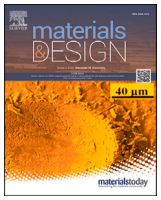
Materials & Design 238 (2024) 112629

Contents lists available at ScienceDirect

Materials & Design

journal homepage: www.elsevier.com/locate/matdes



Vision-based tactile intelligence with soft robotic metamaterial ☆

Tianyu Wu [a,b,1], Yujian Dong [a,b,1], Xiaobo Liu [b,1], Xudong Han [b], Yang Xiao [b], Jinqi Wei [b], Fang Wan [a,c,*], Chaoyang Song [b,*]

[a] Shenzhen Key Laboratory of Intelligent Robotics and Flexible Manufacturing Systems, Southern University of Science and Technology, China
[b] Department of Mechanical and Energy Engineering, Southern University of Science and Technology, China
[c] School of Design, Southern University of Science and Technology, China

ARTICLE INFO

Dataset link: https://github.com/bionicdl-sustech/SoftRoboticTongs

Keywords:
Soft robotic metamaterials
Vision-based tactile sensing
Machine learning

ABSTRACT

Robotic metamaterials represent an innovative approach to creating synthetic structures that combine desired material characteristics with embodied intelligence, blurring the boundaries between materials and machinery. Inspired by the functional qualities of biological skin, integrating tactile intelligence into these materials has gained significant interest for research and practical applications. This study introduces a Soft Robotic Metamaterial (SRM) design featuring omnidirectional adaptability and superior tactile sensing, combining vision-based motion tracking and machine learning. The study compares two sensory integration methods to a state-of-the-art motion tracking system and force/torque sensor baseline: an internal-vision design with high frame rates and an external-vision design offering cost-effectiveness. The results demonstrate the internal-vision SRM design achieving an impressive tactile accuracy of 98.96%, enabling soft and adaptive tactile interactions, especially beneficial for dexterous robotic grasping. The external-vision design offers similar performance at a reduced cost and can be adapted for portability, enhancing material science education and robotic learning. This research significantly advances tactile sensing using vision-based motion tracking in soft robotic metamaterials, and the open-source availability on GitHub fosters collaboration and further exploration of this innovative technology (https://github.com/bionicdl-sustech/SoftRoboticTongs).

1. Introduction

The increasing demand for machine intelligence underscores the need to engage in material design when developing contemporary robotic solutions. Robotic metamaterials offer an innovative approach to crafting three-dimensional (3D) structures with responsive properties, utilizing advancements in additive manufacturing techniques. Recent research [5] has demonstrated the creation of robotic metamaterials capable of programmable motions, self-sensing, and feedback control by incorporating piezoelectric, conductive, and structural elements into a 3D lattice. Another study [28] has developed a material that can sense deformations, transitioning between elasticity and plasticity by integrating sensors. Nevertheless, a design challenge remains in striking a balance between the mechanical richness of metamaterials and the use of robotic intelligence for practical applications [12,16].

A research gap exists in integrating material design and robot learning for a combined effort in fundamental research and engineering applications [18]. Material scientists and applied roboticists are actively exploring the incorporation of tactile intelligence into robotic metamaterials. The skin, the body's largest organ, offers protection, regulation, and sensation. Tactile interaction between the skin and the surrounding environment is vital for dexterous manipulation, and recent research has witnessed the adoption of touch-based tactile intelligence in various robotic applications [8,30,3]. Tactile sensing is commonly encoded as force and torque, with options ranging from basic binary contact detection to advanced 6D force and torque sensing [22,10]. Distributed

☆ This work was partly supported by the Ministry of Science and Technology of China [2022YFB4701200], the National Natural Science Foundation of China [62206119, 52335003], and the Science, Technology and Innovation Commission of Shenzhen Municipality [JCYJ20220818100417038, ZDSYS20220527171403009, and SGDX20220530110804030].

* Corresponding author.
E-mail addresses: wanf@sustech.edu.cn (F. Wan), songcy@ieee.org (C. Song).

[1] Equal contributions as the co-first authors.

https://doi.org/10.1016/j.matdes.2024.112629
Received 30 October 2023; Received in revised form 9 December 2023; Accepted 1 January 2024
Available online 11 January 2024

tactile sensing can be achieved using sensing unit arrays to detect normal pressure over the contact surface [2]. Optical, vision, and magnetic modalities have also gained traction for advanced tactile detection [31,35,33,6].

The growing field of soft robotics allows the development of learning-based solutions that mimic the skin's properties. Interfaces with a soft design, especially those with complex physical forms, can provide rich deformations for vision-based sensors [19]. Recent research leverages magnetic sensors and machine learning to detect tactile interactions with high precision [30]. Another study has demonstrated the integration of visual and tactile modalities through total internal reflection to observe external environments and detect tactile features [34]. Mechanoreceptive designs with soft structures have also shown promise in capturing tactile interactions, such as the Volumetrically Enhanced Soft Actuator (VESA) design, which uses optical sensors to detect external forces on soft pneumatic actuators [27]. 3D metamaterials with a soft design can integrate optical sensors within the soft structure or on the contact surface to provide real-time data for learning-based applications [32,26].

The field of robot design is evolving to integrate sensing, planning, and actuation into a unified unit, addressing the shift towards learning-based approaches where force-related modalities take precedence over accuracy and speed [17]. Tactile perception offers valuable information about objects and environmental properties unaffected by occlusion or lighting, reducing uncertainty in localized perception [21]. Recent designs, such as the insight finger, offer high-performing tactile sensing capabilities [25]. For enhanced fabric defect detection accuracy, a study utilized tactile images captured by a vision-based sensor, exceeding the capabilities of conventional methods [7]. Force-related modalities are increasingly incorporated into learning-based robot control [9]. While various tactile sensor designs have been proposed, making them reproducible and cost-effective remains challenging for efficiently collecting structured touch interaction data for learning [13].

Tactile sensing is also crucial in robotic manipulation, and many studies have adopted vision-based methods to achieve tactile perception [11,1]. A novel study presented a universal jamming gripper integrating tactile perception, facilitated by vision-based methods, for effective underwater manipulation [15]. In studies [14,20], fiducial markers were utilized for visual identification and to detect deformations captured by the camera. There have been various attempts to equip soft grippers with accurate proprioception and tactile sensing. Recent research transformed conventional fin-ray fingers by integrating rigid nodes into their compliant structure, allowing force sensing with vision [29]. GelFlex is a novel exoskeleton-covered soft finger with embedded cameras and deep learning methods that enable high-resolution proprioceptive and rich tactile sensing [24].

This study introduces an omni-adaptive Soft Robotic Metamaterial (SRM) with vision-based pose tracking to enable machine learning for tactile intelligence in a compact, cost-effective form. The SRM exhibits remarkable adaptation to geometric stimuli in all directions, making it suitable for adaptive grasping with multi-fingered grippers. We evaluate its geometric adaptation to external forces using state-of-the-art motion tracking and force-sensing technologies, paving the way for data-driven tactile sensing with machine learning. We present two vision-based tactile sensing methods to meet different cost constraints: one with a compact in-finger vision capturing the SRM's spatial adaptations at a high framerate and the other with a portable design on tongs for capturing both force and motion data at a lower cost. We demonstrate the tactile sensing capabilities of both methods, showcasing the SRM's potential in various tasks involving human-robot interactions in research and education. This work contributes to advanced engineering informatics by enhancing the capacity for tactile sensing in soft robotic metamaterials through vision-based motion tracking.

