RESEARCH ARTICLE

ADVANCED
INTELLIGENT
SYSTEMS
Open Access

www.advintellsyst.com

Unlocking Dynamic Subtle Stimuli Tactile Perception: A Deep Learning-Enhanced Super-Resolution Tactile Sensor Array with Rapid Response

Shuyao Zhou, Depeng Kong, Mengke Wang, Baocheng Wang, Yuyao Lu, Honghao Lyu, Zhangli Lu, Yong Tao, Kaichen Xu, and Geng Yang*

Human skin's touch perception, mediated by mechanoreceptors, can perceive tactile stimuli with a spatial resolution higher than the average spacing between mechanoreceptors, known as super-resolution. This characteristic enables its sensitivity to both stimulus location and velocity. However, existing robotic tactile sensors lag behind human tactile perception, failing to achieve high spatial resolution and rapid response simultaneously. This significantly hinders robots from executing accurate, time-sensitive interaction tasks, particularly during dynamic slight-contact events. Here, a 130 μm-thick flexible tactile sensor array is designed, with spatial resolution enhanced by a tailored deep learning model, multistage attention-based adaptive spatial–temporal graph convolutional networks (MS-AASTGCN), simultaneously achieving a dynamic response of ≈30 ms and a super-resolution factor of 75.19. The tactile sensor array, based on single-electrode triboelectric nanogenerators, can detect dynamic subtle stimuli. It features a bio-inspired topological structure, facilitating super-resolution performance while offering a large perception area with low sensor (taxel) density. Additionally, the MS-AASTGCN enhances positioning resolution by extracting features from sensing data and revealing the hidden relationships among taxels. This research provides new insights into haptic perception systems, enabling the execution of precise, time-sensitive interaction tasks, such as tracking a bouncing table tennis ball.

1. Introduction

The past decades have seen significant development in robot manipulation,[1] in which tactile sensing is crucial for robot's interaction with their surroundings. By mimicking the functionalities of human skin, tactile sensors provide a rich and diverse set of information about the environment in the form of contact-based signals.[2] Accurate and rapid response to external stimuli is preferred for tactile sensors to provide tactile feedback. However, simultaneously achieving high spatial resolution and rapid response is challenging due to the limitations brought by intricate manufacturing processes and the sensing mechanisms of these tactile sensors. In contrast, human skin excels at achieving both high spatial resolution and rapid response. Human skin's touch perception, mainly mediated by four types of mechanoreceptors,[3] can perceive tactile stimuli with a spatial resolution higher than the average spacing between mechanoreceptors (known as super-resolution)[4] and is highly sensitive to both stimulus location and velocity.[3] If a robotic skin is integrated with these two unique features, it will bestow robots with more opportunities to interact with the environment like human skin by providing richer tactile information.[5] Hence, designing an intelligent haptic perception system that mimics human sensory performance would lead to faster and more responsive interactions, especially in dynamic events, which is meaningful for the advancement of robot capabilities.

Many artificial haptic perception systems have been introduced since material science, electronic engineering, and

---

S. Zhou, D. Kong, M. Wang, B. Wang, Y. Lu, H. Lyu, Z. Lu, K. Xu, G. Yang
State Key Laboratory of Fluid Power and Mechatronic Systems
School of Mechanical Engineering
Zhejiang University
Hangzhou 310030, China
E-mail: yanggeng@zju.edu.cn

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/aisy.202400913.

