

GTac-Hand: A Robotic Hand With Integrated Tactile Sensing and Extrinsic Contact Sensing Capabilities

Zeyu Lu  and Haoyong Yu , Senior Member, IEEE

Abstract—In many interaction and manipulation tasks, robots need to estimate the state or properties, such as extrinsic contact states (ECSs), on the object being manipulated. Tactile sensing is an essential sensor modality for robots to acquire this relevant information. In this article, we present a robotic hand integrated with tactile sensors that combine two force sensing principles as skin-like heterogeneity, piezoresistive, and Hall effect, and can obtain normal and shear contact force feedback simultaneously from the fingers and palm, resulting in 285 tactile sensing measurements at a rate of 150 Hz. In the contact-rich manipulation task, there would be complicatedly coupled contact relationships between the robotic hand and the grasped objects during external contact. Sectionwise features (SWFs) are implemented to leverage these tactile features by means of incorporating distributed forces, and acquiring the dynamic time-varying rate and tactile events. A quantitative evaluation of the performance of recognizing nine ECSs is provided such as 99.6% accuracy and the SWFs resulted in a significant reduction (30%) in the error rates. Given the human-like tactile sensing integration design and high-quality real-time models of patterning and learning, the robotic hand can perform challenging tasks, such as delicate object grasping, object handovers, and ball-hit recognition.

Index Terms—Force and tactile sensing, human–robot interaction, multifingered hands.

I. INTRODUCTION

HUMAN hands can effectively perform daily activities, such as grasping and handovers of fragile objects, utilizing the tactile sensing capabilities which simultaneously perceive normal and shear forces via the mechanoreceptors highly integrated into the fingers and palm. In contact-rich manipulation

Manuscript received 10 August 2022; revised 22 November 2022 and 30 January 2023; accepted 25 March 2023. This work was supported by Agency for Science, Technology and Research, Singapore under the National Robotics Program, with A*star SERC under Grant 192 25 00054. Recommended by Technical Editor M. Grebenstein and Senior Editor H. Qiao. (Corresponding author: Haoyong Yu.)

The authors are with the Department of Biomedical Engineering, National University of Singapore, Singapore 117575 (e-mail: zeyu.lu@u.nus.edu; biehy@nus.edu.sg).

This article has supplementary material provided by the authors and color versions of one or more figures available at <https://doi.org/10.1109/TMECH.2023.3264650>.

Digital Object Identifier 10.1109/TMECH.2023.3264650

scenarios where human hands are holding an object, a single external contact stimulus imposed on it may cause a complex combination of normal and shear contact forces to act on multiple tactile sensing units distributed over the hands. Humans can interpret the contact states of held objects imposed by external stimulus through tactile sensing feedback; this is referred to as the extrinsic contact states (ECSs) identification capability [1]. Manipulating or interacting with unseen objects would be related to many robotic tasks. To perform those tasks effectively and efficiently, robots need to identify some properties, such as ECSs, on the object in contact-rich manipulation. The contacts are not directly on the tactile sensors but on the held object. Those information can be obtained by the using tactile sensors on the end-effectors of robots via their raw signals or extracting relevant features [2].

The sense of touch in human hands is perceived by four different mechanoreceptors, i.e., fast-adapting type I and II (FA-I and FA-II), and slow-adapting type I and II (SA-I and SA-II) [3]. FA-I and SA-II types of mechanoreceptors are mainly involved in perception of normal and shear forces, respectively. The second-order neurons in the spinal cord can collect valid and essential time-varying patterns and tactile events via neuronal afferent associated with the FA-I and SA-II processing of the large somatosensory information part of synchronous inputs before they are processed by the human brain (called somatosensory pathways); this increases the efficiency and accuracy of central neurons in the brain in planning and executing manipulation tasks [3], [4], [5]. For humans to accurately perform the contact-rich manipulation tasks, the sensory-motor controller requires the identification of the contact states of the objects along with the involved external environment, according to the decoded tactile patterns and events [3]. Due to the rich tactile sensory feedback on their fingertips and palm and the somatosensory system that efficiently processes tactile features, humans can grasp objects, perceive their contact states, and react accordingly when performing manipulation tasks (e.g., grasping and handover) [6]. Therefore, combining multiple modalities, while retaining high integrability and contact force sensing capabilities like those of human-skin, and interpreting the ECSs of the grasped object from leveraging a large amount of tactile feedback, are critical to realize human-like manipulation capabilities with robotics.

Tactile sensing is an essential part of dexterous robotic manipulation and has been widely integrated into robotic hands.

Meanwhile, coverage and shear force sensing have emerged as challenges of tactile robotics [2], [7], [8]. Some previous review papers well be summarized a number of types of tactile sensors and sensor integration with robotic end-effectors [9], [10], [11]. There are various tactile sensors adopting different sensing principles including piezoresistive, capacitive, magnetic, barometric, piezoelectric, and optical. Piezoresistive and capacitive sensors are generally more customizable regarding spatial resolution, sensitivity, and shape, but they would only measure normal forces. Magnetic and optical sensors are able to measure shear forces. However, magnetic sensors could be disturbed by magnetic objects and optical sensors are limited by their bulkiness in some cases of integration with space-constrained robotic hands. There are some multifingered robotic hands/grippers, including underactuated, that have been integrated with tactile sensors only at fingers [12], [13], [14] and some that can only perceive normal forces [15]. There are also some alternative tactile sensing methods used in robotic hands, such as intrinsic [16] and proprioceptive [17]. Researchers have frequently focused on interpreting the tactile feedback. For instance, robotic hands integrated with tactile sensors can perform haptic object identification by extracting tactile features of grasped objects using Bayesian-based method [18]. When robotics are interacting with objects and external environment for contact-rich tasks, such as peg insertion, deep neural networks model can efficiently learn from complex features that are not able to be modeled [19]. Heyneman et al. [20] implemented the dynamics event classification on tactile array data to discriminate between slips of object/environment or hand/object. Calandra et al. [21] and Chebotar et al. [22] extracted tactile features using learning-based models to classify grasping success probability. In such previous works, data-driven models are employed to implicitly process tactile feedback to learn the mapping between lower level features (e.g., raw sensing signals) to object-level features (e.g., slippage event) [2]. Ma et al. [1] explored the localization of external contacts on the held object, which was addressed by kinematic constraint-based analytical approaches, where the tactile features, local motion field at contact, were extracted. A given threshold is also commonly used to detect contact events. For instance, Romano et al. [23] generated human-like SA-I, FA-I, and FA-II tactile features and events from tactile feedback using calculation of threshold, sum, and magnitude with previous contact experiences to perform a series of robotic tasks. Nonetheless, analytical approaches generally depend on precise models of structured interactions and assumption. Since contacts could occur at any direction and location in an unstructured environment, some previous works about tactile features/events interpretation are constrained by the richness of lower level signals, e.g., only normal force. Considering tactile features with directional, spatial-temporal, and quantified properties would facilitate to interpret more comprehensive object-level information, such as ECSs, from complex tactile feedback. Human somatosensory systems leverage complex tactile feedback into patterns and identify ECSs of objects in contact-rich tasks. Similarly, these features are still open issues and are critical for robots in acquiring such human skills.

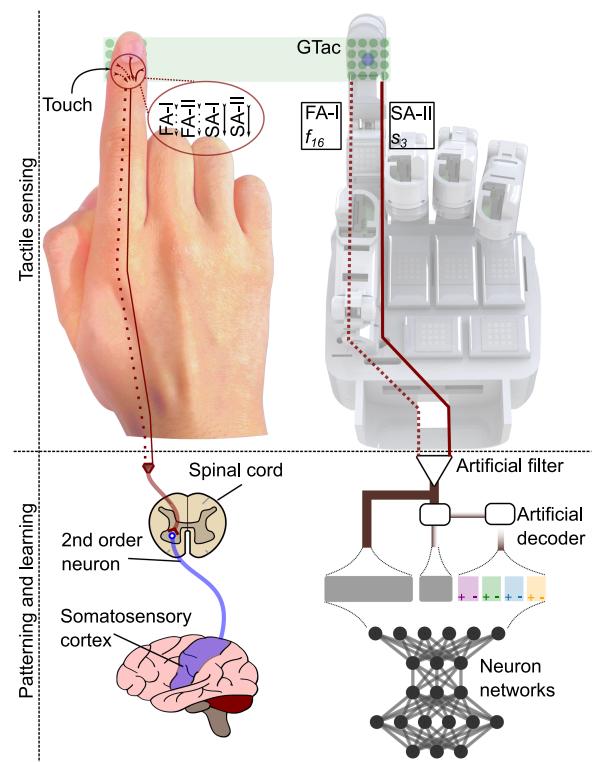


Fig. 1. Illustration of the tactile sensing and processing schematics in a human hand analogous to the GTac-Hand.

