data set remained the same for all test cases as we used the same hardware setup for all our experiments. The values of the hyperparameters can also be seen in Appendix B. The choice of the parameters such as the kernels (Wilson 2014) and hyperparameters (Chen and Wang 2018) can also be replaced with other methods, however, the systematic analysis of their influence on learning performance is beyond the scope of our current study.

In the learning experiments that compare the standard BO and four TL approaches, we chose to represent learning performance with median and IQR instead of mean and standard deviation (as seen in Figure 4), since IQR is a robust measure of scale, as it is less sensitive to the outliers in the data. Moreover, dissimilar to standard deviation, IQR can represent the skewness in the distribution of the walking performance results, which becomes more apparent as the performance values get closer to the ends of possible performance ranges. In addition to its advantages in the statistical distribution representation, IQR does not report any unachievable result according to the gait definition in Section 2.2.

5.1. Conclusion and Future Work

In this study, we investigate the use of BO with GPs to experimentally learn the controller parameters for the walking gait of a magnetic soft

millirobot. We create benchmark data sets consisting of 750 experimental results using an exhaustive search to find the walking gait function space for five different test cases. We then use these data sets to compare the effectiveness of four different TL methods to complement the standard BO learning. In these experiments that involve 10^4 learning steps for each test case, we show that the transfer of the prior mean information increased the BO-learning performance the most in terms of increasing the number of explored sub-optimum controller parameters and decreasing the number of required experiments. Based on these findings, we apply BO learning together with the prior mean transfer method on different task spaces with changing surface adhesion, surface roughness, and medium viscosity. We show that controller learning with a BO that utilizes prior mean transfer demonstrates successful adaptation to task spaces in a data-efficient way by exploring the function space of the robot in fewer experiments to find a larger group of controller parameters that yield successful walking gaits.

Our approach is not only limited to walking gait learning and it can further be applied to different locomotion and manipulation controllers for soft robots (Chin et al. 2020). In future, studies focusing on small-scale fabrication with higher magnetization resolution may address the fabrication reproducibility issues (Kim et al. 2018; Xu et al. 2019; Alapan et al. 2020). However, especially for robots designed for biomedical operations, the interaction with the dynamic task environment may still have degrading robot material and performance effects. For such scenarios, a data-efficient controller learning system may adapt optimum controller parameters to these changes in the robot. For example, such an

approach may be applied to endoscopic soft robots within or outside the gastrointestinal (GI) tract (Yim et al. 2014; Son et al. 2020) using a small number of trials. Our study can be further extended to involve the design parameters, such as the magnetic particle density in our robots, and guide the task-oriented design strategies for future soft mobile robots. Our approach can be used to reveal design guidelines to improve the kinematic models of the small-scale robots while utilizing the constant curvature (CC) approximations (Webster III and Jones 2010), analytical models (Renda et al. 2014), and FEM methods (Largilliere et al. 2015). However, as the BO we are using is an episodic algorithm, meaning that each suggested parameter set must be evaluated first in an experiment, the adaptation to design optimization will require the experiments to be run either in a simulation environment or an automated rapid fabrication system that can be integrated within the actuation architecture. The systematic comparison of our experimental approach to alternative optimization and control methods supported with simulations such as intelligent trial and error (Cully et al. 2015), evolution algorithms (Kriegman et al. 2020), or policy gradients (Sehnke et al. 2010) is beyond the scope of our current study but is an interesting future work. We believe that the benchmark datasets available in this study can be used to compare these different methods. Our long-term vision is to build fully autonomous systems that can control, track, evaluate, and optimize soft robots operating in changing complex real-world environments, with minimum human involvement.

6. Declaration of conflicting interests

The author(s) declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

7. Data Accessibility Statement

The experimental data from the benchmark data sets are available for access in here:

https://github.com/sozgundemir/softrobotwalkingdataset

8. Funding

This work was funded in part by the Alexander von Humboldt Foundation, the Ministry of National Education of the Republic of Türkiye, the Cyber Valley Initiative, the Grassroots Initiative of the Max Planck Institute for Intelligent Systems, the Max Planck Society, and the European Research Council (ERC) Advanced Grant “SoMMoR” Project with Grant No: 834531.

Acknowledgments

S.O.D. thanks the Ministry of National Education of the Republic of Türkiye for the Doctoral Scholarship. U.C. and A.P.F. thank the Alexander von Humboldt Foundation for the Humboldt Postdoctoral Research Fellowship and the Federal Ministry for Education and Research.

References

Alapan Y, Karacakol AC, Guzelhan SN, Isik I and Sitti M (2020) Reprogrammable shape morphing of magnetic soft machines. *Science Advances* 6(38): eabc6414.

Alexander RM (1984) The Gaits of bipedal and quadrupedal animals. *The International Journal of Robotics Research* 3(2): 49–59.

Bern JM, Schnider Y, Banzet P, Kumar N and Coros S (2020) Soft Robot Control With a Learned Differentiable Model. In: *2020 3rd IEEE International Conference on Soft Robotics (RoboSoft)*. pp. 417–423.

Brochu E, Cora VM and de Freitas N (2010) A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. *arXiv:1612.06830 [cs/LG]*.

Calandra R, Seyfarth A, Peters J and Deisenroth MP (2016) Bayesian optimization for learning gaits under uncertainty. *Annals of Mathematics and Artificial Intelligence* 76(1-2): 5–23.

Calisti M, Picardi G and Laschi C (2017) Fundamentals of soft robot locomotion. *Journal of The Royal Society Interface* 14(130): 20170101.

Chatzilygeroudis K, Vassiliades V, Stulp F, Calinon S and Mouret JB (2019) A survey on policy search algorithms for learning robot controllers in a handful of trials. *IEEE Transactions on Robotics* 36(2): 328 – 347.

Chen Z and Wang B (2018) How priors of initial hyperparameters affect Gaussian process regression models. *Neurocomputing* 275: 1702–1710.

Chenevier J, Gonzalez D, Aguado JV, Chinesta F and Cueto E (2018) Reduced-order modeling of soft robots. *PloS One* 13(2): e0192052.

Chin K, Hellebrekers T and Majidi C (2020) Machine learning for soft robotic sensing and control. *Advanced Intelligent Systems* 2(6): 1900171.

Cianchetti M, Laschi C, Menciassi A and Dario P (2018) Biomedical applications of soft robotics. *Nature Reviews Materials* 3(6): 143–153.

Culha U, Demir SO, Trimpe S and Sitti M (2020) Learning of sub-optimal gait controllers for magnetic walking soft millirobots. In: *Proceedings of Robotics: Science and Systems (RSS)*.

Cully A, Clune J, Tarapore D and Mouret JB (2015) Robots that can adapt like animals. *Nature* 521(7553): 503–507.

Delhaisse B, Esteban D, Rozo L and Caldwell D (2017) Transfer learning of shared latent spaces between robots with similar kinematic structure. In: *2017 International Joint Conference on Neural Networks (IJCNN)*. pp. 4142–4149.

Della Santina C, Bicchi A and Rus D (2020) On an improved state parametrization for soft robots with piecewise constant curvature and its use in model based control. *IEEE Robotics and Automation Letters* 5(2): 1001–1008.

Devin C, Gupta A, Darrell T, Abbeel P and Levine S (2017) Learning modular neural network policies for multi-task and multi-robot transfer. In: *2017 IEEE International Conference on Robotics and Automation (ICRA)*. pp. 2169–2176.

Duvenaud DK, Nickisch H and Rasmussen C (2011) Additive Gaussian Processes. In: *Advances in Neural Information Processing Systems*, volume 24. Curran Associates, Inc.

Fang G, Wang X, Wang K, Lee KH, Ho JD, Fu HC, Fu DKC and Kwok KW (2019) Vision-based online learning kinematic control for soft robots using local gaussian process regression. *IEEE Robotics and Automation Letters* 4(2): 1194–1201.

Ghahramani Z (2015) Probabilistic machine learning and artificial intelligence. *Nature* 521(7553): 452–459.

Goury O and Duriez C (2018) Fast, generic, and reliable control and simulation of soft robots using model order reduction. *IEEE Transactions on Robotics* 34(6): 1565–1576.

GPy (2012) GPy: A gaussian process framework in python. http://github.com/SheffieldML/GPy.

Grazioso S, Di Gironimo G and Siciliano B (2019) A geometrically exact model for soft continuum robots: The finite element deformation space formulation. *Soft Robotics* 6(6): 790–811.

Gu H, Boehler Q, Cui H, Secchi E, Savorana G, De Marco C, Gervasoni S, Peyron Q, Huang TY, Pane S, Hirt AM, Ahmed D and Nelson BJ (2020) Magnetic cilia carpets with programmable metachronal waves. *Nature Communications* 11(1): 1–10.

Haddadin S, De Luca A and Albu-Schaffer A (2017) Robot collisions: A survey on detection, isolation, and identification. *IEEE Transactions on Robotics* 33(6): 1292–1312.

Helwa MK and Schoellig AP (2017) Multi-robot transfer learning: A dynamical system perspective. In: *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. pp. 4702–4708.

Hiller J and Lipson H (2014) Dynamic simulation of soft multimaterial 3d-printed objects. *Soft Robotics* 1(1): 88–101.

Holsten F, Engell-Nørregard MP, Darkner S and Erleben K (2019) Data driven inverse kinematics of soft robots using local models. In: *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 6251–6257.

Hu W, Lum GZ, Mastrangeli M and Sitti M (2018) Small-scale soft-bodied robot with multimodal locomotion. *Nature* 554(7690): 81–85.

Huang W, Huang X, Majidi C and Jawed MK (2020) Dynamic simulation of articulated soft robots. *Nature Communications* 11(2233): 1–9.

Hughes J, Culha U, Giardina F, Guenther F, Rosendo A and Iida F (2016) Soft manipulators and grippers: a review. *Frontiers in Robotics and AI* 3: 69.

Hyatt P, Wingate D and Killpack MD (2019) Model-based control of soft actuators using learned non-linear discrete-time models. *Frontiers in Robotics and AI* 6(22): 22.

Iida F and Nurzaman SG (2016) Adaptation of sensor morphology: an integrative view of perception from biologically inspired robotics perspective. *Interface Focus* 6(4): 20160016.

Jeong SH, Zhang S, Hjort K, Hilborn J and Wu Z (2016) PDMS-based elastomer tuned soft, stretchable, and sticky for epidermal electronics. *Advanced Materials* 28(28): 5830–5836.

Jones DR, Schonlau M and Welch WJ (1998) Efficient Global Optimization of Expensive Black-Box Functions. *Journal of Global Optimization* 13(4): 455–492.

Kim Y, Parada GA, Liu S and Zhao X (2019) Ferromagnetic soft continuum robots. *Science Robotics* 4(33): aax7329.

Kim Y, Yuk H, Zhao R, Chester SA and Zhao X (2018) Printing ferromagnetic domains for untethered fast-transforming soft materials. *Nature* 558(7709): 274–279.

Kriegman S, Blackiston D, Levin M and Bongard J (2020) A scalable pipeline for designing reconfigurable organisms. *Proceedings of the National Academy of Sciences* 117(4): 1853–1859.

Largilliere F, Verona V, Coevoet E, Sanz-Lopez M, Dequidt J and Duriez C (2015) Real-time control of soft-robots using asynchronous finite element modeling. In: *2015 IEEE International Conference on Robotics and Automation (ICRA)*. pp. 2550–2555.

Laschi C, Mazzolai B and Cianchetti M (2016) Soft robotics: Technologies and systems pushing the boundaries of robot abilities. *Science Robotics* 1(1): eaah3690.

Liao T, Wang G, Yang B, Lee R, Pister K, Levine S and Calandra R (2019) Data-efficient learning of morphology and controller for a microrobot. In: *2019 IEEE International Conference on Robotics and Automation (ICRA)*. pp. 2488–2494.

Lu H, Zhang M, Yang Y, Huang Q, Fukuda T, Wang Z and Shen Y (2018) A bioinspired multilegged soft millirobot that functions in both dry and wet conditions. *Nature Communications* 9(1): 1–7.

Majidi C (2014) Soft robotics: a perspective—current trends and prospects for the future. *Soft Robotics* 1(1): 5–11.

Makondo N, Rosman B and Hasegawa O (2018) Accelerating model learning with inter-robot knowledge transfer. In: *2018 IEEE International Conference on Robotics and Automation (ICRA)*. pp. 2417–2424.

Marco A, Baumann D, Khadiv M, Hennig P, Righetti L and Trimpe S (2020) Robot Learning with Crash Constraints. *arXiv preprint arXiv:2010.08669*.

Pan SJ and Yang Q (2009) A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering* 22(10): 1345–1359.

Patacchiola M, Turner J, Crowley EJ, O'Boyle M and Storkey A (2020) Bayesian Meta-Learning for the Few-Shot Setting via Deep Kernels. In: *Advances in Neural Information Processing Systems (NeurIPS)*.

Perrone V, Shen H, Seeger MW, Archambeau C and Jenatton R (2019) Learning search spaces for bayesian optimization: Another view of hyperparameter transfer learning. In: *Advances in Neural Information Processing Systems (NeurIPS)*. pp. 12771–12781.

Polygerinos P, Correll N, Morin SA, Mosadegh B, Onal CD, Petersen K, Cianchetti M, Tolley MT and Shepherd RF (2017) Soft robotics: Review of fluid-driven intrinsically soft devices; manufacturing, sensing, control, and applications in human-robot interaction. *Advanced Engineering Materials* 19(12):1700016.

Raina R, Ng AY and Koller D (2006) Constructing informative priors using transfer learning. In: *Proceedings of the 23rd International Conference on Machine Learning (ICML)*. pp. 713–720.

Rasmussen CE and Williams CKI (2006) *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press. ISBN 026218253X.

Ren Z, Hu W, Dong X and Sitti M (2019) Multi-functional soft-bodied jellyfish-like swimming. *Nature Communications* 10(1): 1–12.

Renda F, Boyer F, Dias J and Seneviratne L (2018) Discrete cosserat approach for multisection soft manipulator dynamics. *IEEE Transactions on Robotics* 34(6): 1518–1533.

Renda F, Giorelli M, Calisti M, Cianchetti M and Laschi C (2014) Dynamic model of a multibending soft robot arm driven by cables. *IEEE Transactions on Robotics* 30(5): 1109–1122.

Rich SI, Wood RJ and Majidi C (2018) Untethered soft robotics. *Nature Electronics* 1(2): 102.

Rieffel J and Mouret JB (2018) Adaptive and resilient soft tensegrity robots. *Soft Robotics* 5(3): 318–329.

Rus D and Tolley MT (2015) Design, fabrication and control of soft robots. *Nature* 521(7553): 467–475.

Schramm L, Sintov A and Boularias A (2020) Learning to transfer dynamic models of underactuated soft robotic hands. In: *2020 IEEE International Conference on Robotics and Automation (ICRA)*. pp. 4579–4585.

Sehnke F, Osendorfer C, Ruckstieß T, Graves A, Peters J and Schmidhuber J (2010) Parameter-exploring policy gradients. *Neural Networks* 23(4): 551–559.

Shahriari B, Swersky K, Wang Z, Adams RP and De Freitas N (2015) Taking the human out of the loop: A review of Bayesian optimization. *Proceedings of the IEEE* 104(1): 148–175.

Shen Z, Chen F, Zhu X, Yong KT and Gu G (2020) Stimuli-responsive functional materials for soft robotics. *Journal of Materials Chemistry B* 8(39): 8972–8991.

Sitti M (2018) Miniature soft robots—road to the clinic. *Nature Reviews Materials* 3(6): 74–75.

Son D, Gilbert H and Sitti M (2020) Magnetically actuated soft capsule endoscope for fine-needle biopsy. *Soft Robotics* 7(1): 10–21.

Thuruthel TG, Ansari Y, Falotico E and Laschi C (2018) Control strategies for soft robotic manipulators: A survey. *Soft Robotics* 5(2): 149–163.

Thuruthel TG, Shih B, Laschi C and Tolley MT (2019) Soft robot perception using embedded soft sensors and recurrent neural networks. *Science Robotics* 4(26): aav1488.

Trimmer BA and Lin HT (2014) Bone-free: Soft mechanics for adaptive locomotion. *Integrative and Comparative Biology* 54(6): 1122–1135.

von Rohr A, Trimpe S, Marco A, Fischer P and Palagi S (2018) Gait learning for soft microrobots controlled by light fields. In: *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. pp. 6199–6206.

Walsh C (2018) Human-in-the-loop development of soft wearable robots. *Nature Reviews Materials* 3(6): 78–80.

Wang K, Chen J, Xie L and Su H (2020) Transfer learning based on incorporating source knowledge using Gaussian process models for quick modeling of dynamic target processes. *Chemometrics and Intelligent Laboratory Systems* 198: 103911.

Webster III RJ and Jones BA (2010) Design and kinematic modeling of constant curvature continuum robots: A review. *The International Journal of Robotics Research* 29(13): 1661–1683.

Wilson AG (2014) *Covariance kernels for fast automatic pattern discovery and extrapolation with Gaussian processes*. PhD Thesis, University of Cambridge Cambridge, UK.

Wu Y, Yim JK, Liang J, Shao Z, Qi M, Zhong J, Luo Z, Yan X, Zhang M, Wang X, Fearing RS, Full RJ and Lin L (2019) Insect-scale fast moving and ultrarobust soft robot. *Science Robotics* 4(32): eaax1594.

Xu T, Zhang J, Salehizadeh M, Onaizah O and Diller E (2019) Millimeter-scale flexible robots with programmable three-dimensional magnetization and motions. *Science Robotics* 4(29): eaav4494.

Yang B, Wang G, Calandra R, Contreras D, Levine S and Pister K (2018) Learning flexible and reusable locomotion primitives for a microrobot. *IEEE Robotics and Automation Letters* 3(3): 1904–1911.

Yim S, Gultepe E, Gracias DH and Sitti M (2014) Biopsy using a magnetic capsule endoscope carrying, releasing, and retrieving untethered microgrippers. *IEEE Transactions on Biomedical Engineering* 61(2): 513–521.

Appendix A. Pseudo-code for robot tracking and motion evaluation algorithm

Algorithm 1 Robot Tacking and Motion Evaluation

While Test is running **do**
    Image ← Capture robot's image by orthogonal camera
    Image ← Correct distortion and misalignment in the *Image*
    RobotImage ← Apply threshold on *Image* to find robot pixels,
    RobotLegs ← Find the leg positions running Grassfire algorithm on *RobotImage*,
    For $Leg \in RobotLegs$ **do**
        LegState ← Label leg state as 'touchDown', 'liftOff', or 'slipping',
    RobotState ← Label robot state using *LegState*
    MotionType ← Label robot's motion either as walking or not by comparing *RobotState* throughout the actuation sequence with the four states in Figure 1(f)

Appendix B. Hyperparameter and Controller Parameter Sets

Below are the exact values of the GP hyperparameters used in the TL experiments and the controller parameter sets in the experiments for the adaptation to task spaces.

Table 7 Hyperparemeter sets used in the benchmark tests

Learning Method	σ_n	σ_f	$l_c^{\alpha_1}$	$l_c^{\alpha_2}$
Standard	0.29	1.85	17.75	15.25
HP-4	0.29	0.80	10.12	9.99
HP-l_c	0.29	1.85	12.47	12.74
Mean	0.29	1.85	17.75	15.25
Hybrid	0.29	1.85	12.47	12.74

Table 8 Best-performing controller parameter sets and the corresponding stride length averages for the changing adhesion strength characteristics of the test surfaces within 20 physical experiments in 3 independent learning runs (60 experiments in total) with and without utilizing the prior information.

Surface	**Type**	**Controller Parameters**				**Stride length** S (mm) (avg±std)
		B_{max} (mT)	f (Hz)	α_1 $(°)$	α_2 $(°)$	
Paper	Standard	11.0	0.5	0.0	80.0	1.48±0.15
	TL	12.0	0.5	0.0	90.0	1.29±0.12
PS	Standard	8.0	0.5	35.0	65.0	0.94±0.03
	TL	12.0	4.0	35.0	90.0	1.46±0.05
PDMS-0	Standard	12.0	1.0	50.0	85.0	1.18±0.05
	TL	11.0	0.5	10.0	80.0	1.49±0.19
PDMS-1	Standard	5.0	0.5	0.0	90.0	0.56±0.06
	TL	12.0	0.5	0.0	85.0	1.37±0.34
PDMS-2	Standard	12.0	0.5	30.0	90.0	1.29±0.06
	TL	12.0	0.5	20.0	90.0	1.24±0.05

Table 9 Best-performing controller parameter sets and the corresponding stride length averages for the changing roughness values of the test surfaces within 20 physical experiments in 3 independent learning runs (60 experiments in total) with and without utilizing the prior information.