2. Materials and methods

2.1. Vision-based robotic metamaterial design

The proposed design consists of two primary components shown in Fig. 1, including a soft pyramid network forming the structure of the metamaterial and a miniature motion tracking system that visually records the spatial changes of the metamaterial. Our design begins with assembling pyramid polyhedrons, stacking them into multiple layers to create a single unit. We then transform this assembly into a network by removing all external surfaces. Subsequently, we establish a solid metamaterial body by sweeping all edges with a predefined cross-sectional geometry. Flexure joints are added to all mid-layer edges to enhance their adaptability during physical interactions.

The specific design under analysis features two vertices at the tip and a square-shaped bottom. This design includes a primary interaction surface in the form of a trapezoid optimized for improved grasping and a secondary interaction surface in a triangle, enabling omnidirectional adaptation. The resulting metamaterial can be fabricated using soft materials like TPU or rubber through 3D printing or molding. As depicted in Fig. 1A(i), this metamaterial exhibits remarkable omni-adaptivity when in contact with external objects, resulting in significant bending in the interaction surfaces, twisting about the central z-axis and complex volumetric deformations.

The soft nature of the material allows force/torque information generated from touch-based events to be inferred from compliant deformations. This study presents two approaches for integrating a miniature motion tracking system to capture the compliant deformations and sensitize the metamaterial. When viewed from beneath, the hollow interior provides an unobstructed view of the metamaterial's geometric adaptation, making it ideal for housing a miniature motion-tracking system. In Fig. 1A(ii), a high-performing method involves affixing a miniature camera (S-YUE WX605 from Weixinshijie) with a high framerate of up to 330 Hz to the underside of the metamaterial. In contrast, an alternative approach is shown in Fig. 1A(iii), where two fiducial marker plates are attached to the back of the metamaterial's primary interaction surface. An external miniature camera (H200S2.9mm from JIERUI-WEITONG), mounted on a foldable stand, tracks the markers, encoding the metamaterial's adaptive deformations into a six-dimensional pose vector. Both methods transform the metamaterial into a robotized form by visually encoding its spatial adaptation into the pose movements of the fiducial markers in six dimensions.

2.2. Tracking omni-adaptive motion with 6D force and torque sensing

We comprehensively evaluated our proposed methods utilizing state-of-the-art (SOTA) motion tracking and force and torque sensing technologies. Three distinct sets of time-series data were collected for thorough analysis, encompassing both motion and force measurements, which we refer to as the *baseline*, *internal-vision*, and *external-vision* datasets, with the description summarized in Table 1. Here, the baseline dataset is collected using a SOTA motion tracking system with the highest measurement accuracy for benchmarking purposes. In contrast, the internal and external vision datasets are compared against performance evaluations towards differentiated engineering benefits in design and applications (see the supplementary materials in Appendix A for further details).

To compile the *baseline* data, we established our experimental setup by fixing the robotic metamaterial on a table and mounting it on a 6D Force/Torque (F/T) sensor (Nano25 from ATI), as illustrated in Fig. 1B(i). Five markers were affixed to the metamaterial to monitor its adaptive deformation, utilizing a motion-tracking system with 12 Raptor-12 digital cameras from Motion Analysis. We introduced random pushes to the metamaterial using a 3D-printed pushrod, varying the pushes' angles, heights, and depths. Concurrently, we recorded the positions of the five markers ($[x_i, y_i, z_i], i = 1, ..., 5$) and the forces and

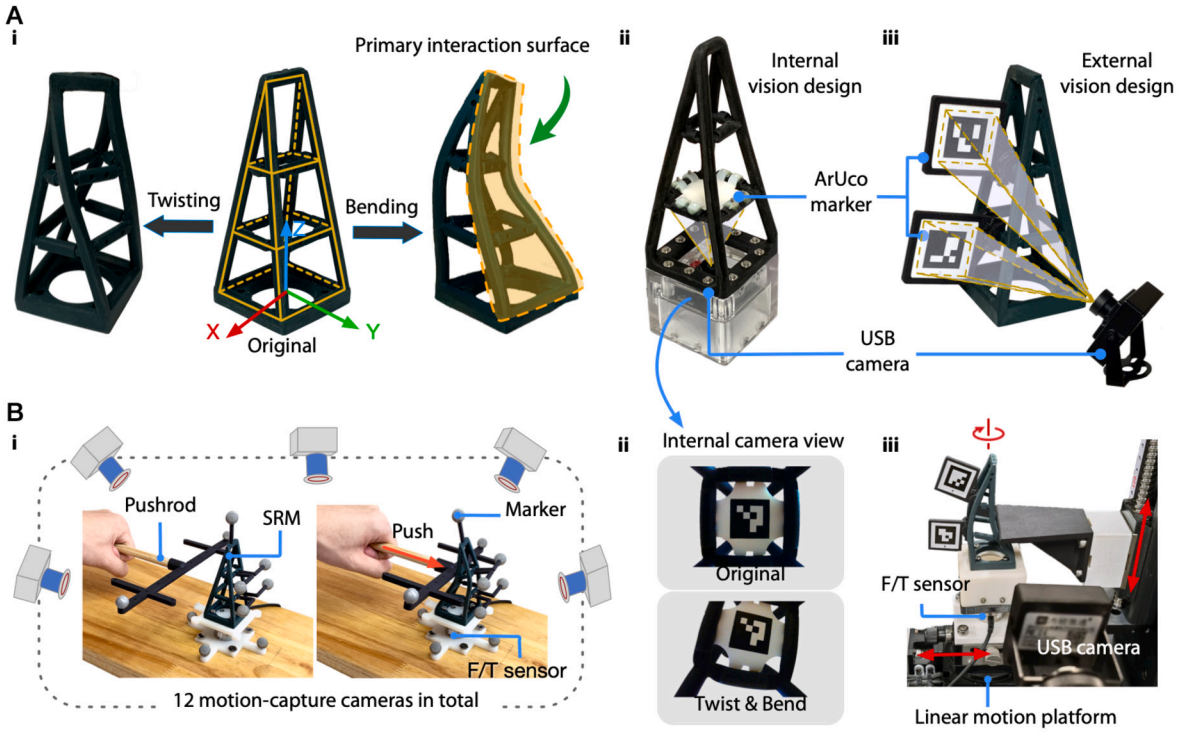


Fig. 1. Omni-adaptive soft robotic metamaterial (SRM) with vision-based machine learning for tactile intelligence. (A) The design of SRM with a pyramid network structure i) before sensorization and after with ii) internal and iii) external visions. (B) Experimental of SRM's force-displacement relationships: i) baseline setup (obtaining SRM's deformation using a motion tracking system). ii) internal vision setup (with an F/T sensor under the camera). iii) external vision setup (settled on a linear motion platform).

Table 1
Summary of the three datasets collected.

Experiment	Data type	Dimension	Training set size	Validation set size
Baseline	The positions of the five motion tracking markers: $[x_i, y_i, z_i], (i = 1, ..., 5)$	15	20,500	1,500
Internal Vision	The pose of the ArUco marker: $[X, Y, Z, roll, pitch, yaw]$	6	21,000	1,500
External Vision	The relative pose of two ArUco markers: $[X, Y, Z, roll, pitch, yaw]$	6	21,617	1,601

torques acting on the metamaterial at a rate of 120 Hz for 10 minutes. Given the metamaterial's symmetrical design, we exclusively recorded interaction data from one of its primary interaction surfaces.

Temporal synchronization between the data from the motion tracking system and the F/T sensor was a crucial consideration, as they were recorded on separate computers. To address this, we attached two markers to the pushrod and initiated multiple impacts on the experimental platform using the pushrod. This approach allowed us to identify the peak values of the pushrod's velocity and the applied force. The simultaneous occurrence of these peak values ensured their precise temporal alignment.