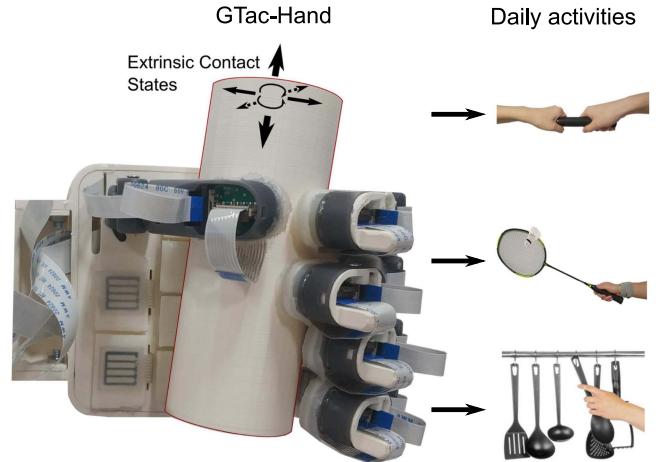


Fig. 2. Left: GTac-Hand holds a cylinder-shaped bar for ECS capability identification; Right: Example photographs of daily human activities implementing cylindrical object grasping and ECS capabilities (from top to bottom), including bar handovers, racket sports, and hanging tools.

Therefore, developing a robotic sensory system consisting of i) effective sensory interpretation methods such as those of the human somatosensory system (patterning and learning as illustrated in Fig. 1), and ii) tactile sensors with human skin-like normal force (distributed) and shear force (gross) sensing capabilities is crucial for robots expected to acquire human skills [3], [9]. When grasping and manipulating an object, humans can leverage a large amount of and complicated coupled tactile force feedback to identify the ECS of the held object (see Fig. 2) to perform daily tasks such as handover, racket sports, and

hanging tools. Motivated by this fact, we have developed an anthropomorphic robotic hand, GTac-Hand, integrated with the heterogeneous tactile (GTac) sensors, which is a biomimetic multimodal tactile sensor that we previously built for distributed normal force and shear force sensing in [24] into the fingers and palm. Thus, the proposed GTac-Hand can acquire 285 tactile signals consisting of normal and shear force. Our results show that GTac-Hand could accurately identify the nine typical ECSs of grasped objects in real time via human-like patterning and deep learning model. Similar to the human sensory-motor controller [3], the GTac-Hand can perform tasks such as closed-loop grasping and object handovers according to the identified ECS of the grasped object.

II. METHODS

A. Tactile Feedback Acquisition and Processing

The GTac sensors [24] are integrated into each phalanx of fingers and the palm of the GTac-Hand. Inspired by the FA-I and SA-II mechanoreceptors in the multilayered structure of human skin and their force sensing function, each GTac sensor consists of 4×4 matrix of piezoresistive sensors for normal force sensing (refer to FA-I layer) and a 3-D magnetic sensor realized by a Hall sensor for gross 3-D forces sensing (refer to SA-II layer). The average root mean square error (RMSE) and maximum hysteresis error of the triaxis force sensing were 0.11 N and 2%. The GTac-Hand could obtain 285 tactile signals from the fingers and palm. The sensing PCB can acquire all the signals at 150 Hz. First, we collected $L_1 = 300$ samples and calculated the mean value of the tail 100 samples (\bar{X} , 0.67 s data) to initialize GTac signals. According to the two sensing principles of GTac sensors, the change of analog signals and magnetic field density are used to estimate the contact forces. The relative values of the tactile signals can be acquired by subtracting the mean values \bar{X} . We added a moving average (window size: 6) to reduce the noise of the signals. We implemented a filtering algorithm to reduce the magnetic disturbance to the SA-II signals of GTac. This signal processing algorithm was used only in the closed-loop grasping phase. Two thresholds on both the FA-I and SA-II data on each GTac unit were used to estimate the contact status. If either of the leveraged FA-I signals

$$\tilde{g}_{s,f|FA-I} = \sum_{r=1}^4 \sum_{c=1}^4 R_{s,f}^{r,c} \quad (1)$$

or the SA-II signals

$$|\tilde{g}_{s,f|SA-II}| = \sqrt{\Delta B_{s,f}^x{}^2 + \Delta B_{s,f}^y{}^2 + \Delta B_{s,f}^z{}^2} \quad (2)$$

was above the thresholds (T_1 and T_2), the section s of a finger f was regarded as contacted. $\Delta B^{x,y,z}$ is the 3-D magnetic force feedback from the SA-II layer and $\Delta B^{x,y}$ are the feedback of shear forces and ΔB^z is the part of the normal force. Index $f \in [0, 1, 2, 3, 4]$ and $s \in [0, 1, 2]$ indicate the GTac sensors in each phalanx (see Fig. 6). $R^{r,c}$ are the matrix of normal force signals in row (r) and column (c) from the FA-I layer. The SA-II signals were valid only when the responding finger section was in contact. Therefore, the change in magnetic flux density due

Algorithm 1: GTac Signal Processing.

INITIALIZATION

$i = 0$

while $i < N_0$ **do**

$g_i^{raw} = readout()$

$i = i + 1$

$\bar{X} = \sum_{i=-L_1/3}^{-1} g_i^{raw} / (L_1/3)$

MAIN LOOP

while true **do**

$g^{raw} = readout()$

$\hat{g} = g^{raw} - \bar{X}$ (initialized signals)

 put \hat{g} in tail of moving average g^6 , then

$\tilde{g} = \sum g^6 / 6$, (filtered signals).

 if $|\tilde{g}_{s,f|FA-I}| < T_1$ and $|\tilde{g}_{s,f|SA-II}| < T_2$ then

$\tilde{g}_{s,f|SA-II} = 0$

to the finger orientation relative to the global coordinate was discarded when no contact was detected. The pseudocode for data processing is in Algorithm 1.

B. Patterning and Learning Models

1) Leveraging Tactile Feedback: Based on the characteristics of GTac, we converted the 285 tactile sensing signals, consisting of 240 FA-I type signals and 45 SA-II type signals from 15 GTac units ($GTac\# = f \times 4 + s$, $f \in \{0, 1, 2, 3, 4\}$, $s \in \{0, 1, 2\}$) to a 15×19 signal matrix (Feature #: 0–18). According to the signaling scheme of GTac, encoding neuron-inspired tactile representations is suitable for implementation in it [3] because GTac can incorporate both FA-I type and SA-II type tactile signals while maintaining synchronized temporal precision in each finger section. For instance, the contact location can be estimated by human fingers. Here, we incorporated the tactile signals from the FA-I layer of GTac sensor to estimate the features of contact location (x_e and y_e) of which the performance has been validated in [24].