Surface	**Type**	**Controller Parameters**				**Stride length** S (mm) (avg±std)
		B_{max} (mT)	f (Hz)	α_1 (°)	α_2 (°)	
P800	Standard	12.0	0.5	0.0	90.0	1.28±0.03
	TL	12.0	0.4	30.0	85.0	1.36±0.05
P400	Standard	12.0	0.5	15.0	70.0	1.18±0.11
	TL	12.0	0.5	40.0	90.0	1.25±0.10
P240	Standard	12.0	0.5	20.0	85.0	1.34±0.08
	TL	12.0	0.5	25.0	90.0	1.28±0.04
P120	Standard	12.0	0.5	0.0	90.0	1.17±0.11
	TL	12.0	2.0	30.0	80.0	1.27±0.13
P60	Standard	12.0	4.0	30.0	85.0	1.26±0.11
	TL	12.0	6.0	30.0	80.0	1.50±0.11

Table 10 Best-performing controller parameter sets and the corresponding stride length averages for the changing viscosity values of the test medium within 20 physical experiments in 3 independent learning runs (60 experiments in total) with and without utilizing the prior information.

Surface	**Type**	**Controller Parameters**				**Stride length** S (mm) (avg±std)
		B_{max} (mT)	f (Hz)	α_1 (°)	α_2 (°)	
cP1	Standard	12.0	0.5	25.0	90.0	1.41±0.03
	TL	11.0	4.0	25.0	75.0	1.45±0.05
cP3	Standard	10.0	1.0	25.0	80.0	1.69±0.04
	TL	12.0	0.5	0.0	80.0	1.69±0.04
cP6	Standard	12.0	0.5	15.0	85.0	1.72±0.09
	TL	12.0	0.5	10.0	90.0	1.61±0.08
cP9	Standard	12.0	0.5	5.0	80.0	1.54±0.02
	TL	12.0	0.5	30.0	75.0	1.53±0.03
cP20	Standard	9.0	0.5	35.0	65.0	1.18±0.05
	TL	9.0	0.5	0.0	85.0	1.55±0.14
cP35	Standard	12.0	0.5	1.0	60.0	1.26±0.10
	TL	11.0	0.5	25.0	65.0	1.22±0.06
cP60	Standard	12.0	20.0	30.0	75.0	0.44±0.14
	TL	12.0	0.5	25.0	65.0	1.16±0.03
cP90	Standard	5.0	0.5	10.0	65.0	0.97±0.16
	TL	8.0	0.5	30.0	65.0	1.09±0.11

Publication 3

Learning Soft Millirobot Multimodal Locomotion with Sim-to-Real Transfer

Abstract

With wireless multimodal locomotion capabilities, magnetic softmillirobots have emerged as potential minimally invasive medical robotic platforms. Due to their diverse shape programming capability, they can generate various locomotion modes, and their locomotion can be adapted to different environments by controlling the external magnetic field signal. Existing adaptation methods, however, are based on hand-tuned signals. Here, a learning-based adaptive magnetic soft millirobot multimodal locomotion framework empowered by sim-to-real transfer is presented. Developing a data-driven magnetic soft millirobot simulation environment, the periodic magnetic actuation signal is learned for a given soft millirobot in simulation. Then, the learned locomotion strategy is deployed to the real world using Bayesian optimization and Gaussian processes. Finally, automated domain recognition and locomotion adaptation for unknown environments using a Kullback-Leibler divergence-based probabilistic method are illustrated. This method can enable soft millirobot locomotion to quickly and continuously adapt to environmental changes and explore the actuation space for unanticipated solutions with minimum experimental cost.

1. Introduction

Among the existing external actuation methods, such as heat,[6] light,[6,7] electric,[8] and magnetic field,[6,9] magnetic actuation stands out due to its high precision, dexterity, speed, penetration depth, and biological safety features.[10] Magnetic soft millirobots have been demonstrated to perform various locomotion modes, such as walking, rolling, crawling, jumping,

tumbling, swimming, and climbing.[11–13] While multiple robots with different magnetization profiles can generate these modes separately,[12] a single robot design with a preprogrammed magnetic profile can also achieve multimodal locomotion under different periodic actuation signals.[6,13–15] Their multimodal locomotion capability and compliance enable them to adapt to the physical changes in their complex environments and perform diverse medical functions, such as on-demand drug delivery, sensing, and embolization, in a target location.[12–14,16]

However, designing adaptive multimodal locomotion strategies exploiting the compliant soft body dynamics is still a challenge for the robust and safe operation of these small-scale magnetic soft robots. In the case of large-scale robotic systems, adaptive locomotion strategies are built on closed-loop controllers utilizing the feedback of onboard shape-sensing sensors and high-fidelity physical models. The soft robot's size scale of less than a centimeter, however, prevents the integration of onboard sensors to obtain robot shape feedback due to the added rigidness and difficulty of scaling down power and communication modules.[17–19] Moreover, factors such as the magnetic and elastic property variations due to available fabrication techniques, material property changes during operation, and complex physical interaction with surroundings make model-based closed-loop control strategies unfit for small-scale magnetic soft robot locomotion.[20] Therefore, the most common approach for magnetic soft millirobots' locomotion strategy is to build simplified quasistatic physical models of the robot and to manually tailor the open-loop locomotion strategies by designing an actuation signal for desired locomotion behavior.[13] Although this approach generates effective

locomotion strategies in engineered environments, it fails to provide robust locomotion strategies in different environmental conditions, such as surface roughness, medium, or varying confinements.[21]

To address the locomotion challenges in different environments, we have previously proposed optimizing the periodic actuation signal for the maximum stride length of a soft millirobot using data-driven Bayesian optimization (BO).[20] Moreover, we have demonstrated the possibility of transferring the learned experience among different robots and environments to overcome the challenges of time- and material-dependent performance variations using the Gaussian process (GP) model with the mean transfer approach.[21] The transfer learning accelerated the domain adaptation of the magnetic soft millirobot in various environmental conditions, such as in high-viscosity mediums or on sticky surfaces. However, relying on physical tests prevents using the proposed learning approach for multimodal locomotion due to varying robot performance through the prolonged experimental time with enlarged search space.

An alternative approach to multimodal locomotion learning is to use sim-to-real transfer. Simulation environments are commonly used in reinforcement learning (RL) for large-scale robotic tasks,[22] such as legged robot locomotion,[23–27] and object manipulation tasks.[28–31] However, the success of sim-to-real transfer approaches in large-scale robots has not been fully reflected in small-scale magnetic soft robots due to the lack of high-speed and high-accuracy simulation systems. Finite-element methods (FEM)-based simulation environments, such as

COMSOL, enable to build of high-quality soft body simulations, even capturing fluid interactions with jellyfish-like magnetic soft robots.[1,14] However, these complex simulations require high computation time and precise knowledge of the environment, which prevents us from using them in sim-to-real transfer learning for adaptive multimodal locomotion in changing environments. A compromise between speed and accuracy could be achieved by using 1D models, such as Cosserat rod theory, to simulate the soft body dynamics of a small-scale soft robot. For instance, Yao et al. have recently demonstrated that 1D Cosserat rod model-based simulations could be used to learn periodic magnetic actuation signals for magnetic soft robot locomotion.[32] However, their method has been limited to relatively simple environments with flat surfaces and could not be generalized to locomotion in more complex 3D environments due to the limitation of the 1D Cosserat rod model. As an alternative, Hiller et al. achieved computationally efficient dynamic soft-body simulation with 3D interactions using coarse structural elements.[33] This simulation environment is further utilized to learn shape and control policy pairs in a given environment for a pneumatically actuated large-scale soft robot.[34] However, this coarse simulation method fails to model small-scale soft robots' dynamic behavior accurately.

In this study, we developed a data-driven simulation environment that accurately models the magnetically actuated soft millirobot in complex environments without compromising the computational efficiency (Figure 1). Next, we introduced a versatile periodic magnetic actuation signal to generate parameterized multimodal locomotion modes. Finally, using the simulated experience in the proposed data-driven magnetic soft

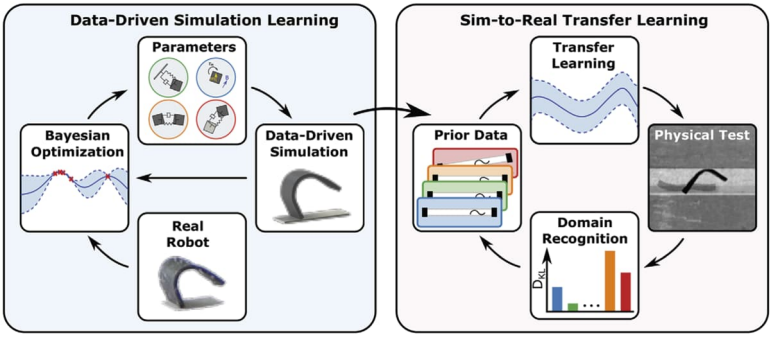


Figure 1. Data-driven magnetic soft millirobot simulation and sim-to-real transfer learning framework. Data-driven magnetic soft millirobot simulation learned the simulation parameters by running BO with GPs to maximize the JI between the simulated and experimental behavior of the sheet-shaped magnetic soft millirobot. The prior data for the sim-to-real transfer learning was generated by running an exhaustive grid search in the data-driven simulation environment for all the given test environments. The domain recognition algorithm continuously compared the observed performance values to the simulated test cases through KLD and identified the environment. The robot's locomotion was learned by sim-to-real transfer using the simulated data of the identified environment as a priori knowledge.

body simulation environment with the transfer learning framework based on BO with GP, we demonstrated that sim-to-real transfer learning can learn the magnetic soft millirobot's locomotion in different environments (Figure 1). Moreover, through the Kullback-Leibler divergence (KLD)-based domain recognition approach, we showed the efficacy of the automated locomotion adaptation to changing environmental confinements. The adaptive magnetic soft millirobot multimodal locomotion framework introduced here fills the gap between simulation and real-world performance, enabling soft millirobot locomotion to

quickly and continuously adapt to environmental changes, thus unlocking the potential of magnetic soft millirobots toward real-world application.

2. Experimental Results

2.1. Actuation Signal Parameterization for Multimodal Locomotion of Magnetic Soft Millirobots

The first challenge in locomotion learning and adaptive control of the magnetic soft millirobots is the parametrization of the magnetic actuation signal. The general strategy in magnetic soft millirobot studies is to use hand-crafted periodic actuation signals to generate desired locomotion modes, making the mode optimization a laborious manual process.[11,14] Previously, we demonstrated that the walking mode could be parameterized for autonomous Bayesian learning-based gait optimization using frequency, field strength, and oscillating field directions in a predefined actuation profile.[20,21] However, generalization to a more comprehensive set of locomotion modes, such as rolling and crawling, requires a higher degree of freedom in actuation signal parameterization. Therefore, in this study, we proposed a generic piece-wise defined periodic magnetic actuation signal for magnetic field magnitude and direction using the parameter set:

$$\theta_{act} = [f, B_{max}, \alpha_1, \Delta\alpha, \Delta T], \quad (1)$$

where frequency (f), maximum field strength (B_{max}), initial field direction (α_1), change in the field direction ($\Delta\alpha$), and duration of direction change (ΔT) define the actuation signal profile (Figure 2a). By allowing the field direction to wrap over 2π and adding additional ΔT, we enabled

rolling and walking modes to be performed by a single function. Using the proposed generic actuation signal, we replicated previously reported actuation signals and the locomotion modes on the sheet-shaped magnetic soft millirobot (Figure 2b; Figure S1a–c and Video S1, Supporting Information).[11] Furthermore, we compared our periodic actuation signal parameterization with the periodic actuation signals learned by an RL-based incremental magnetic field generation approach.[32] We observed that similar signals could be achieved (Figure S1d–f, Supporting Information). Besides replicating the actuation signal in a forward signal generation, i.e., calculating the periodic signal for a given robot and environment, the proposed parameterization also enabled us to define an inverse problem of identifying the environment. By defining a probabilistic framework through the GPs and KLD, we could predict the environmental change and adapt to the new environment by switching between different gaits.

2.2. Data-Driven Magnetic Soft Millirobot Simulation

Following the actuation signal parameterization, the next challenge in learning the optimal locomotion for the magnetic soft millirobot was to generate repeatable training data. Previously, we demonstrated that the walking gait could be optimized efficiently in different environments using the transfer learning approach with automated physical experiments.[21] However, as parameter space enlarges and environmental topologies substantially vary, automated physical experiments become impractical.[20,35]

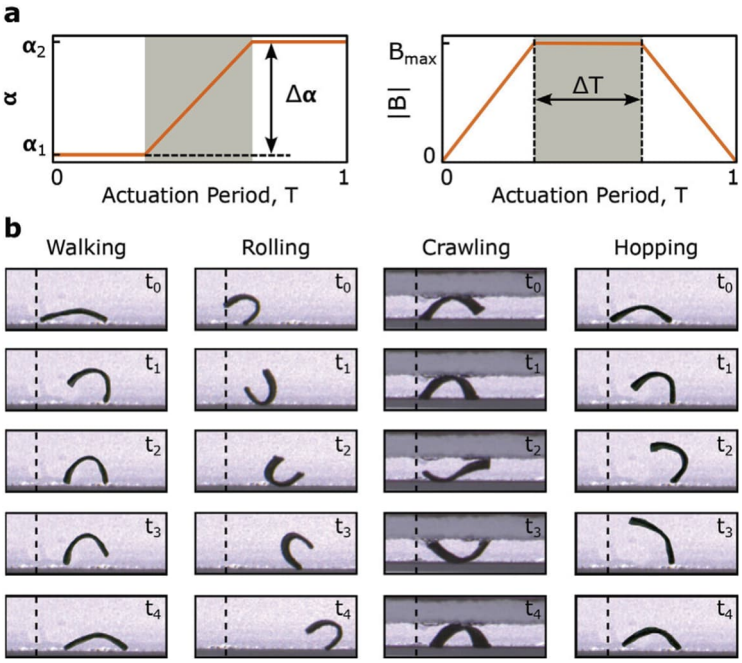


Figure 2. Generic magnetic actuation signal and achieved various locomotion modes. a) Generic actuation signal acting on the yz-plane (parallel to the longitudinal plane of the robot) defines the direction angle, α, and magnitude, $|B|$, of the homogeneous magnetic field B by controlling the actuation parameters $\theta_{act} = [f, B, \alpha_1, \Delta\alpha, \Delta T]$. b) Experimental results for walking, rolling, crawling, and hopping locomotion modes.

A possible solution could be using soft robot simulations instead of physical experiments. However, they have a clear trade-off between accurately capturing the robot dynamics, including physical interactions with the surroundings, and the simulation speed. For instance, FEM-based methods can accurately model the dynamic behavior of a jellyfish-like magnetic soft robot inside a fluidic environment,[1] while

sacrificing computational efficiency.[36,37] In contrast, the Cosserat rod model-based simulation can achieve higher simulation speeds but cannot capture 3D interactions within the environment and the robot body.[32] Alternatively, data-driven methods are proposed to replace computationally inefficient models without compromising accuracy. However, the training datasets' size increases with the modeled system's complexity.[38] Therefore, hybrid approaches combining the analytical model on the high level, such as rigid body dynamics, and the data-driven model on the low level, such as actuator dynamics, are proposed for large-scale robotic systems.[24] For small-scale soft robotic systems, on the other hand, using a modular hybrid design is not possible.

To bridge the gap between these methods, we used a data-driven simulation approach in this study. We implemented a magnetic soft millirobot simulation environment based on the open-source software *Voxelyze*,[33] which can capture the soft body dynamics. Then, we learned the simulation parameters by maximizing the similarity between the simulated and actual behavior of the robot using BO (Figure 1).

We started with implementing the multi-body interaction and magnetic actuation to the *Voxelyze*. To validate the multi-body interaction, we simulated the motion of the magnetic soft millirobot for 400 randomly generated, generic actuation signals (Figure 2a) on a flat surface, which was modeled first by the default floor definition available in *Voxelyze* and then by voxels. Statistical analysis by t-test on the net displacement values of two test cases showed no significant difference (Figure S2a, Supporting Information).

Next, we evaluated the effect of multi-body interaction on computation time by running simulations with the same actuation signals and varying numbers of voxels ranging from 500 to 2500 to define the floor. The results showed that adding multi-body interaction did not affect the simulation speed per voxel. On the other hand, the average computation time for a single simulation step scaled linearly with the increasing number of voxels (Figure S2b, Supporting Information). However, since Voxelyze can run multiple simulations on separate CPU cores simultaneously, running them in parallel overcame the low computational speed problem.

We then focused on the modeling accuracy for the magnetic soft millirobot made of silicone rubber with neodymium-iron-boron (NdFeB) magnetic microparticles with a size of 3.7× 1.5×0.185 mm^3 (Figure S3a, Supporting Information). *Voxelyze* models the dynamic behavior of heterogeneous 3D rigid and soft bodies using a mass-spring-damper system, as shown in Figure 3a (i–iv). While the spring coefficients were derived from material properties, the damping coefficients defined the interaction between connected voxels (c_{bond}), colliding voxels ($c_{collide}$), and voxel and surrounding (c_{global}) could not be measured or derived. Besides, the static (μ_s) and dynamic (μ_d) friction coefficients between the robot and surface materials could not be determined due to the adhesive characteristics of the material (Figure S4, Supporting Information). Therefore, we determined the simulation parameters $\theta_{sim} = [c_{bond}, c_{collide}, c_{global}, \mu_s, \mu_d]$ by running GP-BO with the optimization goal set to the Jaccard index (JI) maximization, which

compared the simulation outputs with the ground truth data and measured their similarity (Figure S5, Supporting Information). To find the optimum simulation parameter set θ_{sim}, we ran BO following Algorithm 1 with the 3246 physical experiments given as the ground truth. The ground truth dataset contained the experimental results from our previous work, which was generated by testing the walking performance of two robots for 150 different controller signals with five repetitions on a flat paper surface for varying field direction angles and constant magnetic field strength ($|B| = 10mT$) and frequency ($f = 1Hz$) values.[21] To enlarge the dataset and include the dynamic behavior of the robot with higher frequency actuation, we tested a new robot by running an exhaustive grid search with a wider range of field strength ($|B| \in [7,10]mT$) and frequency ($f \in [1,5]Hz$) values and collected physical data for 582 different controller signals with three repetitions. While defining the search space, we set the range of each simulation parameter in θ_{sim} according to their definition range and physical limitations.[33] Thus, damping coefficients ($c_{bond}, c_{collide}, c_{global}$) ranged from 0.0001 to 1.0 and were discretized by dividing the whole range into 20 steps. Friction coefficients (μ_s, μ_d) were defined between 0.3 and 1.5, with a step size of 0.025. Then, we filtered out the friction coefficients that were not satisfying $\mu_s > \mu_d$. As a result, we obtained a total number of 9.8 million possible parameter sets in Θ_{sim}.