The *internal-vision* dataset was acquired by applying diverse forces to the metamaterial while simultaneously collecting data from the F/T sensor. Additionally, the pose data of fiducial markers ($[X, Y, Z, roll, pitch, yaw]$) was captured at a high frame rate by a miniature camera running at 330 Hz for 4 minutes, as displayed in Fig. 1B(ii).

For the *external-vision* setup, depicted in Fig. 1B(iii), the metamaterial was fixed on a linear motion platform. Two fiducial markers were attached to the back of the primary interaction surface. An external miniature camera affixed to a foldable stand was employed to record the pose data of these fiducial markers. The linear motion platform was then utilized to induce horizontal and vertical movements of the metamaterial, leading to deformations in various directions. Throughout this process, we recorded both the relative pose of the two fiducial markers and the forces and torques acting on the metamaterial. All recorded data was synchronized during post-processing for subsequent analysis and machine learning applications.

2.3. Tactile robot learning

We analyzed the robotic metamaterial's performance using the three recorded datasets and developed learning-based models for predicting tactile interactions. Fig. 2A illustrates our use of a multi-layer perceptron (MLP) model with three hidden layers, comprising 1,000, 100, and 50 neurons, to evaluate the tactile prediction capabilities using different vision-based sensorization methods for the SRM.

For the experiment configuration outlined in Fig. 1B, we collected data for the SRM's force-displacement relationship in its initial *baseline* state and after sensorization with internal and external vision. The input features for the network encompass positional and kinetic attributes of the markers, encapsulating the SRM's geometric adaptations across approximately 20,000 samples for each model. The ground truth labels consist of 6D force and torque measurements obtained through the force/torque sensors. The network's output provides predicted tactile data through 6D forces and torques. We employed this network to train three models tailored to the respective dataset from these separate experiments. We randomly selected around 1,500 samples in each case to create a test set for evaluating the model's predictive performance.

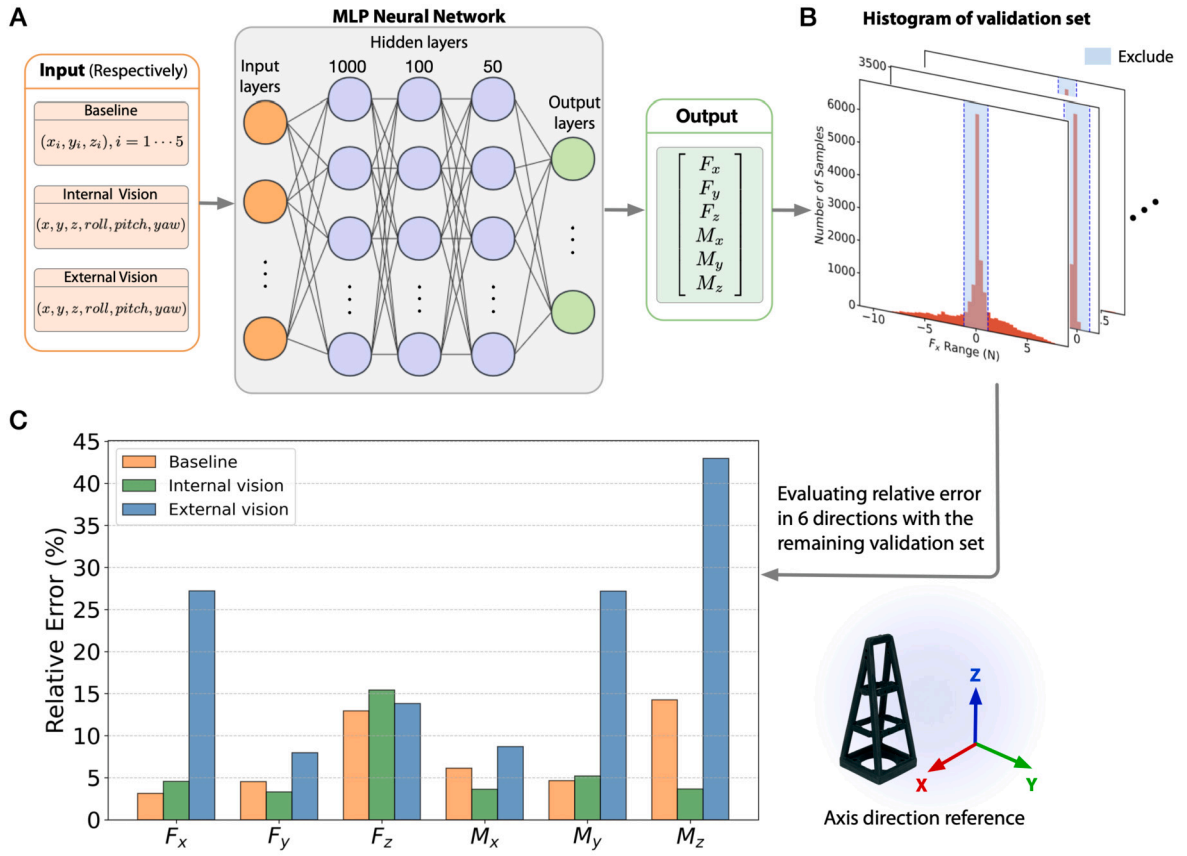


Fig. 2. Neural network for learning tactile intelligence and experimental results. (A) Multilayer perceptron with three hidden layers for learning tactile intelligence with positional and kinetic terms as inputs for estimating the tactile force and torques as the outputs. (B) Data distribution of the validation set and schematic of excluding data around zero values. (C) Relative errors of the robotic metamaterial with baseline, internal, and external vision.

3. Experimental results

3.1. Vision-based tactile intelligence with SRMs

The experiments involving the *baseline*, *internal-vision*, and *external-vision* setups yield data with varying ranges. The *internal-vision* method allows the SRM to enable unrestricted deformations in various directions, whereas the designs of the *baseline* and *external-vision* setups restrict the range of SRM's movements due to limitations imposed by the markers. To ensure fair and consistent error assessment across these experiments despite the differences in data ranges, we employ Mean Absolute Percentage Error (MAPE) calculations to facilitate a standardized approach for error evaluation. However, due to the experimental setup, the input data for the neural network exhibit a skewed distribution, with a significant portion of the data clustering around zero, as shown in Fig. 2B (details of the data distribution and the excluded range can be found in Fig. 7 in the Appendix A.). Calculating MAPE directly from this data would lead to inaccurate results due to the abundance of values near zero. We employ a data reduction technique for the neural network's output results to address this issue. This method involves excluding data points that fall within a range of ± 0.1 times the overall range of the dataset centered around zero. We then calculate the relative error based on the remaining dataset. The relative error is calculated as:

$$RelativeError = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|, \quad (1)$$

where y_i represents the true values of the remaining test set, and $\hat{y}_i$ represents the predicted values.

The results presented in Fig. 2C reveal that all three vision-based sensorization methods demonstrate their best performance in terms of F_y and M_x, with a relative error falling within the range of [3.31%, 8.69%]. These two force/torque components specifically result from forces applied to the primary interaction surface, making their accurate estimation particularly significant in these specific directions. As indicated in Fig. 2C, the *internal vision* model closely mirrors the *baseline* results in predicting forces and torques across various directions, excelling in predicting M_z. However, the *external vision* model falls short in predicting the same torque component.