$$\begin{bmatrix} x_e \\ y_e \end{bmatrix} = \begin{bmatrix} \sum_{r=1}^4 \sum_{c=1}^4 (e \cdot c \cdot R^{r,c}) / 16 \\ \sum_{r=1}^4 \sum_{c=1}^4 (e \cdot r \cdot R^{r,c}) / 16 \end{bmatrix} \quad (3)$$

where $R^{r,c}$ is tactile feedback from the FA-I layer at row r and column c , and the FA-I spatial resolution is $e = 2.5$ mm. Inspired by neural tactile pattern representations in the human somatosensory system [3], we used several decoders to extract biomimetic tactile information including incorporating distributed FA-I signals (SR), integrating FA-I and SA-II signals ($SFA^{x,y,z}$), obtaining the dynamic time-varying rate (dFA & $dSA^{x,y,z}$), and producing tactile events at the n_{th} sampling moment (Fig. 3). The tactile events of the corresponding layer were captured when the time-varying rate, i.e., dFA and $dSA^{x,y,z}$, exceeded the boundary in a predefined threshold, where $|F| = 20$ for the FA-I layer and $|S^i| = 20$, $i \in \{x, y, z\}$ for the SA-II layer in the tri-axis

$$\begin{bmatrix} SR = \sum_{r=1}^4 \sum_{c=1}^4 R^{r,c} \\ SFA^{x,y,z} = \Delta B^{x,y,z} / SR \end{bmatrix} \quad (4)$$

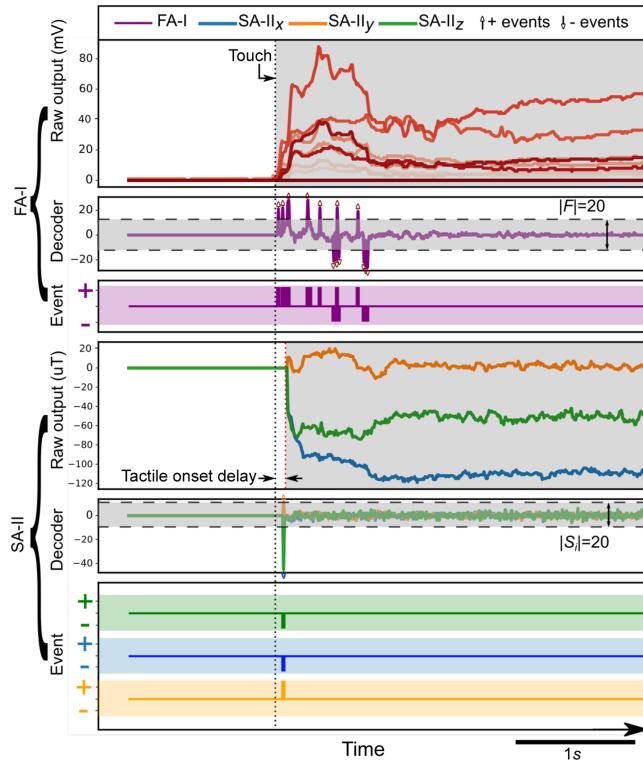


Fig. 3. GTac data flow at the index distal finger section when the GTac-Hand is grasping a cylinder-shaped object in a force-closed manner.

$$\begin{bmatrix} dFA = SR_n - SR_{n-1} \\ dSA^{x,y,z} = \Delta B_n^{x,y,z} - \Delta B_{n-1}^{x,y,z} \end{bmatrix}. \quad (5)$$

Specifically, the extracted tactile information was independent of each finger section. Therefore, we referred to them as sectionwise features (SWFs). As shown in Fig. 4(a), we had 14 SWFs consisting of ten decoder outputs and four tactile events as the optional input (w/ SWF or w/o SWF) of a machine learning-based classifier with the original 19 GTac signals. In total, there were 33 channels of data, namely flattened $R^{r,c}$ (Feature #: 0–15), $\Delta B^{x,y,z}$ (Feature #: 16–18), x_e and y_e (Feature #: 19–20), SR (Feature #: 21), dFA (Feature #: 22), $SFA^{x,y,z}$ (Feature #: 23–25), $dSA^{x,y,z}$ (Feature #: 26–28), and tactile events (Feature #: 29–32) as shown in Fig. 4.

2) Supervised Learning Model Training and Validation: To identify the nine ECSs via tactile feedback, two supervised learning models, i.e., convolutional neural networks (CNNs) and quadratic discriminant analysis (QDA), were implemented. The Keras library in Python, based on TensorFlow, was used to construct and train the CNN-based model [see Fig. 4(b)]. After tuning the parameters of the CNN model to reduce its complexity (number of trainable parameters) and maintain its performance (validation accuracy), we established an architecture consisting of two main convolutional blocks, a flattened layer regulated with a “dropout” rate of 0.5 and a fully connected dense layer. Each convolutional block was assembled by a 3×3 convolutional kernel, a rectified linear unit (ReLU), to introduce nonlinearity, and a 2×2 max-pooling layer. There were 2121 and 2985 trainable parameters in the CNN model for the input

options, w/o SWF and w/ SWF, respectively. During the learning process, we set the learning rate as 0.001 and batch size as 12 using the Adam optimizer and trained the model for 15 epochs (about 100 s spent). To compare the performance with the CNN-based model, the QDA-based classifier was implemented using the scikit-learn library that determines the quadratic decision boundary between the classes in the feature spaces based on the Bayes theorem. The configuration of the QDA model was the default configuration. Similarly like Calandra et al. [21], which has applied CNN with tactile feedback in a form of an array, the input of the supervised learning models is in a compact manner (15×33) by mixing two different sensing modalities (FA-I and SA-II). The features with different signal output ranges would be leveraged by the weights of the models and be used to conduct the identification. Because our approach learns features end-to-end entirely from raw inputs, no prior model is required. To verify the supervised learning models, we implemented fivefold cross-validation, wherein the dataset was split into five equal pieces after randomization, 80% were used to train the model, and the remaining 20% was used to validate the trained model. This estimation process was iterated five times.

C. Mechatronic Design of the GTac-Hand

1) Mechanical Design: The anthropomorphic hand is underactuated and cable driven (see Fig. 5), where ten rotational joints are actuated by six motors. The shape, size, kinematics, and appearance are similar to those of the human hand. The robotic hand integrates all the electronics and actuation transmissions within the hand. A cable-driven mechanism for underactuation is a common design for a robotic hand in a confined space [25]. Each rotational joint consists of a tension spring, two bearings, and finger phalanges structures for embedding the PCB of GTac. The index, middle, ring, and little fingers have the same design, except for the varied length of the distal section and have two rotational joints, namely proximal interphalangeal joint and metacarpophalangeal joint. To reduce the redundancy of the actuated joints for grasping and manipulation, the distal and middle phalanges are connected for each finger. For the thumb, there are two degrees of freedom, namely the radial abduction/adduction of the carpometacarpal joint and the flexion/extension of the metacarpophalangeal joint, to obtain human-like dexterity and mechanical synergies. The actuation mechanism is separated from the electronics for ease of debugging and assembly.

2) Electronics Design: The robotic hand integrates the electronics for tactile sensing and actuation on the wrist. There are two PCBs, where one PCB is used for collecting the signals from the tactile sensors (see Figs. 5 and 6), sensing PCB, and another for power supply, actuation PCB. The sensing PCB has five 20-pins jump cable connectors for each GTac-finger branch. A microcontroller (Teensy 4.1) is part of the sensing PCB to produce the digital data and functions as a low-level controller of the motors and opens the serial port for communication with the core processing unit. The single-pole-double-throw (SPDT) is an ADG333 A integrated with four SDPTs. The 16-channel analog multiplexer is CD74HC4067. The 4-to-1 analog switch was TC4052BF. The TCA9548 A 8-to-1 I2C switch was used.