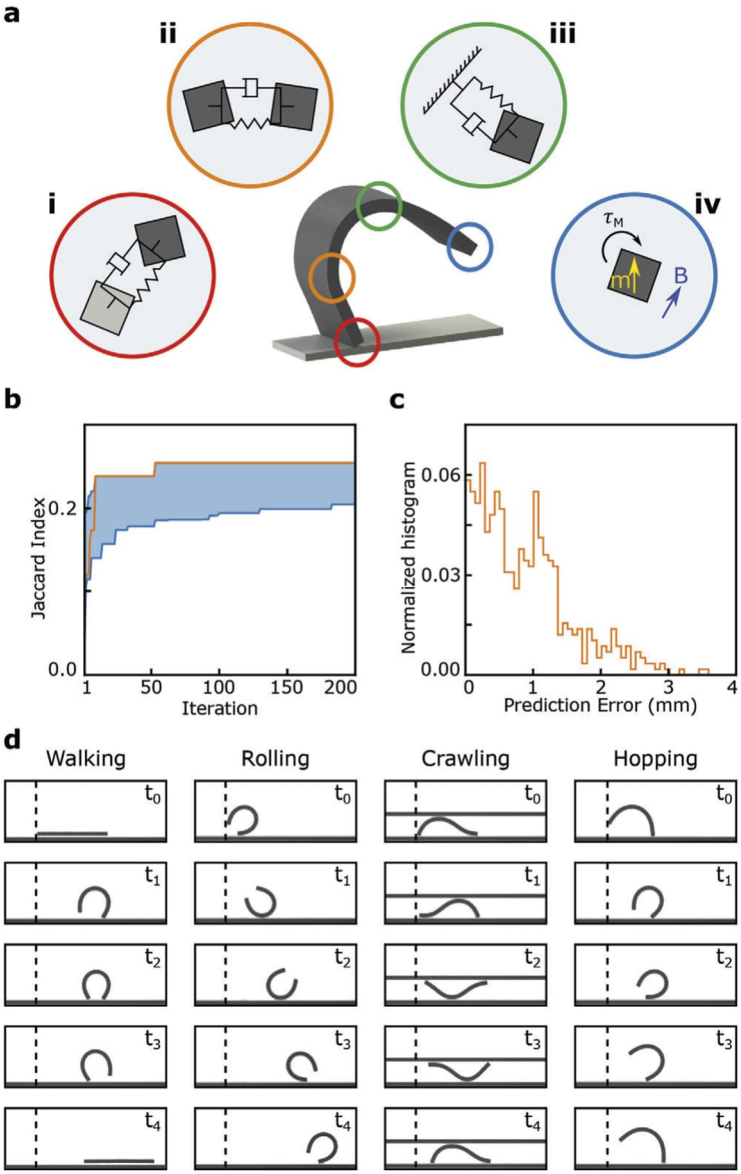

a
i
ii
iii
iv
τ_M
m
B
b
Jaccard Index
0.0
0.2
1
50
100
150
200
Iteration
c
Normalized histogram
0.00
0.03
0.06
0
1
2
3
4
Prediction Error (mm)
d
Walking
Rolling
Crawling
Hopping
t_0
t_1
t_2
t_3
t_4

Figure 3. Simulation parameters and data-driven parameter tuning results for the magnetic soft millirobot simulation. a) Schematic representation of the magnetic soft millirobot simulation by mass-spring-damper model between (i) colliding voxels, $c_{collide}$, (ii) connected voxels, c_{bond}, (iii) voxel and the surrounding, c_{global}, and (iv) magnetic torque acting on a voxel due to external magnetic field B. b) Simulation parameter optimization results obtained by BO with GPs for 30 independent learning runs with 200 iterations. The blue area shows the range of (highest and lowest) JI values obtained by all the optimization runs at any given iteration. The orange line shows the performance of the optimization run, which found the most successful simulation parameter set among all the learning runs. c) Distribution of the average stride length error between the simulation and experimental test results for 582 cases (Figure S6, Supporting Information). d) Data-driven simulation results for walking, rolling, crawling, and hopping locomotion modes.

After completing 30 independent learning runs consisting of 200 iterations, we found out θ^*_{sim} achieving JI= 0.21 (Figure 3b). Next, we evaluated the robot position prediction accuracy of the simulation with θ^*_{sim}. For that purpose, we simulated the robot's motion with the 582 distinct actuation signals used for the ground truth data generation and evaluated the error in the average stride length (Figure S6, Supporting Information). The simulation could predict the robot's position for varying actuation signals with an average error of 0.87 mm, equal to 0.2 body length (BL) (Figure 3c). The accurate deformation and displacement prediction ability of the simulation (Figure 3c,d) enabled us to model the robot's behavior in any given environmental condition and create a priori knowledge about the robot's performance.

Algorithm 1 Data-driven simulation parameter tuning

Inputs: Search space, $\Theta_{sim} = \{c_{bond}, c_{collide}, c_{global}, \mu_s, \mu_d\}$, Experimental data, $ExperimentData$

Output: Best performing simulation parameter set, $\theta^*_{sim} = [c_{bond}, c_{collide}, c_{global}, \mu_s, \mu_d]$

$resultArray \leftarrow$ Initialize an empty array to store JI and θ_{sim}
for $learningRun \in [1,30]$ **do**
  $GP \leftarrow$ Initialize GP with $\mu = 0.5$ and $\sigma_{f,sim} = 0.25$
  for $iteration \in [1,200]$ **do**
    $ActuationSignals,\ TestResults \leftarrow$ Select 8 random experimental data from $ExperimentData$
    $\theta_{sim} \leftarrow$ Select θ_{sim} by BO
    $SimulationEnv \leftarrow$ Create simulation environment with θ_{sim}
    $SimulationData \leftarrow$ Run $SimulationEnv$ with $ActuationSignal$
    $JI \leftarrow$ Compute average JI by Equation 13 using $TestResults$ and $SimulationData$
    $GP \leftarrow$ Update GP with observed JI
    $resultArray \leftarrow$ Add $[\theta_{sim}, JI]$ to $resultArray$
  end
end
$\theta^*_{sim} \leftarrow \underset{\theta_{sim} \in \Theta_{sim}}{\text{argmax}}\ (resultArray)$
return θ^*_{sim}

2.3. Locomotion Optimization with Sim-to-Real Transfer Learning

Next, we used the developed data-driven simulation environment with the BO to learn the optimal locomotion for a specific environment. We set our optimization goal as maximizing the stride length S, i.e., the robot displacement in the forward direction during a complete period of the actuation signal. We simulated the robot locomotion in the given environment to generate the prior data, running an exhaustive grid search.

Testing all the possible actuation signals allowed us to explore all the possible locomotion modes instead of focusing only on user-defined ones, such as walking, rolling, and crawling. Since transferring the GP model's prior mean improves BO's learning performance by increasing the learning speed more than transferring the kernel hyperparameters,[21] we used the simulated data to initialize the GP model for the given task environment. Then, we started running the BO with transfer learning following Algorithm 2 on physical experiments.

We defined the range of actuation signal parameters (θ_{act}) based on the physical limitations of the magnetic actuation setup (Figure S3b, Supporting Information) and the previous findings.[21] Accordingly, B_{max} was defined between 7 and 10 mT, and the actuation frequency (f) ranged from -3 to 3 Hz. We defined the initial field direction (α_1) and the change in the direction ($\Delta\alpha$) as $\alpha_1 \in [0,80]°$ and $\Delta\alpha \in [-30,30]°$, respectively. The duration of direction change (ΔT) ranged from 0.3 to 0.7. We used a step size of 1 mT for B_{max}, 2 Hz for f, $5°$ for α_1, $10°$ for $\Delta\alpha$, and 0.1 for ΔT. To generate rotating actuation signals, we also added $360°$ into the definition of $\Delta\alpha$ and 1.0 into the definition of ΔT. This yielded a total number of 9792 possible parameter sets in Θ_{act}.

Then, we tested the learning performance of the proposed approach for four different test cases with constant profiles (Figure 4). We defined the prior mean function $\mu_{act}(\theta_{act})$ of the GP model for each test case with the corresponding simulation data. We evaluated the effect of the sim-to-real transfer learning approach on learning performance by comparing it to the standard BO in all the task spaces in terms of achieved

stride lengths. For the physical experiments, we set the termination criteria for a learning run as 20 iterations and tested the performance of each learning approach, i.e., standard BO and BO with transfer learning, for three independent learning runs following Algorithm 2.

Algorithm 2 Adaptive locomotion learning with sim-to-real transfer learning

Inputs: Search space, $\Theta_{act} = \{f, B, \alpha_1, \Delta\alpha, \Delta T\}$,
Prior mean function, $\mu_{act}(\theta_{act})$

Output: Best performing actuation parameter set,
$\theta^*_{act} = [f, B, \alpha_1, \Delta\alpha, \Delta T]$

$resultArray \leftarrow$ Initialize an empty array to store θ_{act} and $\tilde{S}$
$GP \leftarrow$ Initialize GP with μ_{act}
for $iteration \in [1,20]$ **do**
  $\theta_{act} \leftarrow$ Select θ_{act} by BO
  $\tilde{S} \leftarrow$ Test θ_{act} by running experiment
  $GP \leftarrow$ Update GP with observed $\tilde{S}$
  $resultArry \leftarrow$ Add $[\theta_{act}, \tilde{S}]$ to $resultArray$
end
$\theta^*_{act} \leftarrow \underset{\theta_{act} \in \Theta_{act}}{\text{argmax}}(resultArray)$
return θ^*_{act}

2.3.1. Different Robots

First, we tested the learning approach for three different robots with the same magnetization profile (i.e., Robots 1, 2, and 3) in a channel with 1 BL height (Figure 4a) to demonstrate the merit of the sim-to-real transfer learning for adapting to changes in the robot. The robots were fabricated following the same procedure (Figure S3a, Supporting Information) but had different characteristics due to stochastic variability during fabrication, deformations during handling, and material

degradation over time. The results showed that BO, both with and without the prior information, could successfully find actuation parameter sets generating forward locomotion for all the robots in the limited number of trials (Figure 4b). Comparing the results of the standard and sim-to-real approaches, we observed that the prior data learned in simulations improved the learning performance for limited physical trials by providing a hot start for parameter optimization (Figure S7, Supporting Information). To further test the sim-to-real transfer learning approach, we repeated the same experiment in a channel with 1 BL height using Robot 1 and two new robots (Robots 4 and 5) with the same structural properties but different magnetic profiles following the study of Yao et al. (Figure S8, Supporting Information).[32] As in the first test case, BO could find actuation parameters for all the robots in the limited number of trials, and also without requiring a redesign of the simulation environment for these new robots, i.e., Robots 4 and 5, the simulated prior knowledge improved the learning performance similar to the previous test case. Improved optimization performance for all robots in these two test cases indicated that the simulation could be used as the a priori knowledge source for the given task. Moreover, using the generic actuation signal instead of the walking gait function allowed the robot to experience different locomotion modes and achieve a higher stride length than the previous studies (see Figure S6 and Table S1, Supporting Information for details).[21]

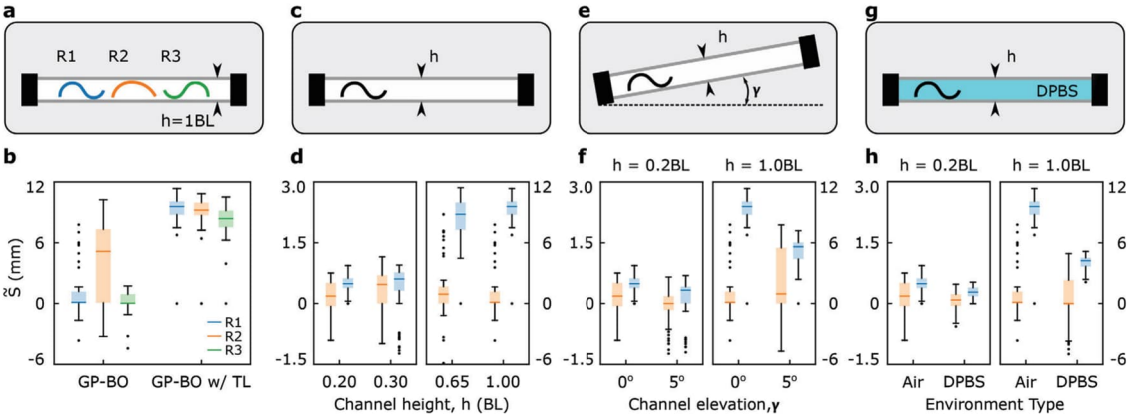


Figure 4. The type and range of task environments (top row) and experimental test results (bottom row). Learning the controller parameters for a, b) three different robots (i.e., Robots 1, 2, and 3) in a channel with 1BL channel height, a single robot (i.e., Robot 1) in a channel with c, d) changing channel heights, e, f) two different channel heights (i.e., 0.2 BL, 1.0BL) and two different elevation angles (i.e., 0°, 5°), g, h) two different channel heights (i.e., 0.2BL, 1.0BL) and two different fluids (Air, DPBS). The results for standard BO are shown on the left (b) and orange (d, f, h) bars. The results for sim-to-real transfer learning are shown on the right (b) and blue (d,f,h) bars. Box plots in the bottom row show the overall performance of the learning approaches as a standard interquartile range (IQR) method, where the horizontal lines are the median of the observed stride lengths $\tilde{S}$ in 60 physical trials for each robot. The box around the median line shows the upper and lower quartiles. The error bars and dots represent the highest and lowest performances and outliers, respectively.

2.3.2. Different Channel Heights

Next, we tested a single robot (Robot 1) in channels with different channel heights $h \in \{0.2, 0.3, 0.65, 1.0\}\, BL$ (Figure 4c). We chose half of the channels to be narrower than $0.38\, BL$ based on the findings of Ren et al.[14] We observed that both standard BO and BO with transfer learning could find the actuation parameters generating forward locomotion in the limited number of trials (Figure 4d). Moreover, similar to the previous test case, using simulation data as the prior mean function improved the learning performance for all the environments by increasing the average stride length of the robot. The difference between the two learning approaches became more evident for the broader channel heights $h \in [0.65, 1.00]\, BL$, as the average achievable displacements increased by an order of magnitude (see Figure S9 and Table S2, Supporting Information for details).

2.3.3. Different Elevation Angles

Later, we put Robot 1 into four different environments with different channel heights $h \in \{0.2, 1.0\}\, BL$ and elevation angles $\gamma \in \{0, 5\}°$ to verify the sim-to-real transfer performance on different slopes (Figure 4e). BO with transfer learning outperformed the standard BO in all the test cases by achieving higher stride lengths (Figure 4f). Especially in the channel with $h = 0.2\, BL$ and $\gamma = 5°$, the difference between these approaches became clearer, where standard BO tended to find less number of parameter sets generating forward locomotion (see Figure S10 and Table S3, Supporting Information for details).

2.3.4. Different Mediums

Finally, we tested Robot 1 in the air and Dulbecco's phosphate buffered saline (DPBS, 14190144, Gibco) filled channels, which decreased the apparent weight of the robot with buoyancy and increased the drag force acting on the robot (Figure 4g). Unlike the previous cases, the robot's motion inside DPBS was not simulated since the simulation parameters were tuned for the robot's behavior in an air-filled environment only. Therefore, we used the prior data generated for the air instead of DPBS. Similar to previous ones, both standard BO (in orange) and BO with transfer learning (in blue) could find the control parameters generating forward locomotion, as shown in Figure 4h (see Figure S11 and Table S4, Supporting Information for details). Although the prior data was generated for the robot moving in the air, BO with transfer learning could still adapt to different environmental conditions and improve learning performance by increasing the stride length achieved.

2.4. Domain Adaptation with Sim-to-Real Transfer Learning

Finally, we deployed our sim-to-real transferred locomotion learning strategy to unknown environments. In this scenario, the learning framework identified the environment without getting explicit information about the environment provided by the user. We developed an automated switching algorithm that continuously compared the observed performance values to the simulated test cases in Figure 4 through the KLD. Then, the simulation data of the chosen environment was used as the a priori knowledge to update the GP (Figure 1).

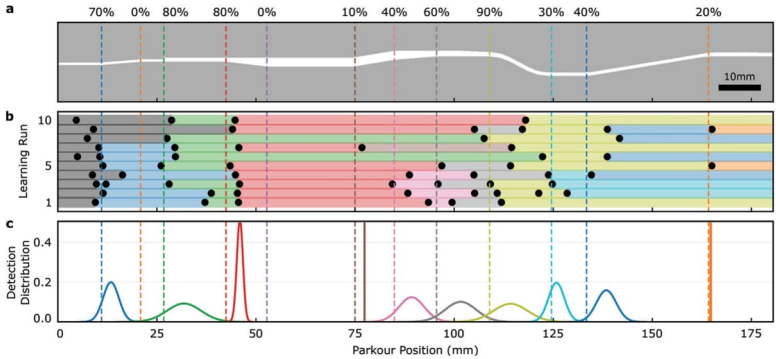


Figure 5. Dynamic task environment and the domain adaptation performance results of the sim-to-real transfer learning for 10 independent learning runs. a) Schematic view of the test environment with 12 domain boundaries shown in different colors. The accuracy of the domain recognition algorithm is reported as percentages on top of the schematic aligned with the corresponding boundary change. b) Detection positions of domain changes for each learning run are shown with dots. c) Domain change detection distance distribution for each domain is shown as the normal distribution. (See Video S2, Supporting Information for a sample learning run).

We tested the domain identification and locomotion adaptation in an environment of 12 varying ceiling heights and elevation angles throughout the path (Figure 5a). We started each learning run by placing the robot on the left entrance of the path and kept iterating Algorithm 3 until the robot reached the other end.

After testing the robot in the channel ten times, we showed that the algorithm could adapt the robot's locomotion and direct the robot to the end of the path in all the trials (Video S2, Supporting Information). During the experiments, the domain identification method demonstrated a 70% accuracy in recognizing the changes in the channel height. Notably

it exhibited a quicker response while detecting an increase in the channel height at ≈4 mm, while the threshold for detecting a decrease was comparatively slower, at ≈5 mm (Figure 5c). In contrast, the algorithm's performance in detecting elevation variations was lower, with an accuracy of only 20% (Figure 5a,b). The disparity in performance became particularly evident in the first, fifth, seventh, and tenth learning runs, where the algorithm could detect all height changes but failed to identify

Algorithm 3 Domain adaptation with sim-to-real transfer learning

Inputs: Search space, $\Theta_{act} = \{f, B, \alpha_1, \Delta\alpha, \Delta T\}$,
Simulation data for *n* many environments,
$SimulationData = \{\mu_{act,1}, \mu_{act,2}, \dots, \mu_{act,n}\}$

Output: Best performing actuation parameter set,
$\theta^*_{act} = [f, B, \alpha_1, \Delta\alpha, \Delta T]$

$resultArray \leftarrow$ Initialize an empty array to store θ_{act}, $\tilde{S}$ and env
$env \leftarrow$ Get random environment $env \in [1, n]$
$GP \leftarrow$ Initialize GP with $\mu_{act,env}$, $\mu_{act,env} \in SimulationData$ and $\sigma_{f_{act}} = 3\,BL$
while *true*
  $\sigma_{act} \leftarrow$ Select θ_{act} by BO
  $\tilde{S} \leftarrow$ Test θ_{act} by running experiment for 5 steps
  $env_{new} \leftarrow \underset{env \in [1,n]}{\operatorname{argmax}} KLD(\mu_{act,new}(\theta_{act}), \tilde{S})$
  if $env \neq env_{new}$
    $GP \leftarrow$ Initialize GP with $\mu_{act,new}$ and $\sigma_{f_{act}} = 3\,BL$
    $env \leftarrow env\_new$
  end
  $resultArry \leftarrow$ Add $[\theta_{act}, \tilde{S}, env]$ to $resultArray$
  if *robot* reached end
    return $resultArray$
  end
end

any elevation changes. Additionally, we observed that the domain identification algorithm could not detect when the robot flipped (Video S2, Supporting Information). However, the locomotion optimization algorithm could still adapt to the new conditions and could find actuation parameters to move the robot in the forward direction.

3. Discussion

This study introduced an adaptive locomotion learning approach based on sim-to-real transfer learning for magnetic soft millirobots. Herein, we developed a high-fidelity data-driven simulation framework to model multi-body interactions and dynamic behavior of magnetically actuated soft millirobots accurately. Using GP-BO, we fine-tuned the physical constants of the simulation environment, minimizing the discrepancy between simulated data and the ground truth dataset of 1746 new and 1500 previously collected physical experimental data.[21] We achieved modeling of the robot's behavior with an average error of 0.87 mm, equal to 0.2 BL. Additionally, we designed a versatile control signal that enabled the magnetic soft millirobot to generate multiple locomotion modes using a single control signal, in contrast to the previous methods relying on distinct control signals or different robot designs.[11,12,14] Furthermore, we demonstrated that our data-driven magnetic soft millirobot simulation could generate a priori knowledge applicable to different robots and environments. Through the sim-to-real transfer method, we bridged the gap between the simulation and real-world performance.[34] Moreover, we showed that the robot could identify unknown environments by matching experimental to simulated data and

could adapt its motion to various conditions it had not previously encountered, thus highlighting its potential for real-world applications.

While we have validated the effectiveness of sim-to-real transfer learning based on BO in various test cases, it may suffer from time inefficiency due to the computational complexity of GP, which is equal to $O(n^3)$, as the a priori dataset gets larger. Hence, we consider the application of data pruning to the simulation-generated a priori knowledge to decrease the data size before defining the prior mean of GP. Alternatively, we will try running the optimization algorithm with selected subsets of the search space, which are chosen according to the expected performance values estimated by the a priori data.

Apart from its time complexity, the proposed sim-to-real-based learning approach may exhibit suboptimal performance in dynamic environments characterized by momentary changes, owing to the episodic nature of BO. One potential solution would be using continuous control algorithms, such as deep reinforcement learning.[27,32,39] These algorithms, especially the neural network-based ones, need a larger training dataset, typically $\approx 10^5$–10^7 timesteps.[32] Therefore, as our next step, we plan to enhance the simulation environment to decrease the average time for a single simulation step. To achieve this goal, we will limit the update in the simulation to a local frame where the robot moves, there by omitting unnecessary calculations in the rest of the environment. Additionally, we will explore the possibility of replacing the simulation with a deep neural network model for higher computation speeds.[38] Besides, the simulation does not model the adhesive interactions, which prevents modeling the

robot's dynamics on sticky surfaces, such as biological tissues covered by mucus. Hence, we plan to implement the adhesion model into the current simulation environment and collect new experimental data to tune the necessary physical parameters for future work.