To provide a practical assessment of the errors in estimating the 6D force and torque by the metamaterial, we also computed the Mean Absolute Error (MAE) for the results of the three sets of experiments, as shown in Table 2. The MAE was calculated within three neural network models we trained: a 3-layer MLP, a 4-layer MLP, and ResNet. The comparison showed that the more complex networks closely approximated the MAE values of the simpler 3-layer MLP, with variations in different directions. This suggests that using a simple 3-layer MLP suffices and offers commendable results for this study while considering computational efficiency. A thorough examination of the results reveals that F_y and M_x exhibit commendable MAE performance, especially compared to their respective data ranges.

3.2. Dexterous teleoperation with internal vision

We demonstrate the versatility of the robotic metamaterial with internal vision in facilitating human-robot interaction via teleoperation, enabling precise dexterous manipulation. An overview of the system is presented in Fig. 3A, where one robotic metamaterial is configured as programmable arcade buttons while another serves as a joystick. These components are connected to a laptop with a browser-based user inter-

Table 2
The MAE results for baseline, internal, and external vision datasets using different networks.

Network	F/T	Baseline		Internal Vision		External Vision	
		MAE	Data Range	MAE	Data Range	MAE	Data Range
MLP with 3 hidden layers (Our method)	F_x (N)	0.132	[-6.211, 5.478]	0.338	[-10.728, 10.759]	0.197	[-1.420, 0.687]
	F_y (N)	0.270	[-18.911, 0.015]	0.293	[-14.040, 13.826]	0.548	[-14.832, -0.717]
	F_z (N)	0.248	[-6.027, 0.839]	0.209	[-4.631, 0.031]	0.342	[-5.570, 0.351]
	M_x (Nm)	0.019	[0.001, 1.002]	0.039	[-1.633, 1.642]	0.045	[0.053, 1.190]
	M_y (Nm)	0.010	[-0.287, 0.292]	0.043	[-1.211, 1.211]	0.020	[-0.124, 0.097]
	M_z (Nm)	0.009	[-0.086, 0.096]	0.009	[-0.280, 0.233]	0.016	[-0.054, 0.058]
MLP with 4 hidden layers	F_x (N)	0.104		0.314		0.170	
	F_y (N)	0.225		0.296		0.447	
	F_z (N)	0.178	Same as above	0.195	Same as above	0.300	Same as above
	M_x (Nm)	0.020		0.040		0.039	
	M_y (Nm)	0.015		0.044		0.021	
	M_z (Nm)	0.01		0.015		0.019	
ResNet18	F_x (N)	0.172		0.398		0.202	
	F_y (N)	0.254		0.545		0.609	
	F_z (N)	0.175	Same as above	0.246	Same as above	0.359	Same as above
	M_x (Nm)	0.023		0.074		0.054	
	M_y (Nm)	0.017		0.062		0.025	
	M_z (Nm)	0.011		0.032		0.017	

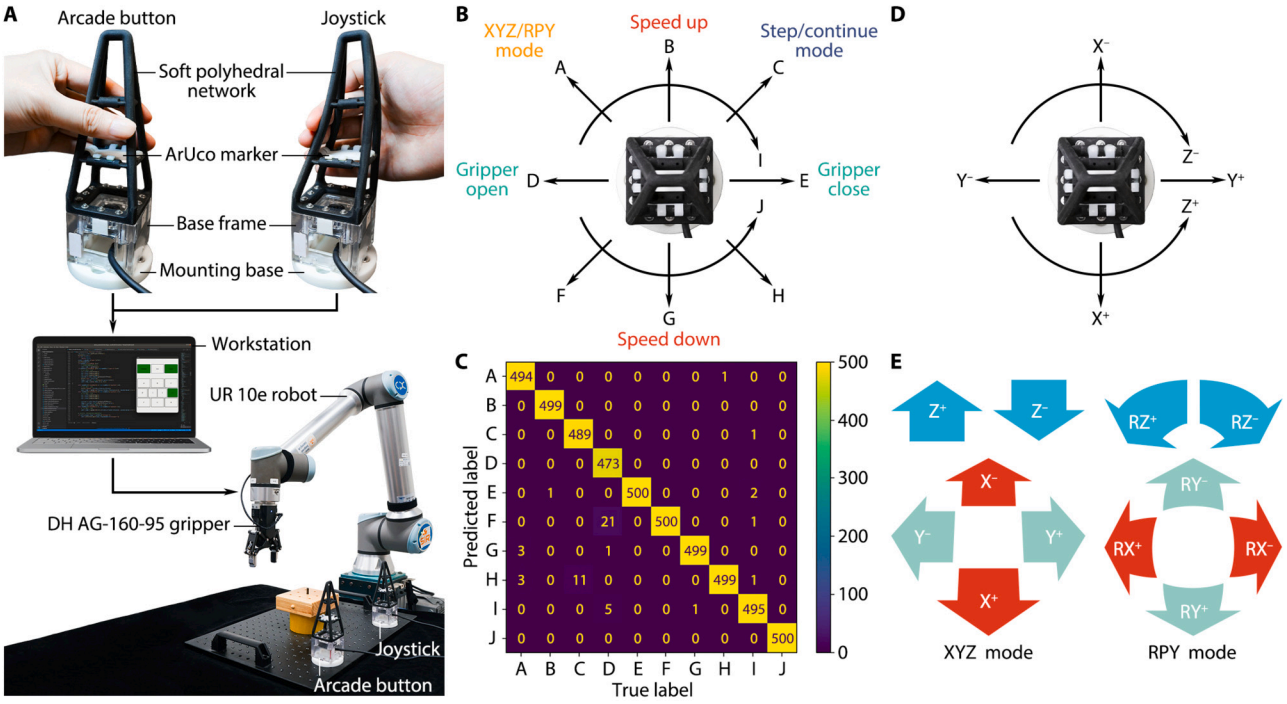


Fig. 3. Teleoperation using the robotic metamaterial with an internal vision for dexterous manipulation. (A) Setup with two robotic metamaterials as controllers with different functions, teleoperating a robotic arm for a peg-in-hole task using YCB objects. (B) Functional mapping as arcade buttons by compressing or twisting them in different directions and (C) Confusion matrix of a neural network trained for classifying ten programmable functions by physically interacting with the soft network. (D) Functional mapping of the soft network as a joystick by compressing or twisting it in different directions and (E) Two modes of interaction for translation in XYZ directions and rotation in RPY angles, corresponding to the typical 6D motion in the Cartesian space, respectively.

face, which communicates with a robot workstation for a peg-in-hole task.

We assign high-level robot commands to ten states of the soft network's whole-body deformations, transforming them into programmable arcade buttons, as depicted in Fig. 3B. To enhance classification accuracy, we trained an MLP model using 20,000 training and 5,000 testing samples, achieving an impressive accuracy of 98.96%, and the confusion matrix for classification accuracy is displayed in Fig. 3C.

Concurrently, we programmed the other robotic metamaterial by mapping the marker's displacement in the x-y plane and its rotation about the z-axis (D_x, D_y, D_z) to the robot end-effector's Cartesian movements in translational (v_x, v_y, v_z) or rotational mode (w_x, w_y, w_z), depending on the user's commands, as shown in Fig. 3E.