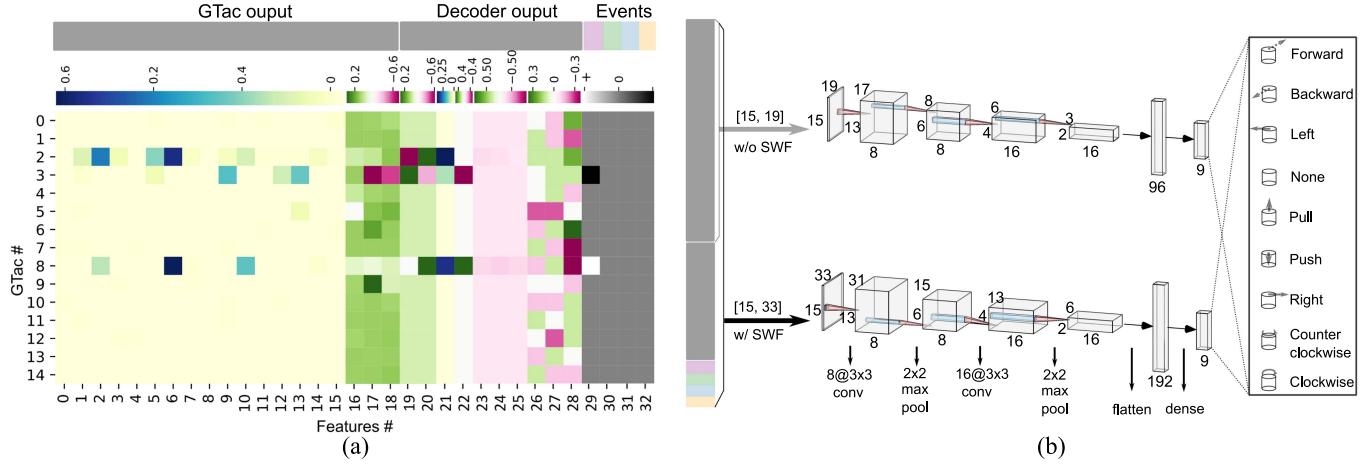


Fig. 4. (a) Sample of complete set of normalized tactile sensing signals (color bars show the scales and the corresponding feature channels) in the 15×33 matrix. (b) Illustration of the CNN architectures used to identify the nine dominant ECSs; the dimensions of the neuron kernel in each layer are indicated.

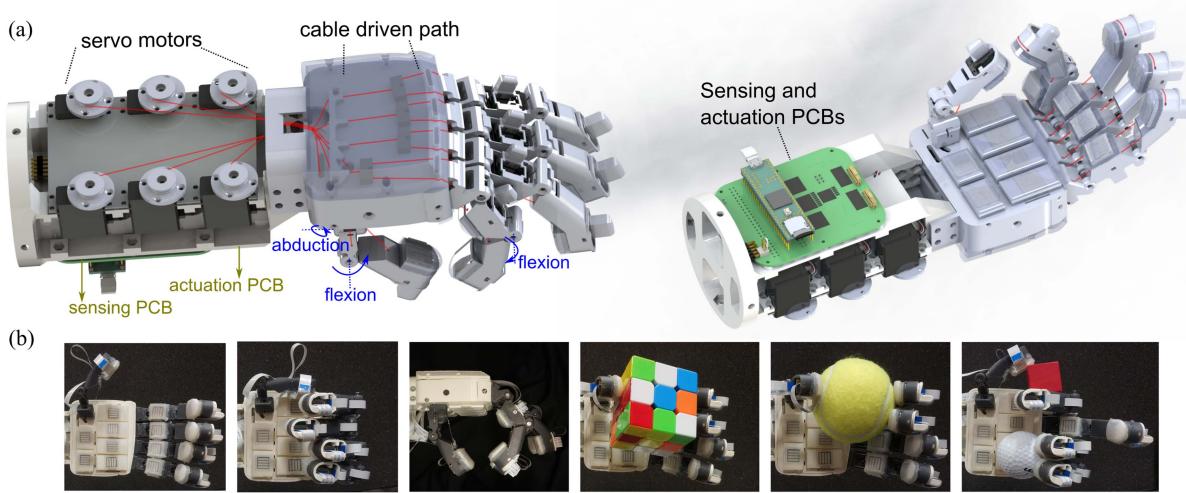


Fig. 5. (a) Mechanical design of the GTac-Hand. (b) GTac-Hand performing various gestures and grasping everyday objects. From left to right, an open gesture, a fist gesture, a pinch gesture, a pinch grasp of a Rubik's Cube and a tennis ball, and a multiobjects grasp comprising a lateral grasp of a wood cube and an ulnar storage grasp of a golf ball.

We used 20-pin 0.5-mm pitch FFC cables to connect each finger branch with the sensing PCB. The actuation PCB mainly powers the six servo motors of the GTac-Hand by connecting to an 8-V dc power supply. It regulates the 8-to-5 V power supply in the sensing PCB. There is a 10-pin jumper cable connecting the sensing PCB and the actuation PCB for the transmittance of the PWM signals and the 5-V dc power supply.

3) Sensing Solution: Each finger consists of the distal, intermediate, and proximal phalanges. All fingers are connected to the palm in parallel. For tactile sensing, mechanoreceptors are primarily distributed at the finger and palm of the hand. Based on its distribution in human hands, the tactile sensors, GTac [24], are integrated into the distal and proximal finger sections and the palm. The tactile sensors are connected serially in each finger. The data are transmitted between the fingers in parallel, as shown in Fig. 6(a). Each finger section can obtain 19 tactile feedback

signals from the GTac sensor (16 from the FA-I layer and 3 from the SA-II layer), thus the robotic hand can perceive 285 tactile sensing signals in total. The frequency of the cycle loop, including tactile sensing readout and sending motor commands, is 150 Hz for the GTac-Hand.

4) GTac Circuits and PCB Design: The PCB design of GTac [24] is anthropomorphic and can be integrated with an anthropomorphic robotic hand. In each GTac PCB, the power source is 5 V and is uploaded serially by the connecting cables. The integrated chips used in each PCB are MLX90393 (Hall sensor) and MCP6004 (4-channel operational amplifiers). As the signals are sensitive to the direction and position of the Hall sensor, the Hall sensor is placed at the center of each PCB. At the back, there are 12-pin, 16-pin, and 20-pin 0.5-mm-pitch jumper cable connectors on the distal, middle and proximal PCB, respectively, for uploading power and downloading digital

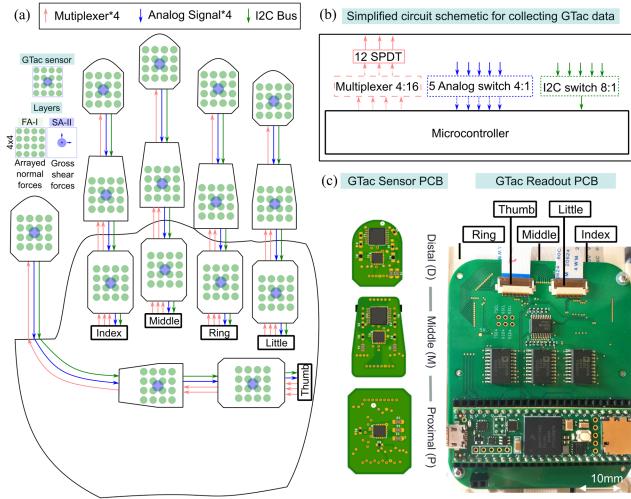


Fig. 6. (a) Illustration of the GTac sensors integration and data transmission. (b) Simplified circuit schematic for collecting GTac data from five finger branches. (c) The PCB layout of GTac sensors and readout.

and analog signals [see Fig. 6(b) and (c)]; 12-pin 1-mm-pitch cable connectors were used to connect the 4×4 matrix of piezoresistive sensors.

D. Statistical Analysis Configurations

We implemented one-way analysis of variance (ANOVA) in one degree-of-freedom (DoF) to verify the significance of the statistical results by calculating the p-value.

III. EXPERIMENTS

A. Experimental Setup

The GTac-Hand was connected to a laptop via a serial port. In the ECS identification experiments, the hollow cylinder bar used for the dataset collection and real-time verification was 3-D printed with Ultimaker S5 and its height, outer radius, thickness, and weight were 102 mm, 40 mm, 2 mm, and 26.2 g, respectively. In the egg grasping and handover experiments, the robotic arm moved in a predetermined trajectory and the egg was 40 g. In the ball-hit recognition experiments, the weight of the shuttlecock was 5.2 g. Two shuttlecocks were hanging on two separate fix points by wires each of length 0.52 m. The shuttlecock was released and swung in a pendulum-like motion with an amplitude of 45° to impose the hit on the left/right half of the grasped racket. The racket was mounted with a cylinder bar (outer radius = 43 mm) for the robotic hand to grasp it stably.