In this work, we have focused on the multimodal locomotion of sheet-shaped magnetic soft millirobot with different magnetization profiles in 2D environments. However, in the future, the proposed optimization method can be applied to soft robots with different shapes and materials[15] and various tasks in more complex environments, such as climbing, path following, and velocity control in 3D confined spaces.[13] Moreover, the proposed simulation environment and the optimization method can be used to design the robots, i.e., their morphological and magnetic properties, and to learn actuation signals for a given task without physical experiments.[22,27,32,34] Besides, the proposed domain identification method can help localize the robot if the performance map is available.

Even though we tested the applicability of the adaptive locomotion and domain recognition algorithms in artificially designed environments, it is not limited to engineered test conditions. Besides the technical aspects, this method will enable medical usage of soft millirobots in patients by achieving robust and safe control. However, one of the critical challenges that needs to be solved is replacing the visual feedback from the camera with medical imaging modalities. As possible solutions, X-ray,[19] ultrasound,[11] and electrical impedance tomography[40] techniques are proposed to track small-scale magnetic soft robots, which have designs

similar to our robot. As our locomotion adaptation and domain recognition methods require the position data of the robot's center of mass over time, similar clinical imaging methods can be adapted to test our approach for clinical applications in the future.

4. Experimental Section

Magnetic Soft Millirobot Fabrication and Actuation: The sheet-shaped elastomeric magnetic soft millirobot design was used, reported by Hu et al., and used in previous works.[11,20,21] It was fabricated by mixing Ecoflex 00–10 (Smooth-On Inc.) silicone rubber with NdFeB magnetic microparticles with ≈5μm diameter (MQP-15-7, Magnequench) with a 1:1 body mass ratio. After curing the pre-polymer mixture on a methylmethacrylate plate, the robots were cut using a high-resolution laser cutter (LPKF Protolaser U4) with dimensions of length $L = 3.7\ mm$, width $w = 1.5\ mm$, and height $h = 185\ \mu m$. To magnetize the robots, they are folded around a cylindrical rod with a circumference equal to L and put inside a magnetic field with a magnitude of $1.8\ T$. The magnetic field was oriented at $45°$ counterclockwise from the y-axis. After separating the robots from the rod, the magnetic particles maintained their magnetization orientation, forming a circular profile along the robot body (Figure S3a, Supporting Information). To actuate the robot, the homogeneous magnetic field was continuously regulated in the environment and created magnetic torque acting on the robot. By controlling the acting magnetic torque during the actuation, the robot's deformation was controlled and created motion (Figure 2; Video S1, Supporting Information).

Magnetic Actuation and Feedback System: Helmholtz coil setup is used with three orthogonal pairs of electromagnets (Figure S3b, Supporting Information) to generate a homogeneous 3D magnetic field within a $13.1 \times 8.5 \times 4.5\ cm^3$ workspace with a maximum value of $12\ mT$. The magnetic field B was modulated, coinciding with the center of the test environment, by controlling the currents on the electromagnetic coils via six independent motor driver units (Maxon ESCON 70/10). An FPGA module (NI PXIe-7847R) was used as the interface to control the motor drivers, receive current readings, and communicate with the master PC. The mapping between the targeted magnetic field and applied electric currents was regularly calibrated to maintain reliable and repeatable experiments.

The robot's motion was tracked using two high-speed cameras (Basler aCa2040-90uc) running at 120 frames per second (fps). The first camera, orthogonal to the robot's movement plane, was used to identify the robot's locomotion mode. The second camera, having a top view of the test area, was used to measure the displacement of the robot. At the end of each experiment, the average stride length of the robot was calculated by tracking the distance covered by its center of mass in three consecutive steps. Then, the robot automatically moved back to its initial position, which minimized human intervention and human-based disturbances on the robot and the test environment.

To enlarge the workspace having a homogeneous 3D magnetic field and to test the robot's performance for longer runs without reaching the workspace's limits, a motorized linear stage is integrated with 150 mm

stroke (Thorlabs LTS150C) to the y-axis of the Helmholtz coil setup (Figure S3b, Supporting Information). The linear stage was continuously repositioned according to the displacement information received from the imaging system to keep the robot in the center of the magnetic field (Video S2, Supporting Information).

All the communication tasks between different elements of the robotic system, such as image capture, coil control, and learning algorithm, were executed on Robot Operating System (ROS) architecture, which allows the system to be scalable for further extensions.

Magnetic Soft Millirobot Simulation: Dynamic behavior of the magnetic and non-magnetic rigid and soft materials was modeled as a mass-spring-damper system with magnetic torques. A version of the *Voxelyze* was modified to integrate multi-body interaction and magnetic actuation into the simulation environment since it can efficiently simulate heterogeneous 3D rigid and soft bodies under a uniform magnetic field by modeling them as a mass-spring-damper system (Figure 3a(i–iv)).[33]

To model the multi-body interaction, the algorithm of *Voxelyze* is adapted, defining the contact mechanics between the robot and the surroundings. For the magnetic actuation of the robot, the magnetic torque was first calculated by acting on a voxel due to the external homogeneous field and then integrated into the dynamic functions. Magnetic torque (τ_t) acting on a voxel of the robot at time step *t* was calculated as follows.

$$\tau_t = m_t \times B_t, \quad (2)$$

$$m_t = M_r d_v^3 R_t \quad (3)$$

where B_t, M_r and d_v denote the homogeneous magnetic field at time *t*, magnetic remanence, and voxel size, respectively. R_t is the rotational matrix defining the magnetic orientation of the voxel at time step *t* (Figure 3a(iv)).

For all the simulations, the voxel size (d_v) was set to $185\ \mu m$. Density (ρ), Young's modulus (E), and magnetic remanence (M_r) values for the magnetic soft millirobot were taken from [11] as 1.86 g cm^{-3}, 8.45 kPa, and 62 kA/m,respectively. The Poisson's ratio was assumed to be 0.49.

The developed simulation engine and the generated dataset for the evaluation are available here.

Gaussian Processes and Bayesian Optimization: A probabilistic learning approach was used for both optimization problems, that was simulation parameter tuning and controller adaptation, based on the BO and GP.

Since the reward functions in both problems do not have an accurate model, based on the collected data are approximated. To overcome the sparsity due to large search spaces, to include uncertainties coming from the experimental data, and to make probabilistic predictions, GPs are used following the previous study[21] as:

$$R(\theta) = GP\big(\mu(\theta), k(\theta, \theta')\big) \quad (4)$$

where $R(\theta)$ is the reward function mapping the input parameter θ to scalar reward values, $\mu(\theta)$ denotes the prior mean for the input parameter θ and $k(\theta, \theta')$ is the kernel function defining the covariance between $R(\theta)$ and $R(\theta')$ for $\theta, \theta' \in \Theta$. For the cases where $R(\theta)$ contains noise due to the measurements, the observed reward value $\tilde{R}$ is defined as

$$\tilde{R}(\theta) = R(\theta) + n \tag{5}$$

where n stands for the zero-mean Gaussian noise with variance σ_n^2 for each measurement. At each iteration of the optimization run, the GP model was updated with $\tilde{R}(\theta)$.

Using the test data $D = \{\theta_i, \tilde{R}(\theta_i)\}_{i=1}^{N}$, where N is the size of the dataset D, R can be predicted for any given θ using the posterior mean and variance defined as:

$$\mu_{post}(\theta) = \mu(\theta) + k^T(\theta)K^{-1}y \tag{6}$$

$$\sigma_{post}^2(\theta) = k(\theta, \theta) - k^T(\theta)K^{-1}k(\theta) \tag{7}$$

$$R_{post}(\theta)|D \sim \mathcal{N}\left(\mu_{post}(\theta), \sigma_{post}^2(\theta)\right) \tag{8}$$

where $k(\theta)$, $y \in \mathbb{R}^N$, and $K \in \mathbb{R}^{N \times N}$ denote $[k(\theta)]_i = k(\theta, \theta_i)$, $y_i = \tilde{R}(\theta_i) - \mu(\theta_i)$, and $K_{i,j} = k(\theta_i, \theta_j) + \delta_{i,j}\sigma_n^2$ with Kronecker delta $\delta_{i,j}$, respectively. Due to its successful results in similar robotic applications,[21,41,42] the squared exponential function is used with automatic relevance detection (ARD-SE) as the kernel function defined for multi-dimensional cases as follows.

$$k_{SE}(\theta, \theta') = \sigma_f^2 \exp\left(-\sum_{d=1}^{d_c} \frac{(\theta_d - \theta'_d)^2}{2l_{c,d}^2}\right) \quad (9)$$

where $l_c \in \mathbb{R}_c^d$ is the length scale defining the rate of change in the modeled function for each parameter space dimension.[43] For slowly-varying functions l_c is set to be high, and for quickly varying functions l_c is set to be low. The signal variance σ_f^2 describes the uncertainty in the predictions for unobserved θ.

To solve both optimization problems, BO is used with GP, which selects the parameter set θ_{next} to be tested based on the acquisition function $\alpha_{act}(\theta)$ value.

$$\theta_{next} = \underset{\theta \in \Theta}{\operatorname{argmax}}\, \alpha_{acq}(\theta) \quad (10)$$

where the acquisition function $\alpha_{acq}(\theta)$ was the expected improvement (EI) due to its better performance than its alternatives.[41] EI is defined as

$$\alpha_{acq}(\theta) = \mathbb{E}\left[\max\left(0, \left(R(\theta) - \tilde{R}(\theta^*)\right)\right)\right] \quad (11)$$

where $\tilde{R}(\theta^*)$ is the highest observed reward function value.[44] The analytical solution for Equation (11) is given as

$$\alpha_{acq}(\theta) = \left(\mu(\theta) - \tilde{R}(\theta^*) - \xi\right)\Phi(Z) - \sigma(\theta)\phi(Z) \quad (12)$$

where Φ and ϕ are the Gaussian cumulative density and probability density functions, respectively.[45] Z is defined as $Z = Z(\theta) = (\mu(\theta) - \tilde{R}(\theta^*) - \xi)/\sigma(\theta)$, with $\mu(\theta)$ and $\sigma(\theta)$ are calculated by Equations (6 and 7). The two terms in Equation (12)

represent the exploitation and exploration weights of the BO, respectively. Their balance is controlled by setting the hyperparameter ξ. As ξ gets higher, BO tends to choose the parameter set in unobserved regions of the search space. BO focuses more on exploitation by testing parameters close to already explored regions as ξ gets lower. In this study, the ξ is set equal to 0.1 to promote exploration over exploitation. Also, the length scales l_c is set for both problems equal one-fourth of the total range of each corresponding parameter following the settings in the previous studies.[20,21]

Data-Driven Magnetic Soft Millirobot Simulation Parameter Tuning: The simulation parameters are tuned $\theta_{sim} = [c_{bond}, c_{collide}, c_{global}, \mu_s, \mu_d]$ to model the robot's behavior accurately. Since the robot's shape deformation directly affects the interaction with its surroundings and how the robot behaves, the optimization goal is set to maximize the similarity between the simulated and actual deformation of the robot. Therefore, the Jaccard Index (JI), commonly used in object detection and image segmentation problems in computer vision, is used to measure the similarity between two frames, i.e., simulated (Figure S5a, Supporting Information) and actual (Figure S5b, Supporting Information). JI is defined by the ratio of overlapped pixels (Figure S5c, Supporting Information) to the union of pixels (Figure S5d, Supporting Information) as follows.

$$JI(F_{sim}, F_{\text{exp}}) = \frac{|F_{sim} \cap F_{exp}|}{|F_{sim} \cup F_{exp}|} \tag{13}$$

where F_{sim} and F_{exp} denote two sample frames of the robot generated by the simulation and physical experiments. Using JI, the reward function is defined as

$$JI: \Theta_{sim} \rightarrow \mathbb{R} \tag{14}$$

which maps the parameter set $\theta_{sim} = [c_{bond}, c_{collide}, c_{global}, \mu_s, \mu_d]$ to scalar reward values, $JI \in [0,1]$. Then, the learning problem given in Equation (10) became

$$\theta^*_{sim} = \underset{\theta_{sim} \in \Theta_{sim}}{\text{argmax}}\ JI(\theta_{sim}) \tag{15}$$

where Θ_{sim} is the complete search space containing all the parameter sets θ_{sim}, and $JI(\theta_{sim})$ denotes the JI for a given θ_{sim}. The reward function is updated in Equation(13) with $JI(\theta_{sim})$ as

$$JI(\theta_{sim}) \sim GP\left(\mu_{sim}(\theta_{sim}), k_{sim}(\theta_{sim}, \theta'_{sim})\right) \tag{16}$$

$\sigma_{n,sim}$ is set in Equation (5) to 0.00 since the JI is calculated without noise. The GP is initialized with a constant prior mean $\mu_{sim} = 0.5$, and signal variance $\sigma^2_{f,sim} = 0.25^2$ so that all the possible values of JI remained inside the 95% confidence interval of the prior.

After defining the hyperparameters of GP-BO and the search space for the simulation parameters, 30 independent learning runs are started in parallel by initializing the GPs with μ_{sim} and $\sigma^2_{f,sim}$. In each independent learning run, eight actuation signal and test result pairs are randomly selected from the training dataset containing 3246 physical tests. Then, the simulation environment is updated with θ_{sim} selected by BO and eight

parallel simulations are run with the selected actuation signals. After completing the simulations, the average JI is evaluated using Equation (13), and the GP model is updated. Until the termination criteria, i.e., 200 iterations, was reached, the learning run is kept iterating by selecting the next θ_{sim}. After completing all the independent learning runs, the best-performing parameter set θ_{sim}^{*} achieving the highest JI is found (Algorithm 1).

The training dataset used for the simulation parameter tuning is available here.

Adaptive Locomotion Learning with Sim-to-Real Transfer Learning: The objective was to design a learning framework to adapt the actuation signal defined by f, B_{max}, α_1, $\Delta\alpha$, and ΔT to maximize the robot's displacement in the forward direction. Hence, the reward function is defined as

$$S: \Theta_{act} \rightarrow \mathbb{R} \tag{17}$$

Which maps the parameter set $\theta_{act} = [f, B_{max}, \alpha_1, \Delta\alpha, \Delta T]$ to scalar reward values, i.e., the stride length of the robot. Using the reward function, the optimization problem Equation (10) is updated as

$$\theta_{act}^{*} = \underset{\theta_{act} \in \Theta_{act}}{\operatorname{argmax}} S(\theta_{act}) \tag{18}$$

where Θ_{act} and θ_{act} denote the complete search space and the parameter set for the actuation signal, respectively, whereas $S(\theta_{act})$ is the average stride length of the robot for a given θ_{act}. As the magnetic soft millirobot does not have an accurate model for its kinematics or dynamics, the reward function is approximated based on the data collected from

physical experiments. In order to include the measurement noises and variations during the experiments into the model, overcome the sparsity in the data, and make probabilistic predictions at unobserved locations of the search space, the reward function $S(\theta_{act})$ is defined by replacing $R(\theta)$ in Equation (13):

$$S(\theta_{act}) \sim GP\left(\mu_{act}(\theta_{act}), k(\theta_{act}, \theta'_{act})\right) \tag{19}$$

To model the measurement noise, $\sigma_{n,act}$ is set to 0.29 based on the previous studies.[20,21] While initializing the GP, two different approaches are employed. In the first one, which is referred to as "standard BO" in the rest of this study, a constant zero mean function is used $\mu_{act} = 0$, i.e., no prior information available about the system and a signal variance $\sigma_{f,act}^2 = (3BL)^2$, so that the highest possible reward value remained in the 95% confidence interval of the prior. In the second approach, referred to as "BO with transfer learning", an exhaustive grid search algorithm testing the robot is first run for all the possible actuation parameter sets in Θ_{act} using the magnetic soft millirobot simulation. Then, the mean function μ_{act} is defined using the simulated data following the study.[21]

After setting the hyperparameters of the GP-BO, defining the search space, and generating simulated data for the given environment, three independent learning runs with physical experiments are run for each optimization approach, i.e., "standard BO", and "BO with transfer learning", for 20 iterations. After completing an independent learning run, the actuation parameter set θ_{act}^* achieving the highest stride length $\tilde{S}$ is found (Algorithm 2).

Task Environments: The proposed adaptive learning strategy is tested for different robots and environmental conditions to show its effectiveness when significant changes happen in the test conditions. In this regard, four different test cases are initially designed with environments having constant profiles: 1) three replicas of the robot (i.e., Robot 1, 2, and 3) in a channel with a channel height equal to 1 BL (Figure 4a), a single robot (Robot 1) in a channel, 2) with changing channel height, $h \in \{0.2, 0.3, 0.65, 1.0\}\ BL$ (Figure 4c), 3) with changing elevation angle $\gamma \in [0,5]°$ (Figure 4e), and 4) filled with different fluids, i.e., air and DPBS (Figure 4g). Next, a more complex and longer environment with changing cross-sectional profiles and elevation angles is designed to further test environment detection and gait adaptation performance (Figure 5).

Domain Recognition: To identify the task environment where the robot is operating, the observed stride length $\tilde{S}$ is compared for a given actuation parameter set θ_{act} to the simulated behavior of the robot (Figure 1) using the KLD value, which is equal to zero for two matching distributions. To calculate the KLD value and evaluate the similarity between two data distributions, i.e., the probability density function of the robot's performance in physical and simulated tests for the given actuation parameter set θ_{act}, separate GP models are first defined ($GP_i, i \in [1, n]$, where n denotes the number of task environments), for each task environment. Then, the stride length measurement is defined as a normal distribution with expected mean and standard deviation equal to the $\tilde{S}(\theta_{act})$, and σ_n, respectively. Next, the stride length ($S_i(\theta_{act})$,

$i \in [1, n]$) for θ_{act} from the GP models GP_i, $i \in [1, n]$. Using these distributions, the KLD value is calculated between the measured and sampled stride lengths for each task environment as follows.

$$D_{KL}\left(P(S|\theta_{act}) \parallel Q(S|\theta_{act})\right) = \sum P(S|\theta_{act}) \ln\left(\frac{P(S|\theta_{act})}{Q(S|Q_{act})}\right) \quad (20)$$

$$D_{KL}\left(P(S|\theta_{act}) \parallel Q(S|\theta_{act})\right) = \log\left(\frac{\sigma_i}{\sigma_n}\right) + \frac{\sigma_n + \left(\tilde{S} - \mu_i\right)^2}{2\sigma_i^2} - \frac{1}{2} \quad (21)$$

where μ_i and σ_i are the mean and standard deviation values sampled from the GP model, respectively. After computing the KLD among the observed and simulated performances, the environment is selected with the minimum value as the robot's working environment.

The accuracy of the domain recognition algorithm is tested with the data collected during the locomotion learning experiments in task environments shown in Figure 4. The first five data points collected through each learning run of standard BO are used for validation since they were chosen with less knowledge. After testing the domain recognition, it was shown that the algorithm could detect the environments with 77.08% accuracy based on a single data point.

Statistical Analysis: All quantitative values were presented as means ± standard deviation. Student's *t*-test was used for the statistical analysis, and statistical significance was set at a 95% confidence level for all tests ($P < 0.05$).

Acknowledgments

This work was funded by the Max Planck Society, and the European Research Council (ERC) Advanced Grant “SoMMoR” Project (grant number 843531). The authors thank Dr. Utku Culha for the initial discussions. S.O.D. thanks the Ministry of National Education of the Republic of Turkiye for the Doctoral Scholarship.

Open access funding enabled and organized by Projekt DEAL.

Author Contributions

S.O.D. contributed to the study design, simulation and algorithm development, experimental procedures, data collection, data analysis, and manuscript writing. M.E.T. participated in the study design and manuscript writing. A.C.K. assisted with the development of the simulation environment. M.S. contributed to the study design, research supervision, and manuscript writing.

References

[1] Z. Ren, W. Hu, X. Dong, M. Sitti, Nat. Commun. 2019, 10.

[2] X. Dong, G.Z. Lum, W. Hu, R. Zhang, Z. Ren, P.R. Onck, M. Sitti, Sci. Adv. 2020, 6.

[3] T. Wang, Z. Ren, W. Hu, M. Li, M. Sitti, Sci. Adv. 2021, 7.

[4] L. Hines, K. Petersen, G. Z. Lum, M. Sitti, Adv. Mater. 2017, 29.

[5] B. Wang, K. Kostarelos, B.J. Nelson, L. Zhang, Adv.Mater. 2021, 33.

[6] Z. Zheng, J. Han, S.O. Demir, H. Wang, W. Jiang, H. Liu, M. Sitti, Adv.Sci. 2023, 10.