We conducted a teleoperation experiment for a peg-in-hole task employing YCB objects [4]. The interaction sequence, as depicted in Fig. 4, involves signals from the arcade button and joystick, both programmed using the two robotic metamaterials equipped with internal vision. Teleoperation unfolds as a series of actions, including pressing or twisting the arcade button for mode switching and manipulating the joystick through pushing and twisting for motion control until the task is completed. This process facilitates discrete and continuous motion mapping from the human operator to robot actions, encompassing movements, mode switches, and gripper control. This enables dexterous

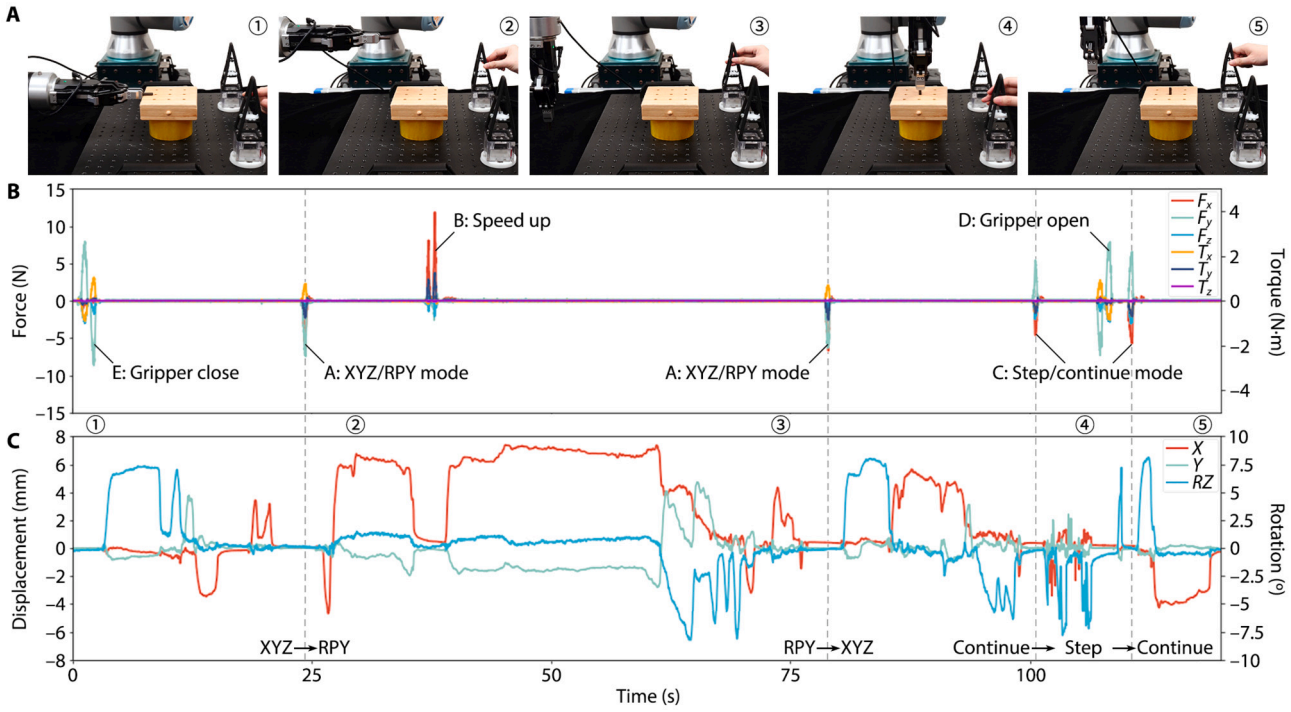


Fig. 4. The teleoperation task by using the SRMs and recorded results. (A) The motion sequence of the teleoperation task while interacting with the two soft networks. (B) Recorded results for the soft network as arcade buttons. (C) Recorded results for the soft network as a joystick.

object manipulation in 6D interactions (see the supplementary movie in Appendix B for further details.).

3.3. Portable interaction with external vision

In this section, we introduce an alternative approach to achieve tactile intelligence using the SRM with external vision, which can gather force and motion data to represent various actions concurrently. Fig. 5A showcases the system design of a portable tong, which incorporates SRMs and an external vision system. This entire setup is conveniently foldable into a compact package and can be swiftly assembled on a tabletop for use with a laptop through a USB connection.

As illustrated in Fig. 5B, the metamaterial's omni-directional adaptability empowers it to grasp objects of various shapes securely. It can function as robotic fingers, enhancing a robotic arm's flexibility and capability to grip items with diverse geometries and levels of fragility securely. In essence, this teleoperated control scheme vastly broadens the horizons for the practical applications of robotics in a wide range of real-world scenarios.

To facilitate this process, we have developed a web-based user interface that efficiently collects tactile data in batches for model training and experimentation, as depicted in Fig. 5C. The interface comprises four data screens, arranged on the right, providing a live camera view, a visualization interface, a data chart, and a control panel. Users can label each marker detected by the camera using the control panel. This labeling procedure associates the marked point with the corresponding object, which is then displayed in a 3D representation on the visualization screen. Specifically, for the six markers affixed to the portable tong, the naming process follows a similar approach, enabling the visualization interface to monitor the three-dimensional movements of the portable tong and display the applied gripping force on the object within the data chart (see the supplementary movie in Appendix B for further details).

Fig. 5D overviews the interaction process between the portable SRM and a human operator. This interaction relies on detecting marker displacements to calculate the SRM's spatial movement and adaptive deformation, which can be further transmitted to a robotic manipulator for teleoperation. The overall action data to be sent to the robotic system involves motion data generated by the teleoperator's hand while holding the tongs and the tactile data represented in 6D F/T generated when the SRMs interact with the objects. At a low cost, this system provides a comprehensive set of action data collected from both motion and force simultaneously, which is a research challenge for learning from demonstrations in robotics [23].

3.4. Teaching tactile intelligence with SRMs

We have taken the development of the portable SRM a step further, transforming it into a cost-effective, shareable, and reproducible educational resource for teaching tactile intelligence. A collection of online tutorials is hosted online for open access.[2] This transformation is visually represented in Fig. 6, where students can actively engage with this tool, gaining practical experience in tactile data collection and interactive learning leveraging the SRM.

Students can effortlessly employ their laptops to access the online interface, establishing a seamless connection through a USB link to the camera embedded in this educational tool. Incorporating this tool into the classroom setting enables students to gain hands-on experience with various aspects of tactile data collection, including the training of models and the assessment of performance. The portable and user-friendly design ensures that the SRM is a convenient tool for acquiring knowledge and skills about tactile sensing and machine learning algorithms, making it an invaluable resource for educational purposes (see the supplementary movie in Appendix B for further details.).

4. Discussion

4.1. Tactile intelligence enabled by SRM design

Our research underscores the potential of the designed Soft Robotic Metamaterials (SRMs) to achieve vision-based, omni-adaptive tactile

[2] https://me336.ancorasir.com/?page_id=457.