B. Dataset Collection

To realize the ECS identification capabilities, we selected nine typical ECSs for a grasped cylindrical object: forward, backward, left, right, none, pull, push, twist clockwise, and twist counterclockwise, in which way the external stimulus was exerted. During data collection, we manually applied the external stimulus at the top of the held object in an intuitive way, and the object is a 3-D printed cylinder bar as shown in Fig. 2.

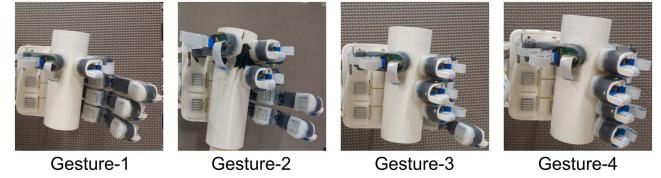


Fig. 7. Four grasp gestures used in dataset collection for NN-ECS model training and validation.

To have a larger diversity in the distribution of dataset and for the sake of real applications, four different gestures (see Fig. 7) were used to collect the same amount of data, 6500 samples for each ECS, to balance the dataset. The total time spent on data collection was about 26 min cumulatively. Therefore, the dataset consisted of 234 000 frames.

C. Control and Recognition Strategy

The control loop for the handover experiments consists of four phases, as shown in Fig. 11(a). The first phase involved grasping the object in a force-closed loop (closed-loop grasping). We implemented a simple threshold (T_g) to achieve a gentle grasping force based on feedback from the integrated GTac. The flexion of each finger was conducted by the corresponding motor (m_f) by rotation of 1 increment (1.5) until the leveraged GTac signals reaching the threshold $\bar{g}_f > T_g$

$$\bar{g}_f = \sum_{s=1}^3 \sqrt{\Delta B_{s,f}^x{}^2 + \Delta B_{s,f}^y{}^2 + \left[a\Delta B_{s,f}^z + (1-a) \sum_{r=1}^4 \sum_{c=1}^4 R_{s,f}^{r,c} \right]^2} / 3. \quad (6)$$

There is a redundancy in terms of the force sensing along the normal direction, thus we use the variable $a = 0.3$ to weight its estimation by the FA-I layer and the SA-II layer. The resultant force estimation was averaged for each finger regardless of the physical units of the two signals in digital form. We set $T_g = 100$ in this study. The gesture used to grasp the object determined which finger flexion was to be conducted first, such as the index finger m_i and middle finger m_m for Gesture-2. Since the fingers were cable-driven and underactuated, the closed-loop control algorithm ceased finger flexion via leveraged tactile signals when $\bar{g}_f > T_g$ in each finger. This was implemented similarly for all fingers except the thumb, i.e., thumb flexion m_t that was conducted after the flexion of the other fingers. The second-phase pre-ECS initialized GTac signals for ECS identification when the object was being grasped (completion of the first phase). The mean average values \bar{X} obtained in this phase were subtracted by data flow. The third phase identified the ECS in the control loop. Inspired by tactile neural representations in the human somatosensory system for closed-loop tasks [3], we implemented artificial decoders to obtain the SWFs. The filtered tactile signals and SWFs were inputted simultaneously to the learned NN-ECS model to identify the nine dominant ECSs of the grasped object. The identification results were simultaneously inputted in the fourth phase, i.e., task-oriented conditioner. The currently identified ECS of the grasped object was used as inputs to different task-oriented conditioners to complete different tasks.

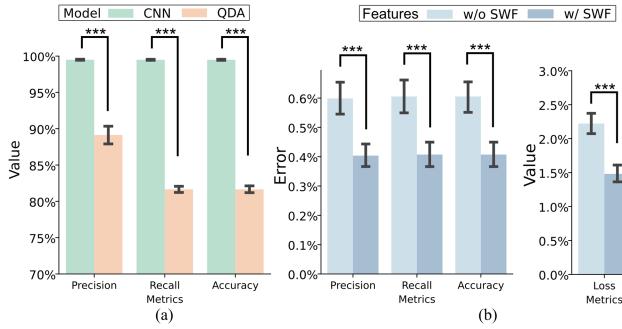


Fig. 8. (a) CNN-based model significantly outperforms the QDA model in terms of precision, recall, and overall accuracy. (b) Input w/ SWF or w/o SWF results in a significant reduction of the error rate of precision, recall, and accuracy and a reduction of the categorical entropy loss. $*p < 0.1$, $**p < 0.01$, $***p < 0.001$. 95% CI is shown by the whiskers.

1) Object Handover: If the ECS output was switched from “none” to interested ECS and remained at T cycles ($T = 10$ in handover experiments), then the handover decision would be triggered. We set the interested ECS as “pull” and “push.”

2) Ball-Hit Recognition: The recognition results were determined by the identified ECS that appeared first after the switch from “none.” Since the shuttlecock hit exerted an instant twist torque on the holder that could produce a tactile response from GTac and be identified by the NN-ECS classifier as “twist clockwise” and “twist counterclockwise,” the ECS identification results switched from “none” lead to “left hit” and “right hit” as the recognition results, respectively, in our experimental setup. The ECS identification results in the subsequent 1 s were neglected as the racket needed to recover from its shaking motion after the last bounce.

IV. RESULTS

A. Learning ECS Identification Capability

Inspired by the human brain’s processing of multisensory information from the somatosensory system in neuron networks and the ability to accurately recognize contact states of a grasped object, we designed and trained a CNN classifier to identify the nine typical ECSs without relying on analytical modeling. The shape of the neurons in each layer varied from the two different inputs (w/ SWF or w/o SWF), as shown in Fig. 4(b). To clarify the choice of using a CNN-based classifier in this task, we compared its performance with a classical machine learning model, QDA, where the inputs of both the classifiers were w/o SWF when using the same dataset explained above. The results are presented in Fig. 8(a). The proposed CNN-based model significantly outperforms the QDA-based model in terms of precision [$p = 7.6 \times 10^{-12} (< 0.001)$, $F = 2.4 \times 10^2$, 95% confidence interval (CI)], recall [$p = 1.3 \times 10^{-23} (< 0.001)$, $F = 5.2 \times 10^3$, 95% CI], and accuracy [$p = 1.3 \times 10^{-23} (< 0.001)$, $F = 5.2 \times 10^3$, 95% CI]. To evaluate how the extracted SWFs improved the ECS identification performance, the datasets of the two options (w/ SWF or w/o SWF) were used to train and validate the stated CNN model. The results [see Fig. 8(b)] showed that the SWFs resulted in a significant reduction (30%) in the error

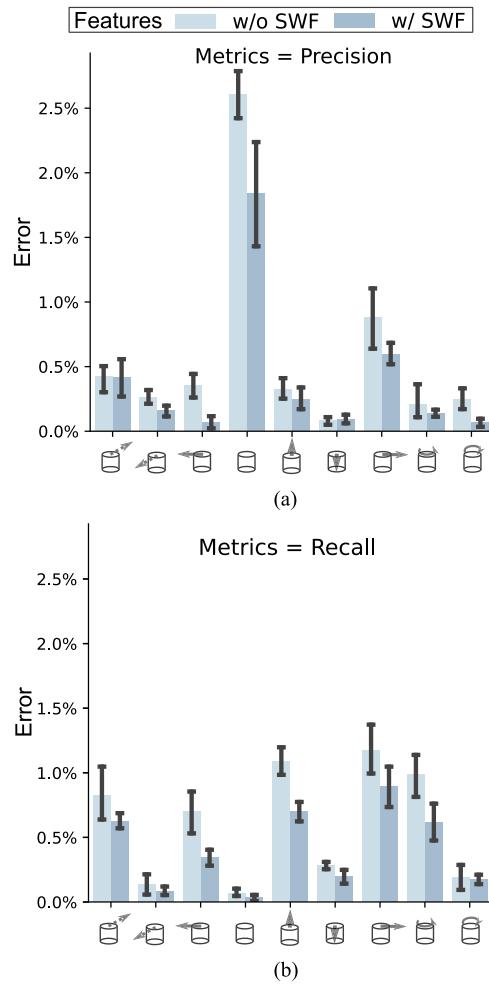


Fig. 9. Precision error and recall error of the CNN model, upon identifying each ECS. Five-fold cross-validation is applied for all validation.