[7] M. Pilz Da Cunha, M.G. Debije, A.P.H.J. Schenning, Chem. Soc. Rev. 2020, 49, 6568.

[8] G. Li, X. Chen, F. Zhou, Y. Liang, Y. Xiao, X. Cao, Z. Zhang, M. Zhang, B. Wu, S. Yin, Y. Xu, H. Fan, Z. Chen, W. Song, W. Yang, B. Pan, J. Hou, W. Zou, S. He, X. Yang, G. Mao, Z. Jia, H. Zhou, T. Li, S. Qu, Z. Xu, Z. Huang,Y. Luo, T.Xie, J.Gu, et al.,Nature 2021, 591, 66.

[9] D. Son, H. Gilbert, M. Sitti, Soft Robot 2020, 7, 10.

[10] M. Sitti, D.S. Wiersma, Adv. Mater. 2020, 32.

[11] W. Hu, G.Z. Lum, M. Mastrangeli, M. Sitti, Nature 2018, 554, 81.

[12] R.H. Soon, Z. Yin, M.A. Dogan, N.O. Dogan, M.E. Tiryaki, A.C. Karacakol, A. Aydin, P. Esmaeili-Dokht, M. Sitti, Nat. Commun.2023, 14, 3320.

[13] Y. Wu, X. Dong, J.K. Kim, C. Wang, M. Sitti, Sci. Adv. 2022, 8.

[14] Z. Ren, R. Zhang, R.H. Soon, Z. Liu, W. Hu, P.R. Onck, M. Sitti, Sci. Adv. 2021, 7.

[15] Z. Zheng, H. Wang, S.O. Demir, Q. Huang, T. Fukuda, M. Sitti, Sci. Adv. 2022, 8.

[16] T. Wang, H. Ugurlu, Y. Yan, M. Li, A.M. Wild, E. Yildiz, M. Schneider, D. Sheehan, W. Hu, M. Sitti, Nat. Commun. 2022, 13, 4465.

[17] H. Lu, Y. Hong, Y. Yang,Z. Yang, Y. Shen, Adv. Sci. 2020, 7.

[18] L. Peng, Y. Zhang, J. Wang, Q. Wang, G. Zheng, Y. Li, Z. Chen, Y. Chen, L. Jiang, C.P. Wong, Nano Energy 2022, 99.

[19] C. Wang, Y. Wu, X. Dong, M. Armacki, M. Sitti, Sci. Adv. 2023, 9.

[20] U. Culha, S.O. Demir, S. Trimpe, M. Sitti, in Robotics: Science and Systems XVI, 2020, https://doi.org/10.15607/RSS.2020.XVI.070.

[21] S.O. Demir, U. Culha, A.C. Karacakol, A. Pena-Francesch, S. Trimpe, M. Sitti, Int. J. Rob. Res. 2021, 40, 1331.

[22] C. Schaff, A. Sedal, M.R. Walter, in Robotics: Science and Systems, XVIII, 2022, https://doi.org/10.15607/RSS.2022.XVIII.062.

[23] J. Bongard, V. Zykov, H. Lipson, Science 2006, 314, 1118.

[24] J. Hwangbo, J. Lee, A. Dosovitskiy, D. Bellicoso, V. Tsounis, V. Koltun, M. Hutter, Sci Robot 2019,4.

[25] Y. Yang, K. Caluwaerts, A. Iscen, T. Zhang, J. Tan, V. Sindhwani, in Conference on Robot Learning 2019, pp. 1–10.

[26] Y.H. Lee, H. Lee, H. Kang, J.H. Lee, J.M. Park, Y.B. Kim, H. Moon, J.C. Koo, H.R. Choi, J. Intell. Robot Syst. 2021, 102, 66.

[27] S. Choi, G. Ji, J. Park, H. Kim, J. Mun, J.H. Lee, J. Hwangbo, Sci. Robot 2023,8.

[28] J. Mahler, M. Matl, V. Satish, M. Danielczuk, B. DeRose, S. McKinley, K. Goldberg, Sci. Robot 2019, 4.

[29] M. Lázaro-Gredilla, D. Lin, J. Swaroop Guntupalli, D. George, Sci. Robot 2019, 4.

[30] N. Fazeli, M. Oller, J. Wu, Z. Wu, J.B. Tenenbaum, A. Rodriguez, Sci. Robot 2019, 4.

[31] F. Ficuciello, A. Migliozzi, G. Laudante, P. Falco, B. Siciliano, Sci. Robot 2019, 4.

[32] J. Yao, Q. Cao, Y. Ju, Y. Sun, R. Liu, X. Han, L. Li, Adv. Int. Sys. 2023, 5.

[33] J. Hiller, H. Lipson, Soft Robot 2014, 1, 88.

[34] D.S. Shah, J.P. Powers, L.G. Tilton, S. Kriegman, J. Bongard, R. Kramer-Bottiglio, Nat. Mach. Intell. 2020, 3, 51.

[35] J. Ibarz, J. Tan, C. Finn, M. Kalakrishnan, P. Pastor, S. Levine, Int. J. Rob. Res. 2021,40,698.

[36] R. Zhao, Y. Kim, S.A. Chester, P. Sharma, X. Zhao, J. Mech. Phys. Solids 2019, 124, 244.

[37] Y. Kim, H. Yuk, R. Zhao, S.A. Chester, X. Zhao, Nature 2018, 558,274.

[38] A. Tariverdi, V.K. Venkiteswaran, M. Richter, O.J. Elle, J. Tørresen, K. Mathiassen, S. Misra, Ø.G. Martinsen, Front Robot AI 2021, 8, 631303.

[39] N. Heess, D. Tb, S. Sriram, J. Lemmon, J. Merel, G. Wayne, Y. Tassa, T. Erez, Z. Wang, S.M.A. Eslami, M. Riedmiller, D.S. Deepmind, arXiv preprintarXiv:1707.02286 2017.

[40] H. Daguerre, S.O. Demir, U. Culha, F. Marionnet, M. Gauthier, M. Sitti, A. Bolopion, IEEE/ASME Transactions on Mechatronics 2022,27, 3506.

[41] A.Von Rohr, S. Trimpe, A. Marco, P. Fischer, S. Palagi, in IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2018, pp. 6199–6206.

[42] M. Neumann-Brosig, A. Marco, D. Schwarzmann, S. Trimpe, IEEE Transactions on Control Systems Technology 2019, pp. 730–740.

[43] D. Duvenaud, H. Nickisch, C.E. Rasmussen, in Proceedings of the 24th International Conference on Neural Information Processing Systems, Curran Associates Inc, Granada, Spain 2011, pp. 226–234.

[44] D.R. Jones, M. Schonlau, W.J. Welch, Journal of Global Optimization 1998, 13, 455.

[45] E. Brochu, V.M. Cora, N. de Freitas, arXiv preprint arXiv:1012.2599 2010.

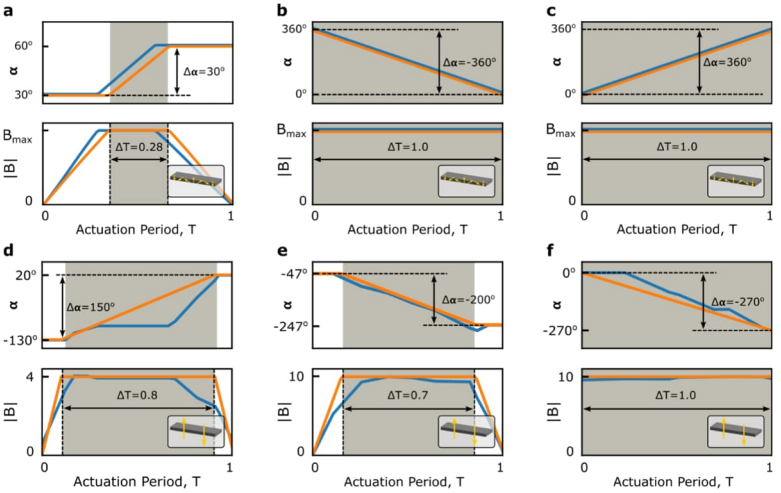


Figure S1 Comparison of previously reported actuation signals to the signals defined by the proposed generic actuation signal. The top row shows the manually defined actuation signals generating (a) walking, (b) rolling, and (c) crawling locomotions.[11] The bottom row shows the actuation signals learned by deep reinforcement learning (DRL) under a magnetic field (d) with a maximum field strength of $4mT$, and (e, f) with a maximum field strength of $10mT$.[32] Reference and replicated signals are shown in blue and orange, respectively. In the bottom row, magnetic field strength |B| is given in mT, and the yellow arrows represent the magnetic profile of the robots, for which the DRL learned the actuation signal.

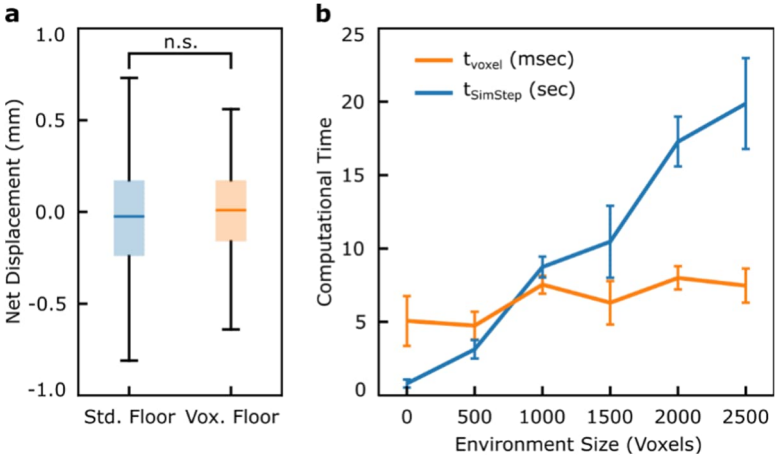


Figure S2 Evaluation of the effect of multi-body interaction on the simulation performance regarding accuracy and computational time. (a) To evaluate the effect of multi-body interaction on the simulation output, the robot's motion was simulated on a flat surface with 400 randomly generated, distinct actuation signals. The floor was modeled by the default floor definition available in *Voxelyze* shown as 'Std. Floor', and by voxels shown as 'Vox. Floor'. No significant difference was observed between two methods (*t*-test, $P > 0.05$). (b) The robot's motion was simulated on a floor defined by different numbers of voxels to evaluate the effect of multi-body interaction on the computational time with 400 randomly generated, distinct actuation signals. The orange and blue trendlines show the average time per voxel, t_{voxel} (msec) and per simulation step, $t_{SimStep}$ (sec), respectively. Each data point reprensents the mean of 400 simulation runs and the error bars show the standard deviation.

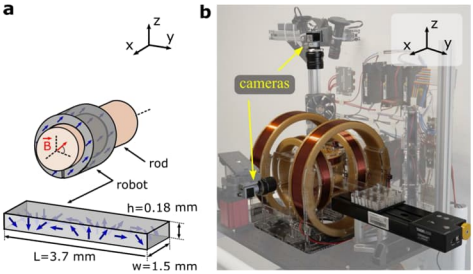


Figure S3 Fabrication and actuation of the magnetic soft millirobot. (a) Magnetic soft millirobot composed of homogeneously distributed non-magnetized ferromagnetic microparticles inside a silicone elastomer sheet was rolled around a cylindrical road and magnetized with $|B| = 1.8T$ field (red arrow) with a 45° angle with respect to the y-axis. The unfolded robot maintained a periodic magnetization profile (blue arrows) along its body. (b) Helmholtz coil setup with three electromagnetic coil pairs allowed generating a homogeneous magnetic field in 3D space up to 12mT. Two high-speed cameras with front and top views were used to observe and evaluate the robot's motion in real-time. The test area with the homogeneous magnetic field on the y-axis was enlarged by the linear stage attached to the xy-plane.

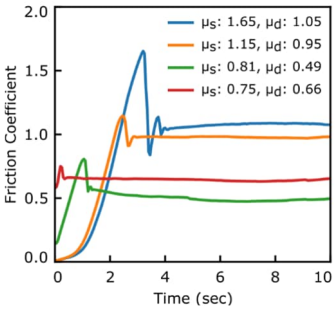


Figure S4 Friction coefficient measurements. The measurements were repeated for four samples cut from the same magnetic soft millirobot material batch on a flat paper surface. The measurements were taken with 600 gr of weight placed on the sample having 20 mm diameter and 0.185 mm thickness, while the sample and the weight were pulled on the test surface with a constant speed of 0.83 mm/sec.

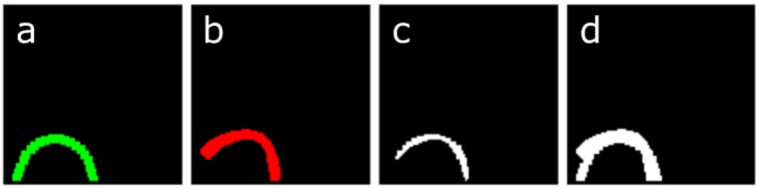


Figure S5 Visualization of the Jaccard index calculation steps (a) Simulated frame, F_{sim}, (b) actual frame, F_{exp}, (c) intersection of frames, $F_{sim} \cap F_{exp}$, and (d) union of the frames, $F_{sim} \cup F_{exp}$.

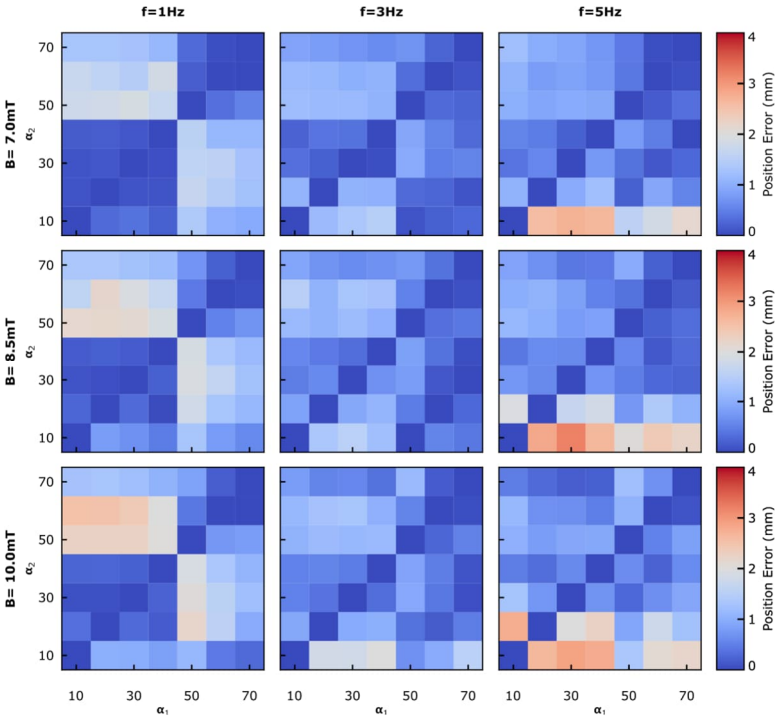


Figure S6 Average stride length error between simulated and experimental results for 582 test cases.

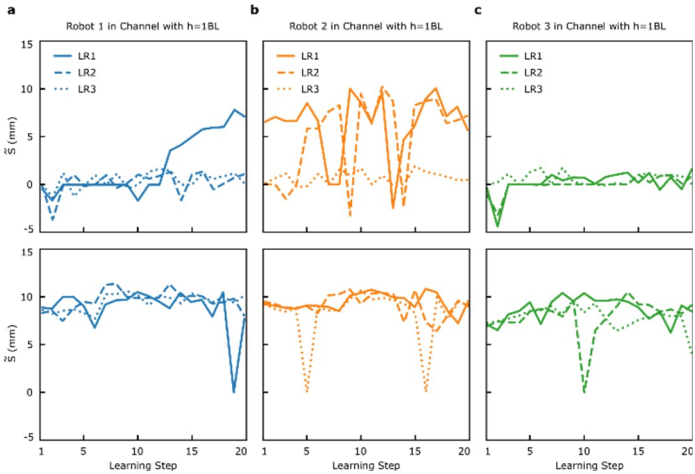


Figure S7 Learning the controller parameters for three different robots (i.e., Robot 1, 2, and 3) in a channel with 1 BL height (Figure 4a) within 20 physical experiments in 3 independent learning runs (i.e., LR1, LR2, and LR3). The learning performance of standard BO (upper row) was compared to BO with transfer learning (lower row) for each robot.

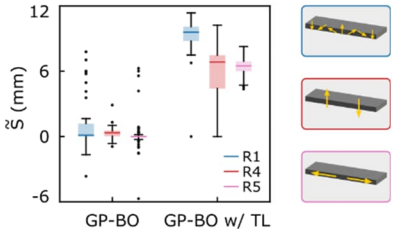


Figure S8 Learning the controller parameters for three robots with the same structural properties and different magnetization profiles (i.e., Robots 1, 4, and 5) in a channel with 1 BL channel height (Figure 4a). Box plot shows the overall performance of the learning approaches as a standard interquartile range (IQR) method, where the horizontal lines are the median of the observed stride lengths $\tilde{S}$ in 60 physical trials for each robot. The box around the median line shows the upper and lower quartiles. The error bars and dots represent the highest and lowest performances and outliers, respectively.

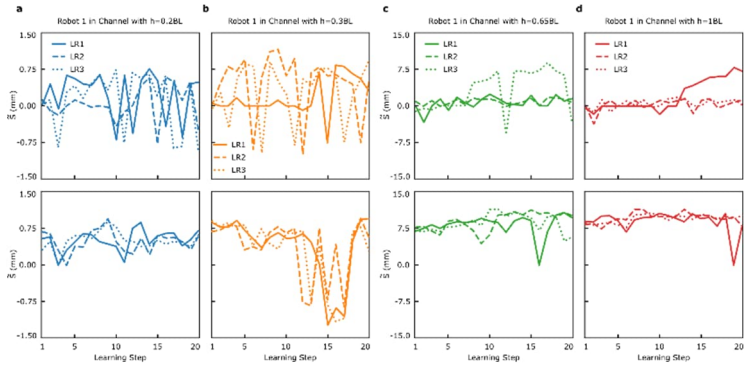


Figure S9 Learning the controller parameters for Robot 1 in a channel with changing channel height $h \in [0.2, 1.0]BL$ (Figure 4c) within 20 physical experiments in 3 independent learning runs (i.e., LR1, LR2, and LR3). The learning performance of standard BO (upper row) was compared to BO with transfer learning (lower row) for each channel height.

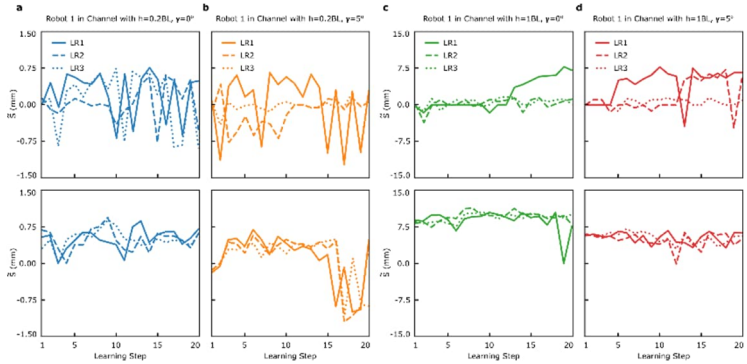


Figure S10 Learning the controller parameters for Robot 1 in a channel with changing channel height $h \in \{0.2, 1.0\}BL$ and elevation angle $\gamma \in \{0, 5\}°$ (Figure 4e) within 20 physical experiments in 3 independent learning runs (i.e., LR1, LR2, and LR3). The learning performance of standard BO (upper row) was compared to BO with transfer learning (lower row) for each channel height and elevation angle.

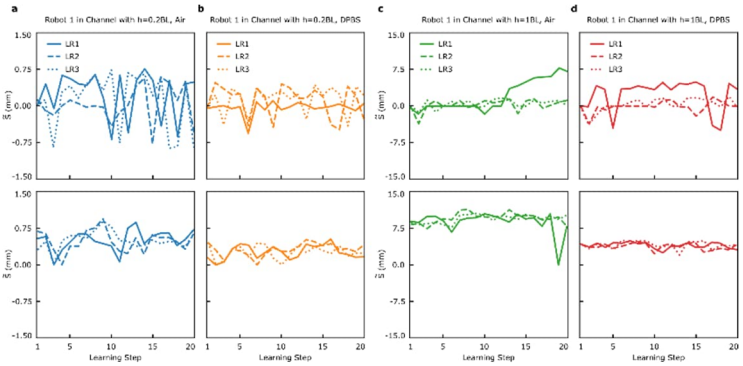


Figure S11 Learning the controller parameters for Robot 1 in a channel with changing channel height $h \in \{0.2, 1.0\}BL$ and the surrounding fluid, i.e., air and DPBS, (Figure 4g) within 20 physical experiments in 3 independent learning runs (i.e., LR1, LR2, and LR3). The learning performance of standard BO (upper row) was compared to BO with transfer learning (lower row) for each channel height and surrounding fluid.