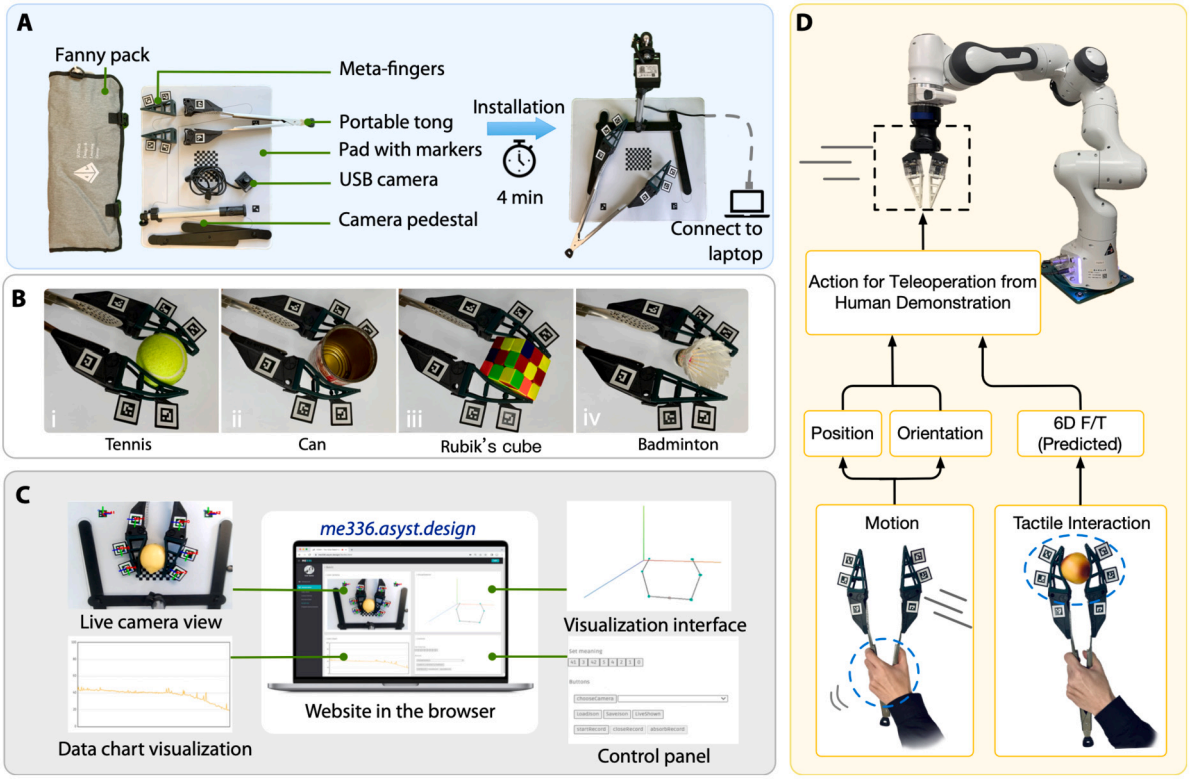


Fig. 5. Portable interaction using the soft robotic metamaterial with an external vision at a low cost. (A) The system design of a portable tong with SRM features an external vision for tactile sensing. (B) Omni-directional adaptation enabled by the SRM for grasping items of various shapes. (C) The web-based user interface for tactile data collection. (D) Integration of the SRM with an external vision for human-robot interaction.

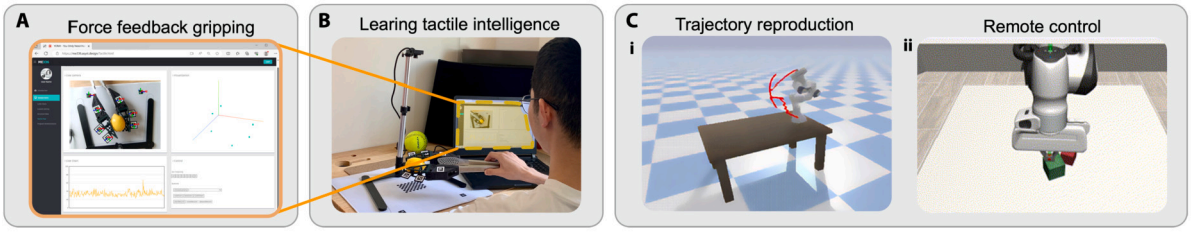


Fig. 6. Pilot program implementation for teaching tactile intelligence with the portable tong. (A) Enlaged view of the browser-based user interface for data collection. (B) Students engaged in tactile intelligence learning with the portable tong. (C) Manipulating a robot with portable tongs in a virtual physics simulation scene: i) Learning trajectory reproduction in PyBullet. ii) Remote control operation of a robot for cube manipulation within the Robosuite environment.

intelligence. The results of this study emphasize this approach's effectiveness in structuring SRMs with unique material properties, enabling adaptability to various tasks, including robotic manipulation, force and torque prediction, human-robot interaction, and more.

By examining our experimental results, we have effectively harnessed the vision-based approach, coupled with machine learning algorithms, to enable the sensorization of the metamaterial through two distinct methods: *internal vision* and *external vision*. Utilizing these methods, we have successfully estimated the magnitude of forces and torques applied to this SRM. To establish a benchmark for evaluation, we initially acquired the force-deformation relationship of our designed SRM using a state-of-the-art motion capture system and a 6D Force/Torque sensor. The *baseline* model, which we trained based on this data, serves as the ground truth for comparison with the other two methods to assess their effectiveness.

Notably, we have observed varying levels of accuracy in predicting forces and torques in different directions. In scenarios where forces were applied to the primary interaction surface, we achieved high predictive accuracy, aligning with the typical use case of the SRM. However, the estimation performance of the *external vision* method for parameters like F_x, F_z, M_y, and M_z exhibited relatively poorer results. We attribute this outcome to two key factors:

- The ArUco markers are placed at a relatively greater distance from the external camera, resulting in a more limited field of view for the markers. This limitation negatively impacts the precision of ArUco marker pose estimation, especially concerning the markers' z-direction.
- The unique structural characteristics of this network-like metamaterial and the placement of markers make it challenging to detect substantial deformations in these specific directions, which are not well-reflected in the markers' pose changes.

These same reasons can also elucidate why the performance of the *internal vision* approach is suboptimal in the F_z direction. However, the *internal vision* design approach demonstrates robust and accurate performance across the remaining five directions. This robustness is due to the proximity of the ArUco marker to the camera and its alignment with the metamaterial's $x-y$ plane. As a result, the changes in the marker's pose are more pronounced, leading to more accurate estima-

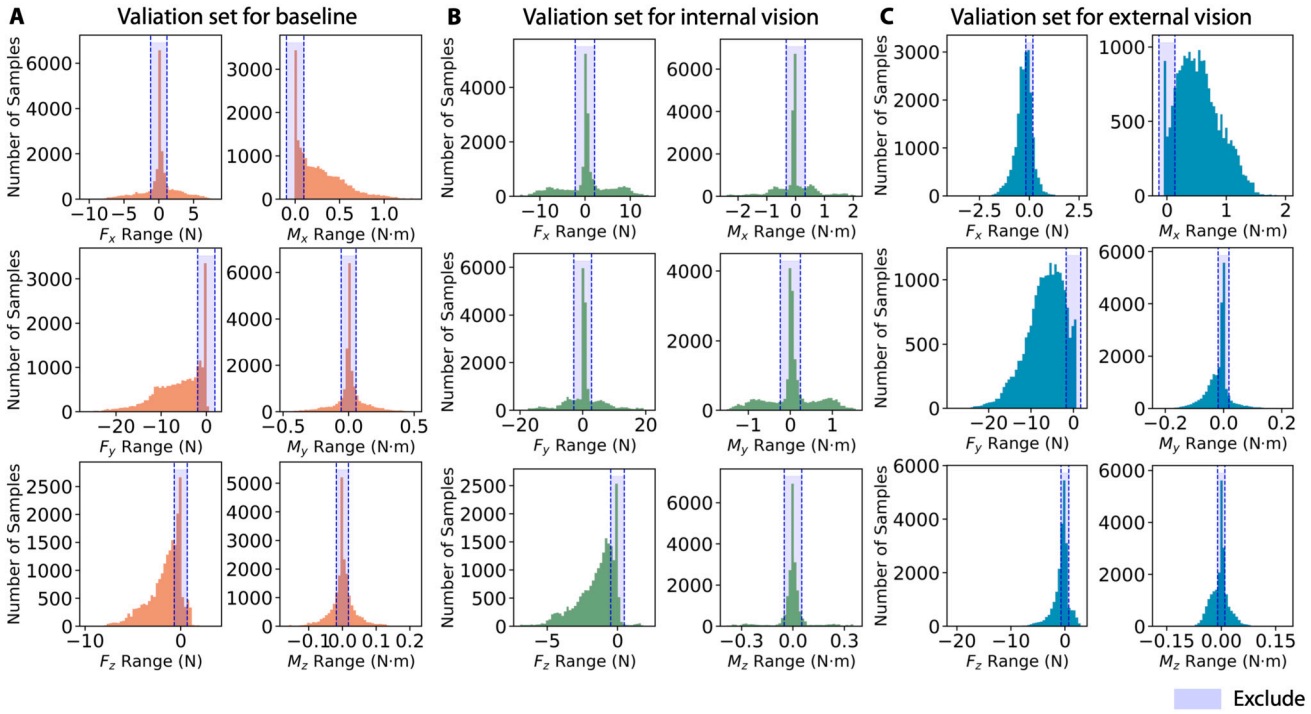


Fig. 7. Histogram of data distribution in validation sets of three experiments. The blue shadow represents the excluded range of data centered around zero.