rates ($E = 100\% - N_t/N_{val}$, N_t is the number of true classifications, and N_{val} is the number of validation samples) of precision [$p = 9.2 \times 10^{-4} (< 0.001)$, $F = 2.6 \times 10^1$, 95% CI], recall [$p = 9.5 \times 10^{-4} (< 0.001)$, $F = 2.6 \times 10^1$, 95% CI], and accuracy [$p = 9.5 \times 10^{-4} (< 0.001)$, $F = 2.6 \times 10^1$, 95% CI] and a reduction in the categorical entropy loss [$p = 1.7 \times 10^{-4} (< 0.001)$, $F = 4.3 \times 10^1$, 95% CI]. The equation of categorical entropy loss is $\ell = - \sum_i y_i \cdot \log \hat{y}_i$ (\hat{y}_i is the i^{th} predicted category possibility, y_i is the corresponding category possibility of the ground truth) and it measures how distinguishable the discrete predictions are between each class. Fig. 9 shows the verification results of the precision and recall in each category, where the error rate of the classifier w/ SWF was significantly lower than that of the classifier w/o SWF. The results indicated that the SWFs improved the capability of identifying the nine ECSs. To accurately verify the models, we implemented fivefold cross-validation to assess the performance of the model according to all the above validation results. Moreover, we evaluated 9000 samples obtained from the validation set in the first fold of the fivefold cross-validation using the trained CNN model w/ SWF. Fig. 10 illustrates the validation results in a confusion matrix containing the prediction and ground truth. Overall, the

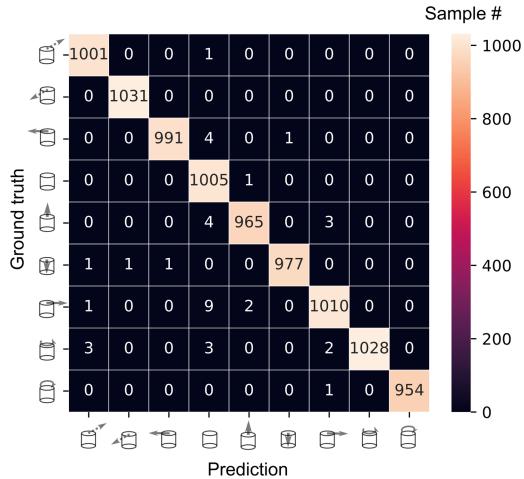


Fig. 10. Confusion matrix of CNN classification results in the validation set of the data (9000 random distributed samples, 99.6% accuracy). The diagonal values indicate the number of correct predictions in each class. The side color bar indicates the number of predicted samples.

results indicate that the CNN model w/ SWF could identify the nine ECSs with 99.6% accuracy. Additionally, we have tested only using the SWFs (Feature #: 19-32) to identify the ECSs offline. The results show an average accuracy of 68.8% with a 1.6% standard deviation (among the fivefold).

B. Applications of NN-ECS for Object Handover and Ball-Hit Recognition

Inspired by how human hands estimate the ECS of a grasped object via tactile feedback and patterns in the somatosensory pathway to conduct motor commands for making handover decisions [2], [6], [26], we incorporated biomimetic tactile sensing capabilities, bioinspired tactile features decoded from GTac, and the neuron network-based ECS (NN-ECS) identification model in GTac-Hand and developed object handover capabilities in it. We also demonstrated recognition capabilities related to the grasped object, i.e., ball-hit, via the NN-ECS model and the GTac-Hand. First, we implemented a control loop for the handover experiments consisting of four phases [see Fig. 11(a)] to demonstrate two daily essential capabilities under two conditioners: 1) handover and 2) ball-hit recognition. As shown in the experiments [see Fig. 11(b)], the handover of an egg was demonstrated by the GTac-Hand using the 4-phases control loop described above. First, the GTac-Hand was mounted on the robotic arm and the robot arm approached the egg. Next, the GTac-Hand started its cycles and entered the grasping phase, when $t = 0$ s. In the grasping phase, the GTac-Hand used a closed-loop grasping controller to grasp the egg, while the thumb flexion was conducted after the index finger flexion (grasping phase ended at $t_1 = 1.2$ s). After initialization for ECS recognition (completion of pre-ECS, $t_2 = 4.5$ s), the ECS recognition ran alongside the task-oriented conditioner to determine the handover decision. Meanwhile, the GTac-Hand ascended to the top ($t_3 = 8.9$ s) and started to descend to hand over the egg to the holder, where the robotic arm was programmed to move in a fixed trajectory (ascend 10 cm and descend 12 cm). When

the handover decision was made ($t_4 = 14.2$ s), the GTac-Hand handed the egg over to the holder. In addition, we also demonstrate that the GTac-Hand hands a bar over to a human. As demonstrated in the handover experiments, our system comprising the biomimetic tactile sensing integrated robotic hand, the GTac-Hand, and the learned NN-ECS model, showed the capabilities of a robot in stably grasping and handing objects over to a receiver.

Additionally, we intend to explore the capabilities of robots in recognition related to grasped objects and commonly seen in daily living, i.e., ball-hit recognition. In the experiments [see Fig. 12(a)], the same 4-phases control loop was used, and conditioner (ii) was applied to recognize the bouncing half of the racket according to the biomimetic tactile sensing feedback from GTac and the learned NN-ECS identification model. In each trial, ten hits were conducted by releasing the shuttlecocks on the left and right. The recognition results and the tactile events detected on the GTac-Hand in the first trial are illustrated in Fig. 12(b). The order to release the ten hits was shuffled before executing each trial. In total, six trials were performed, and the overall recognition accuracy was 96.7% [see Fig. 12(c)].

V. DISCUSSION

Our results showed that the human-like tactile sensing capabilities of GTac using patterning and learning models that mimic human somatosensory systems enable robotic hands to grasp fragile objects and identify their nine essential ECSs, marking the first step of our research work of robots toward utilizing tactile sensing to learn human skills. Moreover, the SWFs that were decoded based on mimicking the human somatosensory pathway as an input to the learned classifier could significantly reduce the error rates in identifying the ECSs of grasped objects along with raw GTac feedback. The identified ECSs could be used to perform challenging tasks, such as ball-hit recognition, and object handovers. The experimental task settings in this study were inspired by situations in our daily living. Ball-hit recognition skills are important for athletes to predict the ball's bouncing trajectory based on the tactile feedback in their hands, particularly when the high-speed flying ball cannot be clearly seen. Object handover is a daily activity; examples of it include handovers to others, tables, holders, and hooks that utilizes the leveraged tactile sensing feedback and estimated ECS of the grasped objects.

To deal with magnetic disturbances from the adjacent magnets, we have implemented a filtering algorithm, called sensing after contact (in Algorithm 1), which discarded the signals from the SA-II layer until the contact condition was triggered on each finger section. However, this solution caused a tactile onset delay in the SA-II signals (see Fig. 3). This delay would only occur at the beginning shortly after the contact, which could be an issue for some robotic systems. This issue could be solved by shielding the Hall sensor from the external magnetic field as an important future work.