Table S1 Performance of the controller parameters learning for three different robots (i.e., Robot 1, 2, and 3) in a channel with 1 BL height (Figure 4a) within 20 physical experiments in 3 independent learning runs.

Robot Name	Learning Approach	Performance (mm)	
		Median	IQR
Robot 1	Standard BO	0.12	1.19
	BO with TL	9.58	1.32
Robot 2	Standard BO	5.16	7.26
	BO with TL	9.22	1.25
Robot 3	Standard BO	0.03	0.92
	BO with TL	8.39	1.64

Table S2 Performance of the controller parameters learning for Robot 1 in a channel with changing channel height $\boldsymbol{h \in [0.2, 1.0]BL}$ (Figure 4c) within 20 physical experiments in 3 independent learning runs.

Channel Height (BL)	Learning Approach	Performance (mm)	
		Median	IQR
0.20	Standard BO	0.18	0.58
	BO with TL	0.49	0.24
0.30	Standard BO	0.47	0.69
	BO with TL	0.61	0.45
0.65	Standard BO	0.93	1.67
	BO with TL	8.81	2.71
1.00	Standard BO	0.12	1.19
	BO with TL	9.58	1.32

Table S3 Performance of the controller parameters learning for Robot 1 in a channel with changing channel height $\boldsymbol{h \in \{0.2, 1.0\}BL}$ and elevation angle $\boldsymbol{\gamma \in \{0, 5\}^\circ}$ (Figure 4e) within 20 physical experiments in 3 independent learning runs.

Channel Elevation Angle	Channel Height (BL)	Learning Approach	Performance (mm)	
			Median	IQR
0.00°	0.20	Standard BO	0.18	0.58
		BO with TL	0.49	0.24
	1.00	Standard BO	0.12	1.19
		BO with TL	9.58	1.32
5.00°	0.20	Standard BO	0.00	0.34
		BO with TL	0.33	0.42
	1.00	Standard BO	0.95	5.50
		BO with TL	5.61	1.64

Table S4 Performance of the controller parameters learning for Robot 1 in a channel with changing channel height $\boldsymbol{h \in \{0.2, 1.0\}BL}$ and the surrounding fluid, i.e., air and DPBS, (Figure 4g) within 20 physical experiments in 3 independent learning runs.

Surrounding Fluid	Channel Height (BL)	Learning Approach	Performance (mm)	
			Median	IQR
Air	0.20	Standard BO	0.18	0.58
		BO with TL	0.49	0.24
	1.00	Standard BO	0.12	1.19
		BO with TL	9.58	1.32
DPBS	0.20	Standard BO	0.09	0.29
		BO with TL	0.28	0.22
	1.00	Standard BO	0.00	2.44
		BO with TL	4.23	0.86

Publication 4

A Localization Method for Untethered Small-Scale Robots Using Electrical Impedance Tomography

Abstract

Untethered small-scale robots can be potentially used in medical applications, such as minimally invasive surgeries and targeted drug delivery. This article introduces a new localization method using electrical impedance tomography, which is an emerging medical imaging technique, to dynamically track small-scale robots. The proposed approach provides the electrical conductivity distribution within the robot workspace from a set of electrical stimulations and voltage measurements gathered from eight electrodes placed at its boundary. The position of the robot can be deduced from the conductivity map that is reconstructed with the contrast in electrical properties between the robot and the background medium. This method is experimentally validated by successfully tracking the 2-D motion of four different magnetically actuated robots within a cylindrical arena (30 mm in diameter and 4.2 mm high). The smallest detected robot is $1.5\times1.5\times1\text{cm}^3$. The proposed tracking method provides a noninvasive technology with low-cost and high-speed potential that would be significant and useful for the position feedback control of untethered devices for biomedical applications in the future.

I. Introduction

Small robots (i.e., robots < 1 cm) capable of navigating in a controlled manner through confined and enclosed spaces have been extensively studied in recent years to enable applications in various domains, such as micromanipulation [1], [2] and healthcare [3], [4]. Among magnetic [2], optical [5], acoustic [6], and biological [7] actuation approaches used for mobile small-scale robots, the magnetic actuation method becomes more

prominent for medical applications due to its potential usage inside nontransparent environments and deep penetration capability in any nonmagnetic media [8]. After a rapid development in the past decade and the demonstration of applications in optically transparent environments, the scientific interest in deploying such small-scale robots into other workspaces is strong and growing [9]. However, ensuring the efficiency of delicate tasks at small scale requires a precise control over robot motion and positioning. Achieving such a fine and consistent control in spite of perturbations and model uncertainties implies the need of a position feedback. Therefore, the potential of untethered small robots to become a real breakthrough in applied medicine strongly depends on the ability to provide suitable localization techniques.

To address this challenge, several methods have been studied in the past few years. These methods are extensively reviewed by recent articles reflecting the interest of the community in the subject [10]–[12]. Various technologies that are well established in biomedical imaging have been explored for the tracking of small-scale robots, such as magnetic-field-based techniques [13]–[15], ultrasound [16], [17], optical techniques [18], [19], and ionizing-radiation-based methods [20], [21]. Although these solutions are promising, a large part of the existing technologies still exhibit some limitations in terms of cost, scanning speed, penetration depth, and adverse health effects [10]–[12].

Alternatively to these conventional methods, the exploitation of electrical impedance variations has recently shown to be relevant for localization purposes at small scale [22], [23]. At the macroscopic level, simulation

results from Snyder et al. [24] suggest that electrical impedance tomography (EIT) would be of interest to track underwater objects. EIT exploits electrical stimulations and measurements made through the surface of an object to noninvasively image the interior of a domain by estimating the internal distribution of the electrical properties. The underlying principles of this technique were first applied in geophysical exploration to detect conductive ore or liquids in the ground [25]. Thereafter, EIT has been used in medicine to image various physiological phenomena, such as breathing [26], cardiac function [27], or brain activity [28]. These applications show that EIT is a safe imaging technique, unlike the prolonged use of ionizing radiations (e.g., X-rays). It also illustrates the possibility to image deep tissues, contrary to optical techniques that are limited to superficial areas [10]. Moreover, EIT systems can provide high-speed imaging [29] with low cost [30] and high portability [31] (i.e., no cumbersome equipment). These aspects make EIT an attractive technique toward microrobot localization in the context of biomedical applications.

While the exploitation of electrical impedance variations in robotics had been so far limited to tactile human–robot interactions [32], [33] and on-board deformability sensing [34], [35], this article describes a new robot tracking system based on EIT. The possibility to localize untethered small-scale robots in 2-D using EIT is demonstrated through in vitro experiments, where the actuation of the robot is achieved by an external magnetic field and a camera is used to verify the tracking performance. It is shown by static experiments that the EIT tracking system and the applied magnetic field do not interfere. Moreover, to evaluate the

influence of the size of the robot on the tracking accuracy, multiple robots in changing sizes are fabricated and tested. The primary contributions of this work are the following:

1) The first use of EIT to track the dynamic motions of small-scale robots.
2) A tracking system decoupled from magnetic actuation.
3) A method suitable for localization in nonhomogeneous environments.
4) The study of the tracking accuracy and the determination of the smallest detectable robot.

The rest of this article is organized as follows. Section II introduces the working principle of the EIT technique. In Section III, the experimental platform, including the robot design, actuation, and optical tracking system, is presented. Section IV provides a detailed description of the proposed EIT localization system. Section V introduces the results of the robot tracking experiments and analyzes the accuracy of the developed system according to the size of the robot, EIT reconstruction parameters, and different environments. Section VI discusses the capacity of the proposed system and presents challenges and future possibilities in the use of EIT technology for monitoring and tracking micromachines.

II. Electrical Impedance Tomography

Tomography is a technique used to estimate the internal structure of an object from signals emitted and measured outside the analyzed domain [36]. Among various tomographic methods based on different physics,

EIT relies on electricity to noninvasively reconstruct the interior of the target. The electrical measurements and excitations (i.e., voltages or currents) are typically performed by means of multiple electrodes placed on the surface of the system to be imaged. This set of signals is then used to determine the distribution of the absolute conductivity (absolute EIT) within a 2-D or 3-D domain, or the distribution of the conductivity changes with respect to a reference state (difference EIT) [36]. This section presents the theoretical fundamentals of EIT. It enters the category of inverse problems, which require finding the model parameters (causes) that correspond to the observations (effects) as well as possible (see Fig. 1).

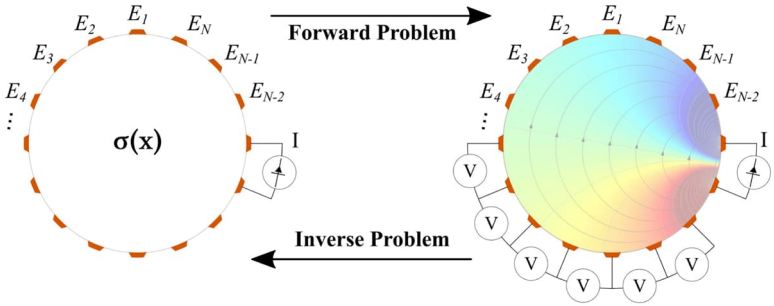


Figure 1. Principle of EIT. An electrical stimulation is applied through the boundary of the studied domain, and the resulting voltages are measured. The forward problem provides the measurements from the knowledge of the medium properties, while the inverse problem estimates the medium properties from the measurements.

A. EIT Forward Problem

The forward problem reflects the influence of the model parameters on the measurements. It determines the electrical voltages V_i measured at different positions i on the boundary of the system from the knowledge of the internal conductivity distribution σ and the stimulating current injection.

From Maxwell's equations, it is possible to show that the electrical potential u inside the studied domain Ω is governed by [37]

$$\nabla.\gamma(x,\omega)\nabla u(x) = 0, \quad (1)$$

where x is a point in the domain Ω, ω is the angular frequency of the applied current, and γ is the electric admittivity given by $\gamma(x,\omega) = \sigma(x,\omega) + i\omega\varepsilon(x,\omega)$, where σ is the electric conductivity and ε is the permittivity. In the conventional frequency range at which EIT systems operate, the imaginary part of the admittivity is usually negligible [36]. Therefore, γ is approximated to the conductivity σ.

In addition to (1), the formulation of the forward problem includes conditions on the current density j, which is related to the potential by Ohm's law $j = \sigma\nabla u$, and satisfies

$$\int_{\partial\Omega} j dS = 0, \quad (2)$$

which implies that electrical charges do not accumulate locally within the medium. In other words, all the currents injected into Ω necessarily flow out.

Additional boundary conditions are set to account for the known electrical stimulation at given electrodes. Their formulation depends on the modeling of the electrode–medium interface. This interface can be modeled in a multitude of ways with varying degrees of precision and complexity [38].

The forward problem can be solved using standard numerical methods including the finite-element method [39]. The goal in EIT is to find a conductivity distribution in Ω such that the resulting forward solution is as close as possible to the signals measured experimentally. This corresponds to solving the inverse problem.

B. Inverse Problem and Image Reconstruction

Solving the inverse problem requires finding a value of the conductivity σ that minimizes the data mismatch between the electrical measurements y and their estimates via the forward solution $F(\sigma)$ [36]

$$\|y - F(\sigma)\|^2. \tag{3}$$

However, unlike the forward problem, the EIT inverse problem is ill-posed, which means that it is particularly complex to solve because of the nonuniqueness and the instability of the solution [38].

Although the determination of an inverse solution remains an open subject, various resolution approaches have been proposed [40]. The most popular is based on the linearization and the regularization of the problem to get an approximated well-posed problem. The equation to minimize becomes [36]

$$\|y - F(\sigma)\|_W^2 + \lambda^2 \|\sigma - \sigma_0\|_Q^2, \quad (4)$$

where W is a data weighting matrix representing the inverse covariance of measurements, σ_0 is an a priori estimate of the solution, Q is the regularization matrix that can contain certain prior assumptions about the solution, and the norm represented as $\|a\|_A^2$ is equal to $a^T A a$. The regularization process introduces extra information to promote some credible solutions. It involves a tradeoff between the prior solution and the exact solution based on the measured data. The balance of this tradeoff is controlled by the hyperparameter λ.

Interestingly, the regularization process can also be defined such that both spatial and temporal information are added. This is possible in the case of nonstationary applications, when successive frames are not independent provided that the acquisition rate is sufficient with respect to the dynamic of the observed phenomenon. Such a spatiotemporal resolution approach has been developed in [41].

This is particularly relevant for robot tracking purposes (see Section V). Herein, the solver developed in [41] is utilized for the proof of concept of an EIT-based position sensor for untethered small-scale robots.

III. Experimental Robot System

The applicability of the proposed EIT-based tracking method on the untethered small-scale robots is tested by using the experimental setup and the small-scale robot design shown in Fig. 2.

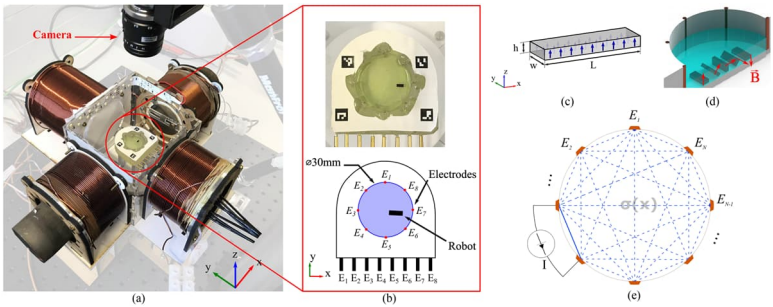


Figure 2. (a) Photo of the magnetic actuation system with five electromagnetic coils (four of them placed on the XY plane, while the fifth is below the test area), the test area surrounded by the eight electrodes of the EIT system, and a high-speed camera that allows generating the ground truth data. (b) Top-view photo and drawing of the test area showing the electrode positions and a sample robot in the test fluid. (c) Uniform magnetization profile of the magnetic small-scale robot along its body (blue arrows). (d) Sample movement of the magnetic small-scale robot by applying a rotating magnetic field B shown by red arrows. (e) Illustration of the full-scan EIT stimulation scheme. All electrode pairs are successively selected as driving electrodes.

A. Robot Design, Fabrication, and Gait Definition

The magnetic small-scale robots used in this study are fabricated following the methods reported in [42] and [43]. To start with, Ecoflex 00-30 (Smooth-On Inc.) and neodymium–iron–boron (NdFeB) magnetic particles with 5-µm diameter (MQFP-15-7, Magnequench) are mixed with a 1:1 body mass ratio. Then, the prepolymer mixture is poured on a methyl methacrylate plate, and two sets of six robots (dimensions are provided in Table I) are cut out of the cured polymer sheet using a high-resolution laser cutter (LPKF Protolaser U4). The first set of robots are kept nonmagnetized and used to avoid any influence of the robot's motion in the experiments studying the impact of the magnetic field on the EIT

data (see Section V-A). The second set of robots, on the other hand, are put into flat molds and magnetized within a magnetic field with a magnitude of 1.8 T pointing parallel to the vertical axis of the robots. Once the robots are taken out of the molds, the magnetic particles maintain their magnetization orientation forming a uniform profile along the longitudinal axis of the robot body as shown in Fig. 2(c).

To test the accuracy of the proposed tracking method for dynamic cases in Section V-B, fabricated robots are magnetically actuated to move in the workspace by rolling locomotion. Each cycle of the applied rotating magnetic field B around an axis orthogonal to the vertical axis of the robot (see Fig. 2(d)) generates a complete turn of the robot around the same axis. Direction of robot's motion is controlled by manipulating the angle between the rotation axis of the B field and the x-axis of the workspace.

Table I Dimensions of the small-scale robots

	L (mm)	**w (mm)**	**h (mm)**	**V (mm^3)**
R1	5.23	2.12	1.00	11.09
R2	3.70	1.50	1.00	5.55
R3	2.77	1.12	1.00	3.10
R4	1.50	1.50	1.00	2.25
R5	1.12	1.12	1.00	1.25
R6	0.78	0.75	1.00	0.56

B. Actuation and Feedback Setup

As the aim of this study is to use EIT to dynamically track untethered magnetic small-scale robots, the workspace is designed as a circular arena with 30 mm diameter considering the average abdomen size of the mice

used in preclinical studies. It is surrounded by eight electrodes positioned equidistantly as shown in Fig. 2(b). After filling the test area with 3 mL of phosphate-buffered saline (PBS, P-5368 with pH 7.4, SIGMA), corresponding to a liquid level of 4.24 mm, which is twice the width of the biggest robot in Table I, the magnetic small-scale robot is placed into this solution. To apply an external magnetic field, the test area is placed in the center of the magnetic coil setup (see Fig. 2(a)) that can generate a 3-D magnetic field within a $4\times4\times4\text{cm}^3$ workspace with a maximum strength of 15 mT. The magnetic field is modulated in the 3-D space by controlling the electric currents running through the electromagnetic coils via motor driver units (SyRen25) and an Arduino microcontroller running at 1.2 kHz. To maintain reliable and repeatable experiments, the mapping between the applied electric currents and the generated magnetic field is calibrated regularly.

The ground truth data required to evaluate the accuracy of the proposed tracking method are collected by a high-speed camera (Basler aCa2040-90uc, shown in Fig. 2(a)) running at 90 frames/s and positioned orthogonal to the xy plane of the workspace. Collected data (see Fig. 2(b)) is first processed to correct the distortions and misalignments in the images, and then, the position of the robot's centroid is extracted as the ground truth data.

A master PC is used to run the image processing code and the robot operating system, which handles all the communication tasks between different elements of the experimental setup (e.g., image capture, electric

current control and EIT data collection) and allows clock synchronization between all these components.

IV. EIT Tracking System

A. Stimulation and Measurement Scheme

As described in Section II, EIT reconstructs the internal conductivity distribution of the test domain from electrical measurements taken at its surface for a given electrical stimulation. The stimulation pattern influences the measured data and thereby the reconstructed images. For the sake of simplifying the implementation of EIT, drive patterns involving a single current source and sink can be preferred to strategies using multiple current sources simultaneously [38]. Bipolar drive patterns consist of injecting the current through a pair of electrodes while taking potential measurements at the remaining adjacent electrode pairs. In this article, the recently developed full-scan scheme [44] is used (see Fig. 2(e)). This scheme uses successively all possible stimulation electrode pairs to maximize the number of measurements. It has been shown to provide improved EIT images compared to the conventional adjacent and opposite strategies that rotate the current carrying electrodes through all successive adjacent and opposite electrode pairs, respectively [45].

B. EIT Instrumentation

As sketched in Fig. 2(b) and (e), the workspace is surrounded by eight evenly distributed electrodes, determined by the number of analog input channels available on the data acquisition system. The electrodes are 0.8 mm diameter rods made out of platinum and connected by shielded wires

to a custom-designed electronic platform (see Fig. 3(a)) integrating two 8-to-1 multiplexers (ADG708, Analog Devices, Wilmington, MA, USA). This multiplexing circuit allows us to select different combinations of current carrying electrode pairs (source and sink) to execute the entire stimulation and measurement strategy. As shown in Fig. 3(a), the multiplexers are controlled by means of the digital output channels of an I/O module (PXIe-7847R, National Instruments, Austin, TX, USA) integrating a field-programmable gate array (FPGA) system. The analog output and input channels of the device are, respectively, dedicated to the generation of the electrical stimulation and the simultaneous acquisition of the resulting electrical potentials. As previously demonstrated in [44], the stimulation signal can be an ac voltage, while the current is measured by means of a sense resistor. In the proposed system, this stimulation signal has an amplitude of 1 V and a frequency of 10 kHz. The FPGA is controlled by a real-time module (PXIe-8840 Quad-Core, National Instruments, Austin, TX, USA) and sends the recorded data to this real-time target through a direct memory access buffer. Both the I/O module and the real-time target are placed in a PXIe-1078 chassis (National Instruments, Austin, TX, USA).