tions. Moreover, while the performance of the *external vision* method may not match that of the *internal vision* in accuracy, it offers an appealing feature with a fully *passive* and *low-cost* design as a portable system. In practical applications, this design does not necessitate an embedded camera as its data source; it solely relies on an external camera to capture SRM tactile deformations and spatial movements. This quality makes it a cost-effective implementation, significantly reducing the overall deployment expenses.

4.2. Applications in robotic teleoperation

Integrating our designed SRM into robotics unlocks many capabilities. The *internal vision* design showcases a remarkable capacity for high-precision force and torque prediction. Our research demonstrates the programmability of the SRMs as a joystick or a push button for teleoperating a robotic arm. As users manipulate or push the SRM, it undergoes deformations in various directions, captured by the camera beneath. Our trained learning-based model accurately predicts the forces and torques applied to the SRMs based on these deformations. We can deduce the nature of the user's actions on the SRMs by harnessing these predicted forces and torques. This encompasses identifying whether the operation involves translation or rotation in a specific direction, push of a button, or even discerning the speed of these movements.

To facilitate tactile intelligence through *external vision*, we developed a 3D-printed adapter that connects the metamaterial to the end of a kitchen tong. The external USB camera plays a pivotal role in tactile sensing, allowing users to manipulate the tong manually while collecting tactile data with the SRMs. This manipulation transfers motion and force data, enabling the teleoperation of a UR10e robot arm, effectively mirroring the user's actions for completing tasks involving object grasping. The teleoperation approach facilitated by the SRM holds tremendous potential across a spectrum of robotic control, manipulation, and educational applications. It provides an intuitive means of controlling a robotic arm, suitable for tasks such as pick-and-place, assembly, sorting, and more, all while circumventing the complexities associated with traditional programming methods.

5. Conclusion, limitations, and future works

This study introduces an innovative Soft Robotic Metamaterial (SRM) with omni-directional adaptability designed for vision-based tactile sensing in robot learning applications. The SRM demonstrates remarkable flexibility in responding to external forces, establishing the foundation for a tactile sensing paradigm grounded in the vision-based observation of SRM deformations. Additionally, the SRM's fabrication is simplified through 3D printing, making it an accessible and shareable manufacturing process.

We have showcased two cost-effective tactile sensing implementations, each harnessing the SRM's visual adaptability. The *internal vision* method, featuring an embedded camera within the SRM, delivers robust tactile sensing performance across most directions, with relative errors below 10% in 5 out of 6 directions. Conversely, employing an external camera, the *external vision* approach exhibits slightly lower accuracy than the internal method. However, *external vision* eliminates the need for an onboard camera and associated wiring, substantially reducing system complexity and cost. Both of these sensing modalities hold the potential to provide flexible and reliable tactile feedback for various robotics applications, including teleoperation. This highlights the promising avenue of advanced materials and design research with soft robotic metamaterials with broad-reaching applications.

While this system has demonstrated promise in robot learning, there are areas where improvements can be made. Expanding the tactile data collection process to encompass a broader range of interaction methods beyond a pushrod could yield a more comprehensive training dataset. Furthermore, implementing lower latency communication and control options could enhance the teleoperation experience over longer distances. With these enhancements to bolster the system's capabilities and performance, the sensing approach could become more versatile, delivering more nuanced tactile interactions for advanced robot learning and teleoperation applications through materials and design.

CRediT authorship contribution statement

Tianyu Wu: Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft. **Yujian Dong:**

Data curation, Formal analysis, Investigation, Methodology, Software, Writing – original draft. **Xiaobo Liu:** Data curation, Formal analysis, Investigation, Methodology, Software, Visualization. **Xudong Han:** Formal analysis, Investigation, Validation. **Yang Xiao:** Formal analysis, Investigation, Methodology, Software, Writing – original draft. **Jinqi Wei:** Formal analysis, Investigation, Methodology, Software, Writing – original draft. **Fang Wan:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Project administration, Resources, Software, Supervision, Validation, Writing – review & editing. **Chaoyang Song:** Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

https://github.com/bionicdl-sustech/SoftRoboticTongs

Appendix A. The three datasets

Please refer to the datasets for more information, including details about our data processing methods, at this link: https://github.com/bionicdl-sustech/SoftRoboticTongs. Refer to Fig. 7 for a histogram of the data distribution.

Appendix B. Video demonstration - Supplementary material

Please refer to the supplementary movie for a demonstration of the proposed Soft Robotic Metamaterial.

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.matdes.2024.112629.

References

[1] A. Alspach, K. Hashimoto, N. Kuppuswamy, R. Tedrake, Soft-bubble: a highly compliant dense geometry tactile sensor for robot manipulation, in: 2019 2nd IEEE International Conference on Soft Robotics (RoboSoft), 2019, pp. 597–604.

[2] N. Bai, L. Wang, Q. Wang, J. Deng, Y. Wang, P. Lu, J. Huang, G. Li, Y. Zhang, J. Yang, K. Xie, X. Zhao, C.F. Guo, Graded intrafillable architecture-based iontronic pressure sensor with ultra-broad-range high sensitivity, Nat. Commun. 11 (2020) 209, https://doi.org/10.1038/s41467-019-14054-9.

[3] J.A. Barreiros, A. Xu, S. Pugach, N. Iyengar, G. Troxell, A. Cornwell, S. Hong, B. Selman, R.F. Shepherd, Haptic perception using optoelectronic robotic flesh for embodied artificially intelligent agents, Sci. Robot. 7 (2022) eabi6745, https://doi.org/10.1126/scirobotics.abi6745.

[4] B. Calli, A. Singh, J. Bruce, A. Walsman, K. Konolige, S. Srinivasa, P. Abbeel, A.M. Dollar, Yale-cmu-Berkeley dataset for robotic manipulation research, Int. J. Robot. Res. 36 (2017) 261–268, https://doi.org/10.1177/0278364917700714.

[5] H. Cui, D. Yao, R. Hensleigh, H. Lu, A. Calderon, Z. Xu, S. Davaria, Z. Wang, P. Mercier, P. Tarazaga, X.R. Zheng, Design and printing of proprioceptive three-dimensional architected robotic metamaterials, Science 376 (2022) 1287–1293, https://doi.org/10.1126/science.abn0090.

[6] E. Donlon, S. Dong, M. Liu, J. Li, E. Adelson, A. Rodriguez, Gelslim: a high-resolution, compact, robust, and calibrated tactile-sensing finger, in: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2018, pp. 1927–1934.

[7] B. Fang, X. Long, F. Sun, H. Liu, S. Zhang, C. Fang, Tactile-based fabric defect detection using convolutional neural network with attention mechanism, IEEE Trans. Instrum. Meas. 71 (2022) 1–9, https://doi.org/10.1109/TIM.2022.3165254.