GTac enables the robotic hand to recognize the nine dominant ECSs of a grasped object (a cylinder bar in this study). Although many domestic objects have a similar bar-shaped handle that

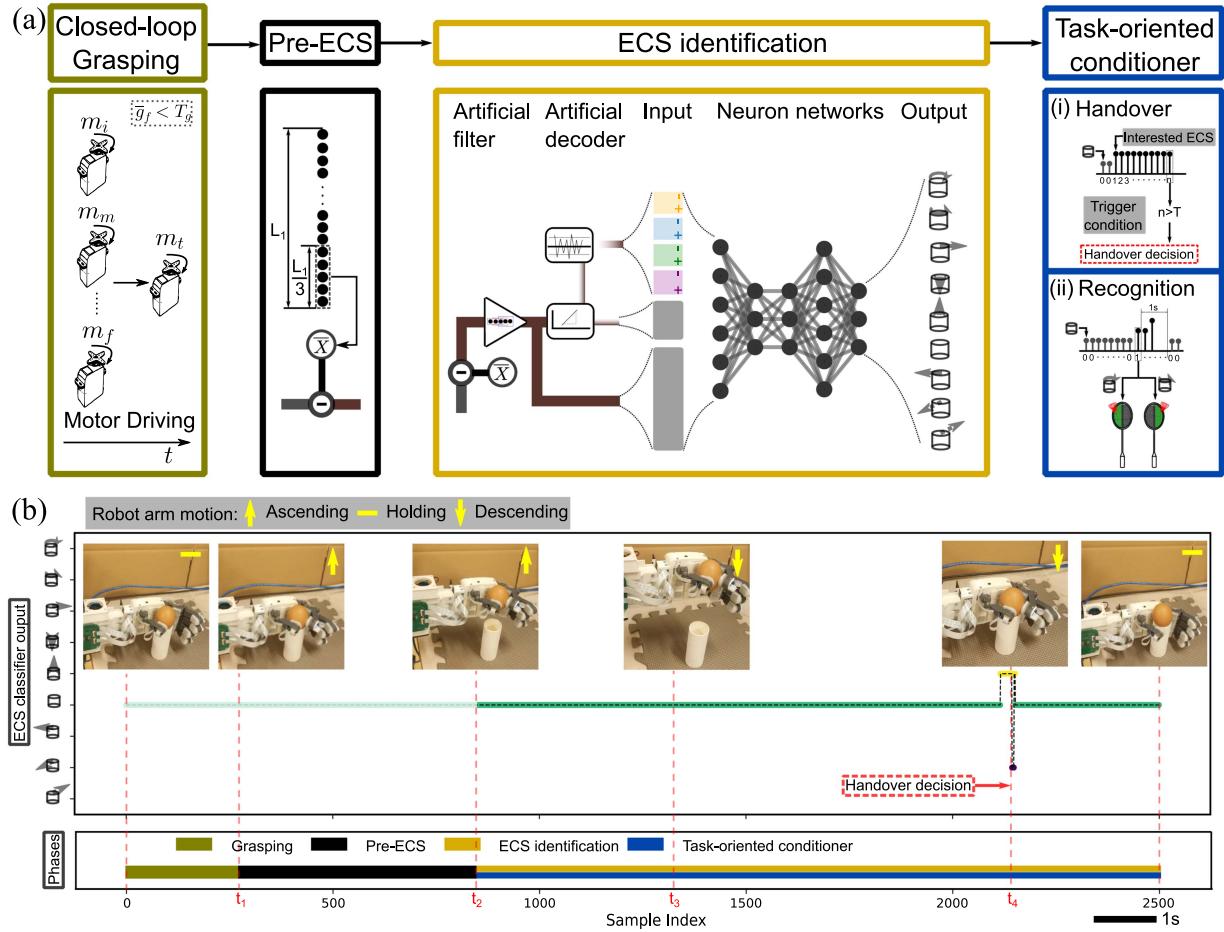


Fig. 11. (a) Illustration of the four phases in the control loop of the GTac-Hand. (b) Experiment of grasping and handover of an egg using the learned NN-ECS model and the handover controller.

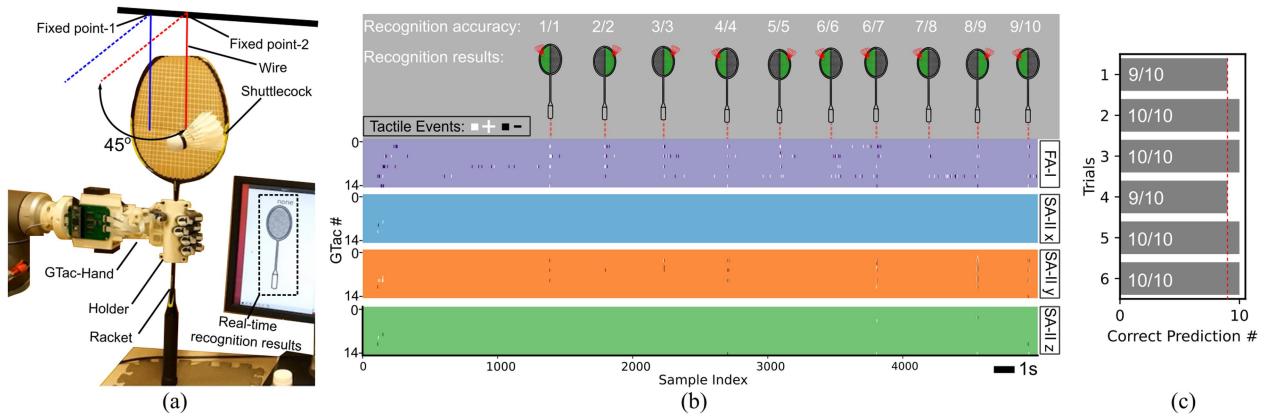


Fig. 12. (a) Experimental setup of applying ECS identification capabilities to perform ball-hit recognition. (b) Recognition results and decoded tactile events representations of one trial experiment. (c) Recognition accuracy of all the six trials of ball-hit recognition experiments conducted (96.7% accuracy over 60 trials in total).

can be grasped, the shapes and sizes of these objects are complicated, and it is not realistic to cover all aspects of them in our experiments. We only considered the bar-shaped cylinder in our ECS recognition experiments. If the object shape or the grasping gesture is changed, the neuron network classifier needs to be retrained. In our experiments of ECSs identification, only raw data were used as input implying there was no conversion to

actual forces since the features remain in the raw tactile signals and the linearity of the raw data slightly varies among sensors depending on the sensor system manufacturing tolerance, installation deviation, etc. (sensor linearity analysis in [24]). For future work, this conversion can be addressed by additional calibration procedures for applications that require related measurements and it would be interesting to investigate how the nonlinearity

of sensors data affects the ECS or other identification tasks to determine the required specification of the sensor. Regarding that the SWFs are derived from GTac outputs according to their spatial and temporal features, it would be important as future work to clarify how much each feature/each type of features contributes to improving performance in some cases more complicated, e.g., requiring to leverage/focus on either type of features. The grasping capability of the robotic hand is limited owing to its cable-driven and underactuated design and the grasping gestures implemented in ECSs recognition were limited; this issue can be solved by designing a fully actuated robotic hand. These thresholds implemented in this study were established by trials-and-errors to remain tactile feedback in a range that keeps the data originality by observing the raw signals, decoder output, and tactile events. The sampling rate of the GTac-Hand is limited to 150 Hz and is constrained by the microcontroller computational power, which may not be ideal for high-speed response applications. This limitation can be addressed by using a customized field-programmable gate array (FPGA) [27] board to reduce unnecessary computation and increase the readout rate.

VI. CONCLUSION

Tactile sensing feedback is critical important for many manipulations tasks using robots, such as tool usage. In this study, taking inspiration from the human somatosensory system, patterning and learning model with the proposed tactile features, SWFs, facilitated GTac-Hand to perform real-time ECSs recognition and several human skills, such as grasping and handover of delicate objects. The proposed GTac-Hand can perceive normal and shear force sensing at the fingers and the palm. Given the cable-driven mechanism and tactile sensing integration, multiple grasping gestures can be performed while the tactile feedback is used for closed-loop grasping. Our results illustrated its potential in intelligent humanoids and human–robot interaction.

ACKNOWLEDGMENT

The authors would like to thank H. Guo, S. Bhattacharya, W. Zhang, and J. Liang for assistance with the execution of the experiments and also D. Carmona for proofreading the manuscript.