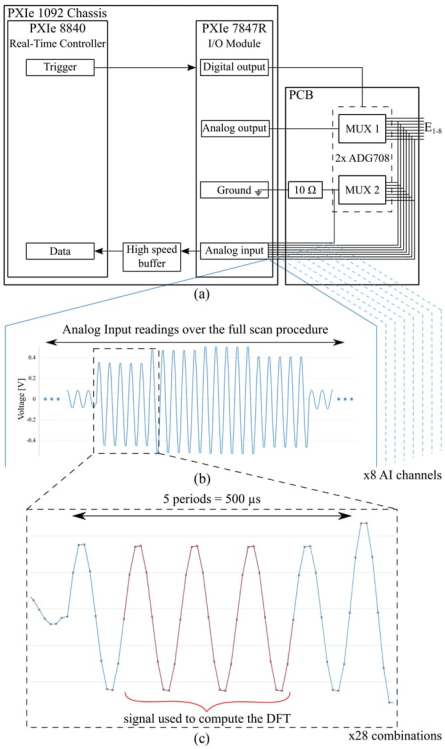


Figure 3. (a) Instrumentation used in the proposed EIT system. A real-time operating system controls the FPGA-based I/O module. This module both generates the stimulation sequence and measures the electrical potential of the electrodes. A 10 Ω sense resistor is used to deduce the stimulating current and normalize the measurements. All the data are sent to the host system through a high-speed buffer. (b) Typical analog input readings recorded in the experiment. The combination of driving electrodes is changed every 500 µs, which corresponds to five periods of the signal. The amplitude of the measured potentials changes accordingly. (c) Sampling rate is 100 kHz, which provides ten data points per period. On the whole, 50 data points are taken at each electrode for each combination of driving electrode pairs. To avoid any influence of possible transient effects, the DFT is calculated over the three central periods only.

C. Signal Acquisition Procedure

While the electrical stimulation is continuously generated by the analog output of the I/O module, the digital outputs were controlled by the built-in blocks in a LabView environment to switch from a stimulation/measurement pattern to the other every 500 µs. This corresponds to five periods of the excitation signal for each of the 28 different stimulation combinations provided by the full-scan strategy in an eight-electrode EIT system. In parallel, the electrical potential measurements are performed via the analog input channels at a sampling rate of 100 kHz (see Fig. 3(b) and (c)).

D. Data Processing

The data are subsequently processed using the built-in libraries in the open-source software EIDORS [39]. For each stimulation combination, the discrete Fourier transform (DFT) of the recorded signal is determined over the three central periods of the recording [see Fig. 3(c)]. The first and last periods are ignored to avoid any influence of possible transient effects due to the switching from a stimulation electrode to another. The filtering of the signal is performed by taking only the element of the DFT that provides the amplitude corresponding to the excitation frequency [44]. As the amplitude of the adjacent voltages involving a current carrying electrode can be biased by the electrode polarization, these data are discarded in EIDORS. In addition, as the first analog input channel of the I/O module is used to monitor the voltage across the sense resistor and deduce the flowing current, the adjacent voltages involving E1 are not measured. Therefore, in each stimulation combination, two, three,

four, or five relevant output adjacent voltages can be obtained. These voltages are normalized by the amplitude of the injected current. Going over the 28 stimulation combinations, a total of 90 (= 2 × 5 + 8 × 4 + 12 × 3 + 6 × 2) normalized voltages are obtained and taken as input of the EIT image reconstruction process.

As this article uses difference EIT to reconstruct conductivity changes rather than the absolute conductivity map, two datasets are needed for image reconstruction. One set is taken before the robot is inserted into the workspace, which will be referred as 'homogeneous dataset'. To reduce the influence of measurement noise in this reference dataset, it is averaged over 356 vectors of data points taken from the homogeneous medium. The other set is taken from the inhomogeneous system containing the small-scale robot, which will be referred as 'inhomogeneous dataset'.

The homogeneous and inhomogeneous datasets are taken as input of the reconstruction algorithm developed in [41] and implemented in the open-source package EIDORS [39]. This solver involves different image reconstruction parameters, which are indicated and studied in the next section. In the images generated by the reconstruction code, the workspace is discretized into multiple elements. An estimated conductivity change σ_{el} is assigned to each element, and the image is generated by converting the conductivity values to a colormap.

Since the robot is known to be more resistive than the PBS solution, it is expected to correspond to a region with a highly negative conductivity change. To improve its distinguishability in the image, a threshold is imposed to the conductivity change values of the elements appearing in

the reconstructed image. In other words, only significant negative conductivity changes are shown in the image, while the estimated conductivity changes σel that are positive or above the threshold are set to 0 (no color). The threshold δ_σ is defined as

$$\delta_\sigma = \overline{\sigma_{el}} - Sd_\sigma \quad (5)$$

where $\overline{\sigma_{el}}$ is the average of the estimated conductivity value over all the elements composing the workspace and Sd_σ is the standard deviation. The thresholded image is then binarized to mark the significant elements. The center of mass of the robot is detected on the binarized image by simple shape detection after basic image processing (dilation and erosion). Remaining small isolated regions (significant negative conductivity changes) are ignored by just selecting the biggest blob in the image.

The next section describes the experiments performed to evaluate the potential of this EIT-based system for the 2-D localization of a small-scale robot and presents the results of this first proof of concept.

V. Results

A. Effect of External Magnetic Field on the EIT-Based Localization System

As a preliminary study, the compatibility of the proposed EIT-based localization approach with the magnetic actuation system has been investigated. To identify the influence of the external magnetic field on the voltage readings, the tests have been performed using the nonmagnetized set of robots (i.e., the first set of robots) whose sizes are

given in Table I. Accordingly, the influence of the magnetic field on the reconstructed EIT image is analyzed while avoiding any robot motion due to magnetic actuation.

For this purpose, four different test cases were studied: 1) $|B| = 0mT$, without robot; 2) $|B| > 0mT$, without robot; 3) $|B| = 0mT$, with a robot; and 4) $|B| > 0mT$, with a robot. In these experiments, the external magnetic field is defined as a rotating field with five different amplitude and frequency sets from 3 to 9 mT and from 1 to 10 Hz. Moreover, throughout the experiments including a robot, the robot is placed into five different positions inside the test area (see Fig. 4(a)–(c) and (g)–(i)) and each experiment is repeated five times. All EIT images are reconstructed using the solver in [41]. The resolution parameters used are $p = 0.5$, $\gamma_c = 0.9$, $\lambda = 0.05$, and $d = 5$, whose influence is explained and studied in the next subsection.

Two metrics have been used to explore the impact of the magnetic field on the EIT-based localization system. First, in the experiments with the robot, the position estimations obtained from the EIT images are compared to the ground truth data obtained from the camera (see Fig. 4). The position detection accuracy of the EIT system is reported in Table II with median and interquartile range (IQR) for each robot with B-field ($B = 3mT$ and $f = 1Hz$) and no B-field. The camera and EIT images for each robot and the position detection accuracy for the remaining B-field cases are presented in the supplementary material. Results for both test cases show that the average localization error is the lowest for R2, and it increases as the size of the robot gets smaller, except for R1 and R6.

Moreover, due to their small size, the localization of R5 and R6 results in high error values considering the size of the workspace.

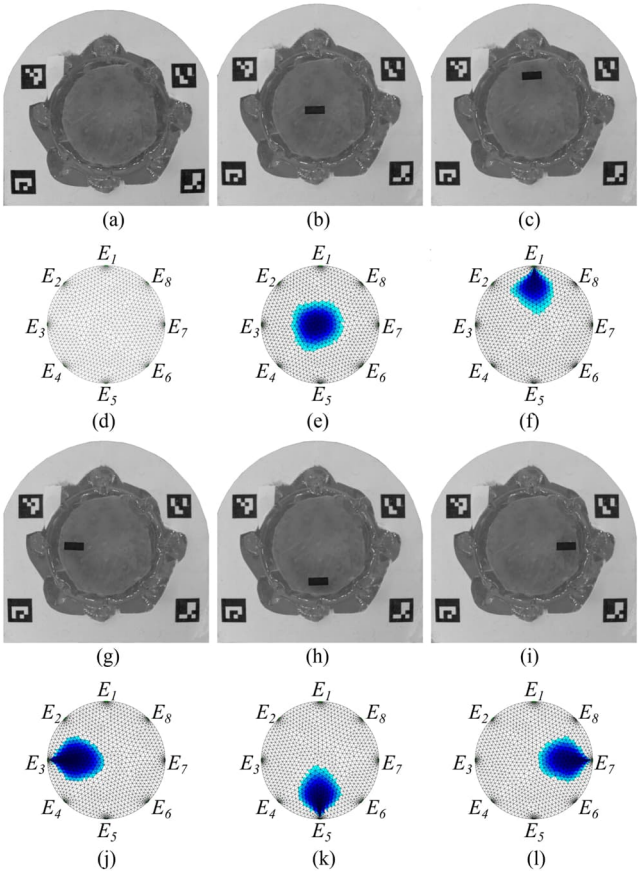


Figure 4. Actual images (first and third rows) and the reconstruction results generated by the collected EIT data (second and fourth rows) for the static tests done with the nonmagnetized robot R1 in Table I and |B| = 0 mT. For all robots with and without external magnetic field, camera and EIT images are presented in the supplementary material.

Table II Position detection accuracy of the EIT system for each nonmagnetized robot in Table 1 throughout the static tests shown in Fig. 4.

	No B-Field		B-Field (3mT 1Hz)	
Robot	**Median (mm)**	**IQR (mm)**	**Median (mm)**	**IQR (mm)**
R1	0.61	0.08	0.55	0.26
R2	0.38	0.21	0.45	0.40
R3	0.56	0.33	0.54	0.51
R4	0.70	0.61	0.58	0.74
R5	8.66	0.42	6.63	8.40
R6	3.80	13.03	7.00	4.96

Second, the estimated average conductivity variation in the workspace and the conductivity threshold values (δ_σ) are computed for all test cases. Table III reports the obtained values. The comparison of δ_σ values for each test case with Case 1 given in the last column shows that application of the external magnetic field without a robot in the test area causes a maximum increase in the conductivity threshold value by 268.76%, whereas the existence of the robot without external magnetic field causes 531.26% increase. Therefore, the impact of the magnetic field on the reconstructed EIT images is shown to be smaller than the impact of the robot itself. Furthermore, the existence of the robot minimizes any visible effect of the external magnetic field on δ_σ.

To summarize, the results in Tables II and III show that the EIT-based localization system is barely affected by the magnetic field, whose influence on the reconstructed EIT images is minor compared to the presence of the small-scale robot itself. Given the current results, the

decoupling could be explained by the four-order of magnitude difference between the frequency of the actuation signals and the frequency of the electrical signals used for localization (1 Hz and 10 kHz, respectively).

Table III Average conductivity and conductivity threshold values for the four different test cases used to identify the influence of magnetic field on the EIT system.

B-Field	Robot	Avg. Conductivity ($\overline{\sigma_{el}} \pm Sd_{\sigma}$)	Cond. Threshold δ_{σ}	Rel. Cond. Threshold
✗	✗	6.17±6.93	-0.76	Ref.
3mT 1Hz	✗	1.67±3.23	-1.56	106.28%
3mT 5Hz	✗	-0.59±0.75	-1.33	76.36%
3mT 10Hz	✗	-0.93±0.99	-1.92	153.13%
6mT 1Hz	✗	-1.22±1.21	-2.43	221.07%
9mT 1Hz	✗	-1.42±1.37	-2.79	268.76%
✗	✓	-0.89±3.89	-4.78	531.26%
3mT 1Hz	✓	-0.10±5.53	-5.63	643.58%
3mT 5Hz	✓	-0.75±3.86	-4.61	508.94%
3mT 10Hz	✓	-0.63±4.40	-5.03	564.15%
6mT 1Hz	✓	-0.61±4.37	-4.98	557.93%
9mT 1Hz	✓	-0.61±4.39	-5.00	560.25%

B. Dynamic Object Detection

Since the magnetic actuation and the electrical localization have been shown to be compatible within the presented experimental system, the next study uses magnetized robots (see Table I) and compares the EIT-based tracking to the ground truth data obtained from the images of the camera.

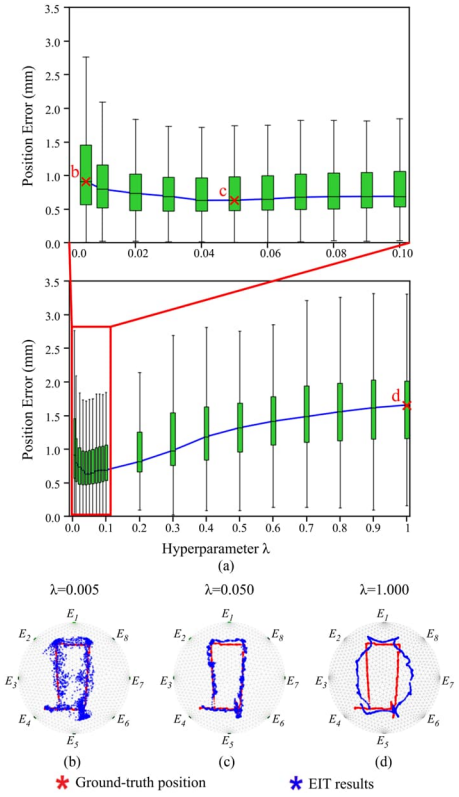


Figure 5. Influence of the hyperparameter on the accuracy of the EIT-based tracking system. (a) Evolution of the position error over the whole trajectory as a function of the reconstruction hyperparameter λ. The whiskers represent the second and 98th percentiles, while the boxes indicate the IQR and contain the median position error (blue line). The points noted b, c, and d refer to the subfigures below (b, c, d) showing the position data obtained from the camera (red) and the position estimated from the electrical signals (blue). In (b), a small hyperparameter generates noise in the reconstructed trajectory. (c) Increasing the hyperparameter improves the tracking accuracy until a local optimum. (d) Smoother reconstructed trajectory is obtained using a larger hyperparameter, however, at the cost of spatial resolution.

1) Reconstruction Parameters:

As mentioned in Section II, the spatiotemporal solver developed in [41] is used in this work to get EIT images of the robot moving within the workspace. Several reconstruction parameters defined therein have to be set, such as the NOSER exponent p, the interframe correlation γ_c, the hyperparameter λ, and the half-width d of the temporal window. In EIT, solving the inverse problem involves a tradeoff between the exact solution based on the measurements and a prior information. The exponent p is related to the spatial part of the regularization term. It tends to push the noise toward the center ($p = 1$) or the edges ($p = 0$) of the arena [41]. The parameter γ_c is involved in the temporal part of the regularization term. It reflects the correlation between successive frames. The number of frames considered in a reconstruction is determined by d. The hyperparameter λ determines the tradeoff between the measurements and the known spatiotemporal information. A high value of λ increases the weight of the spatiotemporal prior information with respect to the electrical measurements, whereas a low value of λ provides an estimated conductivity closer to the exact solution based on the measurements, however, at the cost of a higher sensibility to measurement noise.

The first parameter was set using a standard value from the literature [41] $p = 0.5$. $\gamma_c = 0.9$ is chosen because the data acquisition is known to be fast (14 ms) with respect to the dynamic of the robot. A parametric study is performed to set λ and d and highlight their impact on the accuracy of the EIT tracking. These parameters are successively varied, while reconstructing multiple times a rectangular-shaped trajectory of the robot

R2. The impact of the hyperparameter λ on the tracking accuracy is shown in Fig. 5(a), while $d = 10$. A small hyperparameter degrades the tracking accuracy by generating noisy reconstructed images. In contrast, a smooth reconstructed trajectory is obtained using a large hyperparameter, however, at the cost of the spatial resolution. These results are illustrated by the trajectories obtained in the two extreme cases considered $\lambda = 0.005$ (see Fig. 5(b)) and $\lambda = 1$ (see Fig. 5(d)), while the local optimum in terms of position accuracy is obtained with $\lambda = 0.05$ (see Fig. 5(a) and (c)).

With $\lambda = 0.05$, the influence of the temporal window d can also be studied by a parametric sweep. Although a large temporal window provides a more accurate tracking, it induces a higher temporal delay of the obtained position information (see Fig. 6). This effect is due to the fact that the reconstruction of the frame ft uses the concatenated sequence of measurements $[y_{t-d}, \cdots, y_t, \cdots, y_{t+d}]$ [41]. In this respect, considering 14 ms per frame required by the data acquisition procedure, the delay (ms) induced by this approach is $14 \times d$.

By choosing $d = 5$, the induced delay remains below the 100 ms, which is induced by a position feedback system operating at 10 Hz. This frequency is higher than the frequency reached in some recent microrobotic studies using magnetic resonance imaging (MRI) as localization modality [12].

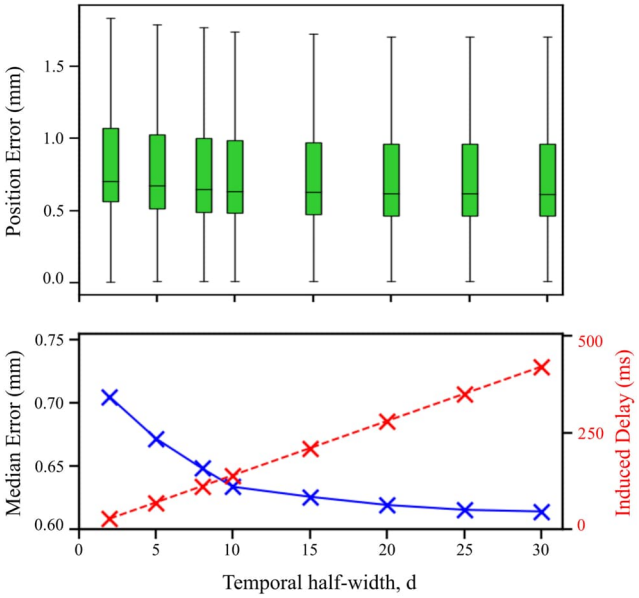


Figure 6. Influence of the reconstruction temporal width on the accuracy of the EIT-based tracking system. While the position error (top plot) decreases with the extension of the temporal window, the delay induced in the localization increases (bottom plot).

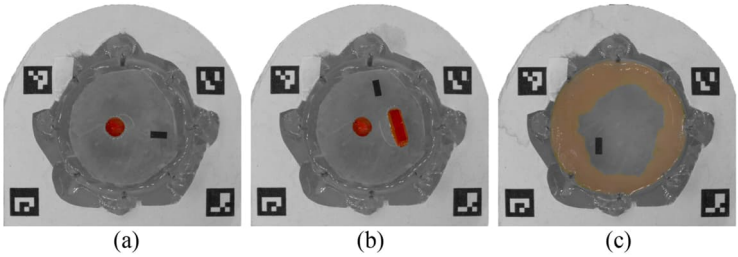


Figure 7. Dynamic test cases with R2 used to evaluate the tracking accuracy of the EIT system in nonhomogeneous environments. (a) Single circular stationary object in the test area. (b) Multiple stationary objects with different shapes in the test area. (c) Electrodes covered by the chicken flesh.

2) Tracking of Different Sized Robots:

To evaluate the capability of the proposed system to track different untethered magnetic robots, each robot (see Table I) is actuated to generate five different trajectories. These trajectories are monitored by both the camera and the EIT tracking system. The EIT reconstruction hyperparameter $\lambda = 0.05$, which is the local optimum found in the above parametric study, while $d = 5$. Other EIT reconstruction parameters are kept same as presented above. In the supplementary video, one trajectory per robot is shown together with the corresponding live EIT reconstruction.

For each experiment, the median EIT tracking error and the number of recorded positions composing the trajectory are reported in Table IV. These results show that R4 is the smallest robot that can be detected with a reasonable accuracy. Until R4, the median tracking error is lower than the length of the robot, whereas it increases significantly when the size of the robot further decreases. The results of the localization of R5 and R6 are consistent with the results of the static experiments in Section V-A (high error values). It is likely that their small size makes them undetectable by our EIT system.

Although it is bigger, R1 generates a higher tracking error than the one obtained for R2 and R3. This can be due to the bigger wavelets that are created by the motion of the robot at the surface of the liquid. These wavelets can change the local propagation of the electrical signal within the PBS, thus degrading the measurements and the reconstructed images. Besides, the setting of the hyperparameter has been done in the light of a

parametric study performed using a trajectory of R2. The parameter found to be optimal may have been different if the study was done using another robot.

Table IV Position detection accuracy of the EIT system given by the median and IQR for each magnetized robot in Table I throughout the dynamic tests.