[8] N. Fazeli, M. Oller, J. Wu, Z. Wu, J.B. Tenenbaum, A. Rodriguez, See, feel, act: hierarchical learning for complex manipulation skills with multisensory fusion, Sci. Robot. 4 (2019) eaav3123, https://doi.org/10.1126/scirobotics.aav3123.

[9] D.V. Gealy, S. McKinley, B. Yi, P. Wu, P.R. Downey, G. Balke, A. Zhao, M. Guo, R. Thomasson, A. Sinclair, P. Cuellar, Z. McCarthy, P. Abbeel, Quasi-direct drive for low-cost compliant robotic manipulation, in: 2019 International Conference on Robotics and Automation (ICRA), 2019, pp. 437–443.

[10] A. Ghafoor, J.S. Dai, J. Duffy, Stiffness modeling of the soft-finger contact in robotic grasping, J. Mech. Des. 126 (2004) 646–656.

[11] D.F. Gomes, Z. Lin, S. Luo, Geltip: a finger-shaped optical tactile sensor for robotic manipulation, in: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2020, pp. 9903–9909.

[12] P. Jiao, A.H. Alavi, Artificial intelligence-enabled smart mechanical metamaterials: advent and future trends, Int. Mater. Rev. 66 (2021) 365–393, https://doi.org/10.1080/09506608.2020.1815394.

[13] M. Lambeta, P.W. Chou, S. Tian, B. Yang, B. Maloon, V.R. Most, D. Stroud, R. Santos, A. Byagowi, G. Kammerer, D. Jayaraman, R. Calandra, Digit: a novel design for a low-cost compact high-resolution tactile sensor with application to in-hand manipulation, IEEE Robot. Autom. Lett. 5 (2020) 3838–3845, https://doi.org/10.1109/LRA.2020.2977257.

[14] R. Li, B. Peng, Implementing monocular visual-tactile sensors for robust manipulation, Cyborg Bionic Syst. 2022 (2022), https://doi.org/10.34133/2022/9797562.

[15] S. Li, X. Yin, C. Xia, L. Ye, X. Wang, B. Liang, Tata: a universal jamming gripper with high-quality tactile perception and its application to underwater manipulation, in: 2022 International Conference on Robotics and Automation (ICRA), 2022, pp. 6151–6157.

[16] Q. Ma, T.J. Cui, Information metamaterials: bridging the physical world and digital world, PhotoniX 1 (2020) 1, https://doi.org/10.1186/s43074-020-00006-w.

[17] M.A. McEvoy, N. Correll, Materials that couple sensing, actuation, computation, and communication, Science 347 (2015) 1261689, https://doi.org/10.1126/science.1261689.

[18] D. Mukherjee, K. Gupta, L.H. Chang, H. Najjaran, A survey of robot learning strategies for human-robot collaboration in industrial settings, Robot. Comput.-Integr. Manuf. 73 (2022) 102231, https://doi.org/10.1016/j.rcim.2021.102231.

[19] D.N. Neshev, A.E. Miroshnichenko, Enabling smart vision with metasurfaces, Nat. Photonics 17 (2023) 26–35, https://doi.org/10.1038/s41566-022-01126-4.

[20] R. Ouyang, R. Howe, Low-cost fiducial-based 6-axis force-torque sensor, arXiv:2005.14250, 2020.

[21] S. Pyo, J. Lee, K. Bae, S. Sim, J. Kim, Recent progress in flexible tactile sensors for human-interactive systems: from sensors to advanced applications, Adv. Mater. 33 (2021) 2005902, https://doi.org/10.1002/adma.202005902.

[22] J. Ramos, S. Kim, Dynamic locomotion synchronization of bipedal robot and human operator via bilateral feedback teleoperation, Sci. Robot. 4 (2019) eaav4282, https://doi.org/10.1126/scirobotics.aav4282.

[23] H. Ravichandar, A.S. Polydoros, S. Chernova, A. Billard, Recent advances in robot learning from demonstration, Annu. Rev. Control Robot. Auton. Syst. 3 (2020) 297–330, https://doi.org/10.1146/annurev-control-100819-063206.

[24] Y. She, S.Q. Liu, P. Yu, E. Adelson, Exoskeleton-covered soft finger with vision-based proprioception and tactile sensing, in: 2020 IEEE International Conference on Robotics and Automation (ICRA), 2020, pp. 10075–10081.

[25] H. Sun, K.J. Kuchenbecker, G. Martius, A soft thumb-sized vision-based sensor with accurate all-round force perception, Nat. Mach. Intell. 4 (2022) 135–145, https://doi.org/10.1038/s42256-021-00439-3.

[26] F. Wan, X. Liu, N. Guo, X. Han, F. Tian, C. Song, Visual learning towards soft robot force control using a 3d metamaterial with differential stiffness, in: A. Faust, D. Hsu, G. Neumann (Eds.), Proceedings of the 5th Conference on Robot Learning, PMLR, 2022, pp. 1269–1278, https://proceedings.mlr.press/v164/wan22a.html.

[27] B. Wang, W. Guo, S. Feng, Y. Hongdong, F. Wan, C. Song, Volumetrically enhanced soft actuator with proprioceptive sensing, IEEE Robot. Autom. Lett. 6 (2021) 5284–5291, https://doi.org/10.1109/LRA.2021.3072859.

[28] X. Wang, Z. Meng, C.Q. Chen, Robotic materials transformable between elasticity and plasticity, Adv. Sci. 10 (2023) 2206637, https://doi.org/10.1002/advs.202206637.

[29] W. Xu, H. Zhang, H. Yuan, B. Liang, A compliant adaptive gripper and its intrinsic force sensing method, IEEE Trans. Robot. 37 (2021) 1584–1603, https://doi.org/10.1109/TRO.2021.3060971.

[30] Y. Yan, Z. Hu, Z. Yang, W. Yuan, C. Song, J. Pan, Y. Shen, Soft magnetic skin for super-resolution tactile sensing with force self-decoupling, Sci. Robot. 6 (2021) eabc8801, https://doi.org/10.1126/scirobotics.abc8801.

[31] L. Yang, X. Han, W. Guo, F. Wan, J. Pan, C. Song, Learning-based optoelectronically innervated tactile finger for rigid-soft interactive grasping, IEEE Robot. Autom. Lett. 6 (2021) 3817–3824, https://doi.org/10.1109/LRA.2021.3065186.

[32] Z. Yang, S. Ge, F. Wan, Y. Liu, C. Song, Scalable tactile sensing for an omni-adaptive soft robot finger, in: 2020 3rd IEEE International Conference on Soft Robotics (RoboSoft), 2020, pp. 572–577.

[33] W. Yuan, S. Dong, E.H. Adelson, Gelsight: high-resolution robot tactile sensors for estimating geometry and force, Sensors 17 (2017), https://doi.org/10.3390/s17122762.

[34] S. Zhang, Y. Sun, J. Shan, Z. Chen, F. Sun, Y. Yang, B. Fang, Tirgel: a visuo-tactile sensor with total internal reflection mechanism for external observation and contact detection, IEEE Robot. Autom. Lett. 8 (2023) 6307–6314, https://doi.org/10.1109/LRA.2023.3306670.

[35] H. Zhao, K. O'Brien, S. Li, R.F. Shepherd, Optoelectronically innervated soft prosthetic hand via stretchable optical waveguides, Sci. Robot. 1 (2016) eaai7529, https://doi.org/10.1126/scirobotics.aai7529.