REFERENCES

- [1] D. Ma, S. Dong, and A. Rodriguez, “Extrinsic contact sensing with relative-motion tracking from distributed tactile measurements,” in *Proc. IEEE Int. Conf. Robot. Autom.*, 2021, pp. 11262–11268.
- [2] Q. Li, O. Kroemer, Z. Su, F. F. Veiga, M. Kaboli, and H. J. Ritter, “A review of tactile information: Perception and action through touch,” *IEEE Trans. Robot.*, vol. 36, no. 6, pp. 1619–1634, Dec. 2020.
- [3] R. S. Johansson and J. R. Flanagan, “Coding and use of tactile signals from the fingertips in object manipulation tasks,” *Nature Rev. Neurosci.*, vol. 10, no. 5, pp. 345–359, May 2009. [Online]. Available: <http://www.nature.com/articles/nrn2621>
- [4] J. J. Hopfield, “Pattern recognition computation using action potential timing for stimulus representation,” *Nature*, vol. 376, no. 6535, pp. 33–36, Jul. 1995. [Online]. Available: <http://www.nature.com/articles/376033a0>
- [5] T. Masquelier, R. Guyonneau, and S. J. Thorpe, “Spike timing dependent plasticity finds the start of repeating patterns in continuous spike trains,” *PLoS One*, vol. 3, no. 1, Art. no. e1377, Jun. 2008. [Online]. Available: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0001377>
- [6] V. Ortenzi, A. Cosgun, T. Pardi, W. P. Chan, E. Croft, and D. Kulic, “Object handovers: A review for robotics,” *IEEE Trans. Robot.*, vol. 37, no. 6, pp. 1855–1873, Dec. 2021.
- [7] Z. Kappassov, J.-A. Corrales, and V. Perdereau, “Tactile sensing in dexterous robot hands—Review,” *Robot. Auton. Syst.*, vol. 74, pp. 195–220, Dec. 2015. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0921889015001621>
- [8] S. Luo, J. Bimbo, R. Dahiya, and H. Liu, “Robotic tactile perception of object properties: A review,” *Mechatronics*, vol. 48, pp. 54–67, Nov. 2017. [Online]. Available: <http://arxiv.org/abs/1711.03810>
- [9] R. S. Dahiya, G. Metta, M. Valle, and G. Sandini, “Tactile sensing—from humans to humanoids,” *IEEE Trans. Robot.*, vol. 26, no. 1, pp. 1–20, Feb. 2010.
- [10] D. Silvera-Tawil, D. Rye, and M. Velonaki, “Artificial skin and tactile sensing for socially interactive robots: A review,” *Robot. Auton. Syst.*, vol. 63, pp. 230–243, Jan. 2015. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0921889014001833>
- [11] B. Shih et al., “Electronic skins and machine learning for intelligent soft robots,” *Sci. Robot.*, vol. 5, no. 41, Art. no. eaaz9239, Apr. 2015. [Online]. Available: <https://robotics.sciencemag.org/lookup/doi/10.1126/scirobotics.aaz9239>
- [12] N. F. Lepora et al., “Towards integrated tactile sensorimotor control in anthropomorphic soft robotic hands,” in *Proc. IEEE Int. Conf. Robot. Autom.*, 2021, pp. 1622–1628.
- [13] M. Lambeta et al., “DIGIT: A novel design for a low-cost compact high-resolution tactile sensor with application to in-hand manipulation,” *IEEE Robot. Autom. Lett.*, vol. 5, no. 3, pp. 3838–3845, Jul. 2020.
- [14] P. Piacenza, K. Behrman, B. Schifferer, I. Kymmissis, and M. Ciocarlie, “A sensorized multicurved robot finger with data-driven touch sensing via overlapping light signals,” *IEEE/ASME Trans. Mechatron.*, vol. 25, no. 5, pp. 2416–2427, Oct. 2020.
- [15] G. Büscher, M. Meier, G. Walck, R. Haschke, and H. J. Ritter, “Augmenting curved robot surfaces with soft tactile skin,” in *Proc. IEEE Int. Conf. Intell. Robots Syst.*, 2015, pp. 1514–1519.
- [16] U. Kim et al., “Integrated linkage-driven dexterous anthropomorphic robotic hand,” *Nature Commun.*, vol. 12, no. 1, Dec. 2021, Art. no. 7177. [Online]. Available: <https://www.nature.com/articles/s41467-021-27261-0>
- [17] B. Jamil, G. Yoo, Y. Choi, and H. Rodrigue, “Proprioceptive soft pneumatic gripper for extreme environments using hybrid optical fibers,” *IEEE Robot. Autom. Lett.*, vol. 6, no. 4, pp. 8694–8701, Oct. 2021.
- [18] Y. Xia et al., “Beta mixture model for the uncertainties in robotic haptic object identification,” *IEEE/ASME Trans. Mechatron.*, vol. 27, no. 4, pp. 1955–1963, Aug. 2022.
- [19] M. A. Lee et al., “Making sense of vision and touch: Learning multimodal representations for contact-rich tasks,” *IEEE Trans. Robot.*, vol. 36, no. 3, pp. 582–596, Jun. 2020.
- [20] B. Heyneman and M. R. Cutkosky, “Slip classification for dynamic tactile array sensors,” *Int. J. Robot. Res.*, vol. 35, no. 4, pp. 404–421, Apr. 2016, doi: [10.1177/0278364914564703](https://doi.org/10.1177/0278364914564703).
- [21] R. Calandra et al., “More than a feeling: Learning to grasp and regrasp using vision and touch,” *IEEE Robot. Autom. Lett.*, vol. 3, no. 4, pp. 3300–3307, Oct. 2018.
- [22] Y. Chebotar, K. Hausman, Z. Su, G. S. Sukhatme, and S. Schaal, “Self-supervised regrasping using spatio-temporal tactile features and reinforcement learning,” in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2016, pp. 1960–1966.
- [23] J. M. Romano, K. Hsiao, G. Niemeyer, S. Chitta, and K. J. Kuchenbecker, “Human-inspired robotic grasp control with tactile sensing,” *IEEE Trans. Robot.*, vol. 27, no. 6, pp. 1067–1079, Jun. 2020.
- [24] Z. Lu, X. Gao, and H. Yu, “GTac: A biomimetic tactile sensor with skin-like heterogeneous force feedback for robots,” *IEEE Sensors J.*, vol. 22, no. 14, pp. 14491–14500, Jul. 2022.
- [25] F. Ficuciello, “Synergy-based control of underactuated anthropomorphic hands,” *IEEE Trans. Ind. Inf.*, vol. 15, no. 2, pp. 1144–1152, Feb. 2019.
- [26] G. Rizzolatti and L. Craighero, “The mirror-neuron system,” *Annu. Rev. Neurosci.*, vol. 27, no. 1, pp. 169–192, Jul. 2004. [Online]. Available: <http://www.annualreviews.org/doi/10.1146/annurev.neuro.27.070203.144230>
- [27] K. Park, H. Lee, K. J. Kuchenbecker, and J. Kim, “Adaptive optimal measurement algorithm for ERT-based large-area tactile sensors,” *IEEE/ASME Trans. Mechatron.*, vol. 27, no. 1, pp. 304–314, Feb. 2022.



Zeyu Lu received the B.S. degree in mechanical engineering from Southern University of Science and Technology, Shenzhen, Guangdong, China, in 2019. He is currently working toward the Ph.D. degree in biomedical engineering with the National University of Singapore, Singapore.

His research interests include bioinspired robots, tactile sensors, multifingered hand/gripper, grasping and manipulation, and robotic learning.



Haoyong Yu (Senior Member, IEEE) received the B.S. and M.S. degrees in mechanical engineering from Shanghai Jiao Tong University, Shanghai, China, in 1988 and 1991, respectively. He received the Ph.D. degree in mechanical engineering from Massachusetts Institute of Technology, Cambridge, MA, USA, in 2002.

He is currently an Associate Professor with the Department of Biomedical Engineering, National University of Singapore. His areas of research include medical robotics, rehabilitation engineering and assistive technologies, system dynamics and control.