Robot	Path Number	Path Error (mm)		Robot Error (mm)		# recorded positions
		Median	IQR	Median	IQR	
R1	P1	0.93	0.58	0.99	0.84	3887
	P2	1.21	0.74			2974
	P3	0.66	1.37			3169
	P4	1.03	0.76			3361
	P5	0.93	0.79			2224
R2	P1	0.68	0.67	0.75	0.67	2234
	P2	0.65	0.65			2058
	P3	0.68	0.50			2236
	P4	0.86	0.77			2237
	P5	0.93	0.62			2536
R3	P1	0.82	0.81	0.87	0.79	1938
	P2	0.86	0.92			1963
	P3	0.89	0.73			2135
	P4	0.94	0.87			2279
	P5	0.82	0.64			1932
R4	P1	1.01	1.05	1.11	1.14	2260
	P2	1.44	1.38			1940
	P3	1.26	1.03			2014
	P4	0.74	0.62			1565
	P5	0.92	0.81			1644
R5	P1	6.49	2.15	6.87	2.16	2118
	P2	6.57	2.24			2194
	P3	6.79	2.31			1892
	P4	6.92	1.73			2126
	P5	7.91	2.24			2213
R6	P1	14.75	9.20	7.36	9.62	2821
	P2	8.10	7.30			2553
	P3	4.07	3.84			2003
	P4	4.92	5.13			2061
	P5	8.04	13.71			2230

3) Tracking of the Magnetic Robot R2 in Nonhomogeneous Environments:

To evaluate the tracking performance of the EIT system for nonhomogeneous cases, R2 is tested for three different scenarios (see Fig. 7). In each of them, two trajectories are performed and monitored by both the camera and the EIT system to quantify the tracking performance as before. The EIT reconstruction result and ground truth data of one trajectory for each test case are presented in the supplementary video. The EIT tracking error reported in Table V shows that the proposed tracking method can localize the robot when stationary objects are in the environment (see Fig. 7(a) and (b)) with a similar performance as in the homogeneous test cases. Moreover, promising results are obtained in the third test scenario (see Fig. 7(c)), where the chicken flesh covers the electrodes. Although it is necessary to tune the EIT reconstruction parameters to achieve better localization performance as in the previous test cases, these results open the door to the application of the proposed method for medical applications.

Table V Position detection accuracy of the EIT system given by the median and IQR for R2 in Table I throughout the dynamic tests in nonhomogeneous environments.

Test Case	Path Error (mm)		Robot Error (mm)		# recorded positions
	Median	IQR	Median	IQR	
Circular Object-1	0.76	0.68	0.92	0.78	2573
Circular Object-2	1.09	0.86			3328
Multiple Object-1	0.89	0.58	0.82	0.61	2206
Multiple Object-2	0.77	0.61			2990
Tissue-1	1.27	1.43	1.24	1.26	3304
Tissue-2	1.20	1.02			2872

VI. Discussion and Outlook

This article introduces a method based on EIT to dynamically track small-scale robots. The proposed approach provides the electrical conductivity distribution within the robot workspace from a set of electrical stimulations and voltage measurements gathered from eight electrodes placed at its boundary. The position of the robot can be deduced from the reconstructed EIT image with the difference in electrical properties between the robot and the background medium. Moreover, the tracking system developed in this article has been shown not to interfere with the magnetic actuation applied to the robot by testing the system with different magnetic field strengths and actuation frequencies. Multiple robots in changing sizes were fabricated and tested, while the performance of the EIT-based tracking was verified using a camera throughout the experiments. It turned out that the smallest robot that can be tracked in the cylindrical arena (30 mm in diameter) with a median position error lower than its length is $1.5\times1.5\times1\text{cm}^3$.

Multiple factors influence the results obtained by the proposed system in terms of tracking accuracy and smallest detectable size. As highlighted in this article, the reconstruction parameters can have a significant impact. In particular, the hyperparameter has been determined based on multiple reconstructions of a trajectory of robot R2 with a varying hyperparameter. As the optimal hyperparameter can be different for the other robots, a systematic parametric optimization can help improving the tracking accuracy. Beyond the heuristic method used herein, the existing advanced methods can be implemented for tuning the hyperparameter such as the

popular L-curve method [46] and Bayesian optimization [43]. Similarly, rather than setting p and γ_c from standard values in the literature, these parameters could be determined from a parametric optimization.

The spatial resolution of the EIT images is also known to be strongly dependent on the number of measurements, which is related to the number of electrodes [38]. Even though many EIT systems in the literature involve at least 16 electrodes [38], [44], we decided to use eight electrodes for the sake of keeping the circuitry and the software as simple as possible. As our data acquisition system has eight analog input channels, providing more electrodes would require to redesign the proposed EIT system. This new version would involve either using additional multiplexers or replacing the data acquisition card by one offering more input channels, which would be more expensive. In either case, this modification would be beneficial for the spatial resolution, at the expense of the robot localization speed since the full-scan scheme would involve more successive configurations. For example, using 16 electrodes leads to 120 successive configurations, inducing an increase by a factor greater than 4 of the acquisition time compared to the 28 configurations obtained with eight electrodes.

Another factor that influences the resolution and the capacity to detect a robot is the contrast in electrical properties between the robot and the background medium. The current system uses PBS, whose conductivity is $1.6 Sm^{-1}$. Although, this value is in the same order of magnitude as the conductivity of blood [47] and other physiological fluids [48], it would

be interesting to evaluate the performance of our localization method in application-relevant medium in the future.

Besides its conductivity property, the liquid level also has a critical effect on the tracking performance. Throughout the preliminary tests, we observed that out of PBS motion of the robot would lead to lower tracking accuracy, as any part of the robot moving outside the liquid does not affect the impedance readings. Because of this, and to keep the experiments as similar to the desired biomedical robotic applications as possible, we determined the height of PBS used in the experiments twice the width of the biggest robot in Table I.

Although optical imaging remains the standard for in vitro experiments thanks to its ease of use and high resolution, the tracking accuracy obtained in this first study using EIT as a localization method in microrobotics (e.g., 0.75 and 1.24 mm median position errors for robot R2 in homogeneous test area and test area covered by chicken flesh, respectively) makes us believe that this technique could be significant and useful for the position feedback control of untethered devices for biomedical applications in the future. Moreover, it is a nonionizing imaging method (unlike X-rays), it is portable and noninterfering with magnetic actuation (unlike MRI).

We will focus on using it together with the magnetic actuation to perform some closed-loop control inside completely enclosed spaces. As the feedback control performance depends partly on the tracking accuracy, we consider combining the EIT data to the knowledge of the expected displacement induced by the actuation (state model) by using a

subsequent Kalman filter. With a data acquisition time of 14 ms, our EIT system has the potential to provide a position feedback at a frequency of 71 Hz. The frequency of the control loop will depend on coding aspects and the computing capacities available.

In the future, the application of this method to detect a small-scale robot in biomedical robotics raises new challenges such as the scale-up from a centimeter scale workspace to the human body and the optimization of the number and the arrangement of the electrodes due to the complex anatomy.

Acknowledgments

The authors would like to thank Alp C. Karacakol for the help in designing and fabricating the soft robots. Hugo Daguerre would like to gratefully thank Dr. A. Dupré for the instructive discussions.

References

[1] F. Ongaro et al., “Control of untethered soft grippers for pick-and-place tasks,” in Proc. 6th IEEE Int. Conf. Biomed. Robot. Biomechatronics, 2016, pp. 299–304.

[2] B. V. Johnson, S. Chowdhury, and D. J. Cappelleri, “Local magnetic field design and characterization for independent closed-loop control of multiple mobile microrobots,” IEEE/ASME Trans. Mechatronics, vol. 25, no. 2, pp. 526–534, Apr. 2020.

[3] J. Choi, J. Hwang, J.-Y. Kim, and H. Choi, “Recent progress in magnetically actuated microrobots for targeted delivery of therapeutic agents,” Adv. Healthcare Mater., vol. 10, no. 6, Dec. 2020, Art. no. 2001596.

[4] P. Shokrollahi et al., “Blindly controlled magnetically actuated capsule for noninvasive sampling of the gastrointestinal microbiome,” IEEE/ASME Trans. Mechatronics, vol. 26, no. 5, pp. 2616–2628, Dec. 2020.

[5] S. Palagi et al., “Structured light enables biomimetic swimming and versatile locomotion of photoresponsive soft microrobots,” Nature Mater., vol. 15, no. 6, pp. 647–653, Jun. 2016.

[6] A. Aghakhani, O. Yasa, P. Wrede, and M. Sitti, “Acoustically powered surface-slipping mobile microrobots,” in Proc. Nat. Acad. Sci., vol. 117, no. 7, pp. 3469–3477, Feb. 2020.

[7] Y. Alapan, O. Yasa, B. Yigit, I. C. Yasa, P. Erkoc, and M. Sitti, “Microrobotics and microorganisms: Biohybrid autonomous cellular robots,” Annu. Rev. Control, Robot., Auton. Syst., vol. 2, no. 1, pp. 205–230, May 2019.

[8] M. Sitti and D. S. Wiersma, “Pros and cons: Magnetic versus optical microrobots,” Adv. Mater., vol. 32, no. 20, 2020, Art. no. 1906766.

[9] H. Ceylan, I. C. Yasa, U. Kilic, W. Hu, and M. Sitti, “Translational prospects of untethered medical microrobots,” Prog. Biomed. Eng., vol. 1, no. 1, Jul. 2019, Art. no. 012002.

[10] B. Wang, Y. Zhang, and L. Zhang, “Recent progress on micro- and nano-robots: Towards in vivo tracking and localization,” Quant. Imag. Med. Surg., vol. 8, no. 5, pp. 461–479, Jun. 2018.

[11] S. Pané et al., “Imaging technologies for biomedical micro- and nanoswimmers,” Adv. Mater. Technol., vol. 4, no. 4, 2019, Art. no. 1800575.

[12] A. Aziz et al., “Medical imaging of microrobots: Toward in vivo applications,” ACS Nano, vol. 14, no. 9, pp. 10865–10893, Sep. 2020.

[13] D. Son, S. Yim, and M. Sitti, “A 5-D localization method for a magnetically manipulated untethered robot using a 2-D array of Hall-effect sensors,” IEEE/ASME Trans. Mechatronics, vol. 21, no. 2, pp. 708–716, Apr. 2016.

[14] I. S. M. Khalil et al., “Magnetic localization and control of helical robots for clearing superficial blood clots,” APL Bioeng., vol. 3, no. 2, May 2019, Art. no. 026104.

[15] O. Erin, M. Boyvat, M. E. Tiryaki, M. Phelan, and M. Sitti, “Magnetic resonance imaging system-driven medical robotics,” Adv. Intell. Syst., vol. 2, no. 2, 2020, Art. no. 1900110.

[16] S. Pane, V. Iacovacci, E. Sinibaldi, and A. Menciassi, “Real-time imaging and tracking of microrobots in tissues using ultrasound phase analysis,” Appl. Phys. Lett., vol. 118, no. 1, Jan. 2021, Art. no. 014102.

[17] C. Heunis, K. Behrendt, E. Hekman, C. Moers, J.-P. De Vries, and S. Misra, “Design and evaluation of a magnetic rotablation catheter for arterial stenosis,” IEEE/ASME Trans. Mechatronics, early access, doi: 10.1109/TMECH.2021.3092608.

[18] X. Yan et al., “Multifunctional biohybrid magnetite microrobots for imaging-guided therapy,” Sci. Robot., vol. 2, no. 12, Nov. 2017, Art. no. eaaq1155.

[19] S. Jeon et al., “Magnetically actuated microrobots as a platform for stem cell transplantation,” Sci. Robot., vol. 4, no. 30, pp. 1–11, May 2019.

[20] D. Vilela et al., “Medical imaging for the tracking of micromotors,” ACS Nano, vol. 12, no. 2, pp. 1220–1227, Feb. 2018.

[21] V. Iacovacci et al., “High-resolution SPECT imaging of stimuli-responsive soft microrobots,” Small, vol. 15, no. 34, 2019, Art. no. 1900709.

[22] R. Reale, A. D. Ninno, L. Businaro, P. Bisegna, and F. Caselli, “High-throughput electrical position detection of single flowing particles/cells with non-spherical shape,” Lab Chip, vol. 19, no. 10, pp. 1818–1827, May 2019.

[23] H. Daguerre, M. Solsona, J. Cottet, M. Gauthier, P. Renaud, and A. Bolopion, “Positional dependence of particles and cells in microfluidic electrical impedance flow cytometry: Origin, challenges and opportunities,” Lab Chip, vol. 20, no. 20, pp. 3665–3689, Oct. 2020.

[24] J. Snyder, Y. Silverman, Y. Bai, and M. A. MacIver, “Underwater object tracking using electrical impedance tomography,” in Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst., 2012, pp. 520–525.

[25] L. A. Allaud and M. H. Martin, Schlumberger: The History of a Technique. New York, NY, USA: Wiley, 1977.

[26] A. Masner, F. Blasina, and F. Simini, “Electrical impedance tomography for neonatal ventilation assessment: A narrative review,” J. Phys.: Conf. Ser., vol. 1272, Jul. 2019, Art. no. 012008.

[27] M. Rapin et al., “Wearable sensors for frequency-multiplexed EIT and multilead ECG data acquisition,” IEEE Trans. Biomed. Eng., vol. 66, no. 3, pp. 810–820, Mar. 2019.

[28] X. Shi et al., “High-precision electrical impedance tomography data acquisition system for brain imaging,” IEEE Sens. J., vol. 18, no. 14, pp. 5974–5984, Jul. 2018.

[29] M. Darnajou et al., “High speed EIT with multifrequency excitation using FPGA and response analysis using FDM,” IEEE Sens. J., vol. 20, no. 15, pp. 8698–8710, Aug. 2020.

[30] Y. Zhang, R. Xiao, and C. Harrison, “Advancing hand gesture recognition with high resolution electrical impedance tomography,” in Proc. 29th Annu. Symp. User Interface Softw. Technol., 2016, pp. 843–850.

[31] Z. Xu et al., “Development of a portable electrical impedance tomography system for biomedical applications,” IEEE Sens. J., vol. 18, no. 19, pp. 8117–8124, Oct. 2018.

[32] E. Zheng, Y. Li, Z. Zhao, Q. Wang, and H. Qiao, “An electrical-impedance-tomography-based interface for human-robot collaboration,” IEEE/ASME Trans. Mechatronics, vol. 26, no. 5, pp. 2373–2384, Oct. 2020.

[33] K. Park, H. Lee, K. J. Kuchenbecker, and J. Kim, “Adaptive optimal measurement algorithm for ERT-based large-area tactile sensors,” IEEE/ASME Trans. Mechatronics, early access, doi: 10.1109/TMECH.2021.3063414.

[34] F. Visentin, P. Fiorini, and K. Suzuki, “A deformable smart skin for continuous sensing based on electrical impedance tomography,” Sensors, vol. 16, no. 11, Nov. 2016, Art. no. 1928.

[35] J. Avery, M. Runciman, A. Darzi, and G. P. Mylonas, “Shape sensing of variable stiffness soft robots using electrical impedance tomography,” in Proc. Int. Conf. Robot. Autom., 2019, pp. 9066–9072.

[36] A. Adler and A. Boyle, “Electrical Impedance Tomography,” in Wiley Encyclopedia of Electrical and Electronics Engineering. Atlanta, GA, USA: Amer. Cancer Soc., 2019, pp. 1–16.

[37] M. Cheney, D. Isaacson, and J. C. Newell, “Electrical impedance tomography,” SIAM Rev., vol. 41, no. 1, pp. 85–101, 1999.

[38] D. Silvera-Tawil, D. Rye, M. Soleimani, and M. Velonaki, “Electrical impedance tomography for artificial sensitive robotic skin: A review,” IEEE Sens. J., vol. 15, no. 4, pp. 2001–2016, Apr. 2015.

[39] A. Adler and W. R. B. Lionheart, “Uses and abuses of EIDORS: An extensible software base for EIT,” Physiol. Meas., vol. 27, no. 5, pp. S25–S42, May 2006.

[40] Z. Zong, Y. Wang, and Z. Wei, “A review of algorithms and hardware implementations in electrical impedance tomography,” Prog. Electromagn. Res., vol. 169, pp. 59–71, 2020.

[41] A. Adler, T. Dai, and W. R. B. Lionheart, “Temporal image reconstruction in electrical impedance tomography,” Physiol. Meas., vol. 28, no. 7, pp. S1–S11, Jun. 2007.

[42] W. Hu, G. Z. Lum, M. Mastrangeli, and M. Sitti, “Small-scale soft-bodied robot with multimodal locomotion,” Nature, vol. 554, no. 7690, pp. 81–85, Feb. 2018.

[43] S. O. Demir, U. Culha, A. C. Karacakol, A. Pena-Francesch, S. Trimpe, and M. Sitti, “Task space adaptation via the learning of gait controllers of magnetic soft millirobots,” Int. J. Robot. Res., vol. 40, pp. 1331–1351, 2021.

[44] A. Dupré, G. Ricciardi, and S. Bourennane, “Novel approach for analysis and design of high-speed electrical impedance tomographic system for void fraction measurements in fast two-phase flows,” IEEE Sens. J., vol. 17, no. 14, pp. 4472–4482, Jul. 2017.

[45] C. Dang, C. Bellis, M. Darnajou, G. Ricciardi, S. Mylvaganam, and S. Bourennane, “Practical comparisons of EIT excitation protocols with applications in high contrast imaging,” Meas. Sci. Technol., vol. 38, no. 8, 2021, Art. no. 085110.

[46] A. Cultrera and L. Callegaro, “A simple algorithm to find the L-curve corner in the regularisation of ill-posed inverse problems,” IOP SciNotes, vol. 1, no. 2, Aug. 2020, Art. no. 025004.

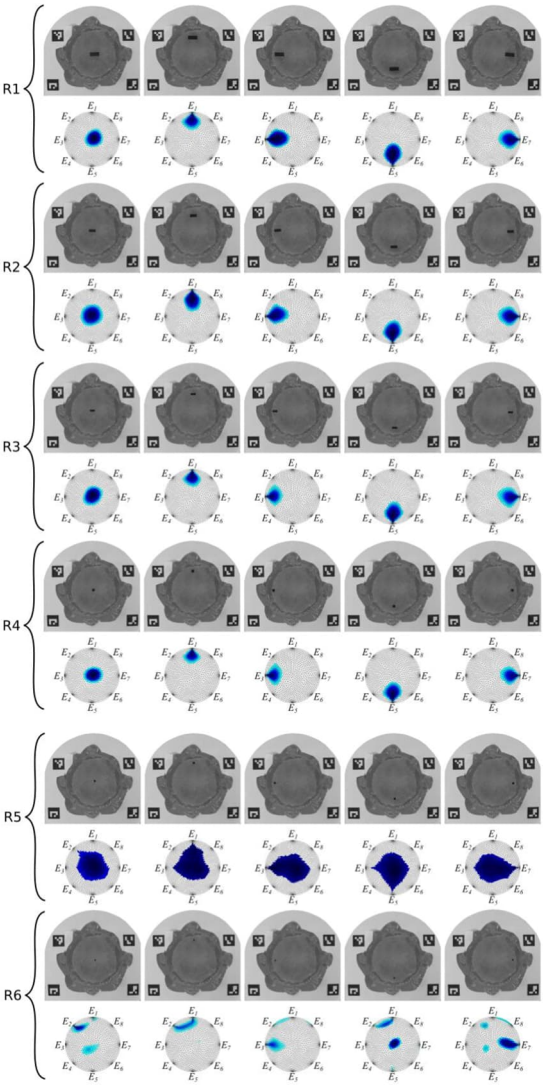
[47] F. G. Hirsch et al., “The electrical conductivity of blood: I. Relationship to erythrocyte concentration,” Blood, vol. 5, no. 11, pp. 1017–1035, Nov. 1950.

[48] Y. Lu et al., “AC electrokinetics of physiological fluids for biomedical applications,” J. Lab. Autom., vol. 20, no. 6, pp. 611–620, Dec. 2015.

Supplementary Material A. Influence of the magnetic field (B) on the position detection accuracy of the EIT system for non-magnetized robots in Table I throughout the static tests in Section V-A

	No B-Field		3mT 1Hz		3mT 5Hz		3mT 10Hz		6mT 1Hz		9mT 1Hz	
	Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR
R1	0.61	0.08	0.55	0.26	0.62	0.37	1.76	0.30	0.62	0.33	0.63	0.32
R2	0.38	0.21	0.45	0.40	0.64	0.39	1.62	0.31	0.75	0.37	0.74	0.43
R3	0.56	0.33	0.54	0.51	0.75	0.31	1.48	0.50	0.73	0.31	0.79	0.27
R4	0.70	0.61	0.58	0.74	0.91	0.65	0.95	0.77	0.95	0.80	0.85	0.75
R5	8.66	0.42	6.63	8.40	1.10	1.04	2.20	0.96	1.19	0.95	1.36	0.89
R6	3.80	13.03	7.00	4.96	11.77	5.76	10.43	7.48	13.61	5.38	15.33	8.98

Supplementary Material B. Reconstructed EIT images of non-magnetized robots in Table I with |B|=0mT throughout the static tests in Section V-A



Supplementary Material C. Reconstructed EIT images of non-magnetized robots in Table I with external rotating magnetic field defined by |B|=3mT and *f*=1Hz throughout the static tests in Section V-A