H. Lyu, K. Xu, G. Yang
Zhejiang Key Laboratory of Intelligent Robot for Operation and Maintenance
Hangzhou 310030, China

Y. Tao
School of Mechanical Engineering & Automation
Beihang University
Beijing 100191, China

Y. Tao
Aero-Engine Research Institute
Beihang University
Beijing 102206, China

ADVANCED
SCIENCE NEWS

www.advancedsciencenews.com

ADVANCED
INTELLIGENT
SYSTEMS

www.advintellsyst.com

artificial intelligence have advanced significantly.[6–8] Tactile sensors, which actively convey mechanical stimuli to a wide range of electrical signals, are essential elements of intelligent haptic perception systems.[6] Various artificial tactile sensors have been proposed based on sensing mechanisms ranging from piezoresistance[9–12] and triboelectric[13–19] to magnetism[20] and optics[21,22] mechanisms. Some of them focus on enhancing spatial resolution in tactile perception, even achieving super-resolution perception. Machine learning technologies, with their prominent ability for big data analysis and fast processing speed,[23] are instrumental in mapping raw sensory information to relevant functional abstractions for these sensors,[9,15,20,24] thereby enabling advances in spatial resolution. A detailed comparison of these artificial tactile perception systems is shown in **Table 1**. Among these, the piezoresistive tactile sensor arrays[9] and the latest optical-based sensors[21,22] have achieved super-resolution comparable to human skin, but the bulky structure[21,22] or limited dynamic responses[12,21] restrict their applications in interaction tasks that require compact sensor structure and rapid dynamic response. In recent years, researchers have increasingly integrated bionic concepts into the design of tactile design, aiming to enhance tactile perception performance by effectively mimicking skin structures or mechanoreceptors. The soft magnetic film-based tactile sensors, featuring a sandwich structure[20] that mimics the functionalized multilayer structure of human skin, have achieved a 60-fold super-resolved accuracy enhanced by multilayer perceptron (MLP) models, but this super-resolution is solely focused on contact under pressure, making it impractical for subtle stimuli. Although a biomimetic elastomeric robot skin[10] has achieved the tactile perception of static and dynamic stimuli by mimicking different types of mechanoreceptors with electrodes and microphones and utilizing MLP models, the performance of super-resolution for dynamic perception has remained limited.

It is worth noting that current super-resolution haptic perception systems primarily concentrate on static tactile stimuli, as the tactile sensors used are based on sustained pressure signals and the induced deformation of the sensing medium.[9–12,20–22,24] In addition, the signal features generated by dynamic subtle stimuli are more challenging to capture and effectively utilize due to their typically brief duration and limited prominence. These limitations hinder effective and rapid responses to external, dynamic subtle stimuli, such as stroking and light touch. Fortunately, the triboelectric nanogenerator (TENG)-based tactile sensors,[13–18] which exhibit high sensitivity to dynamic subtle stimuli with rapid responses, are widely employed in the dynamic tactile perception domain in the form of aligned rectangular arrays. Nonetheless, this type of tactile array cannot offer spatially continuous perception due to the physical isolation between adjacent taxels, resulting in generally low spatial resolution. Most of these arrays strengthened resolution at the physical level, for example, by minimizing taxel size through microfabrication technology[13–15] for higher spatial resolution and by fabricating microstructures that employ a reactive ion etching process[17] for greater touch sensitivity. However, these solutions are limited by complicated manufacturing and dense wiring, which are prone to failures. Additionally, machine learning techniques have also been applied to TENG-based tactile sensor arrays,[15,18] primarily to improve the precision of mapping contact positions to real taxels[15] or classify dynamic stimuli,[15,18] rather than genuinely achieving super-resolution for stimuli contact positions.

To tackle this challenge, this research proposes a flexible tactile sensor array with a thickness of 130 μm (**Figure 1**) that incorporates both super-resolution and rapid response capabilities, mimicking the location-sensing mechanism of human skin. These features of this tactile array are attributed to the effective combination of the single-electrode TENGs (SETENG)-based

Table 1. Comparison of state-of-the-art artificial tactile perception systems.

Tactile perception mode	Sensing principle	Spatial resolution enhancement methods	Super-resolution	Spatial resolution[a]	Response time of system	Feature
Static tactile perception	Piezoresistive[9]	Physics-based modeling + data-driven learning model	Yes	From ≈100 to ≈3.5 mm	–	Scalable
	Magnetic[20]	Bioinspired structure of tactile sensors + MLP model + table	Yes	From 6 to 0.1 mm	–	High resolution but complex
	Optical (fiber Bragg gratings based)[21]	Bioinspired tactile sensor mechanism + MLP model	Yes	3.2 mm[b]	≈500 ms	Simple structure but bulky
	Optical (Taptic)[22]	Bioinspired structure of tactile sensors + Bayesian perception	Yes	From 2.4 to 0.20 mm	–	Bulky, rigid but high resolution
Static and dynamic tactile perception	Piezoresistive and acoustics[10]	Bioinspired multilayer structure + MLP model	Yes	From 40 to 4.2 mm (static); From 55 to 6.6 mm (dynamic)	–	Soft, complex, and bulky
Dynamic tactile perception	TENG-based[15]	Dense sensor array + the miniaturized physical taxel's size	No	6 dpi (4.23 mm)	≈160 ms	Flexible, highly sensitive but complex
	SETENG-based (This work)	Bioinspired optimized topological structure of sensor array + graph neural network model	Yes	From 15 to 1.50 mm	≈30 ms	Flexible, highly sensitive, and simple

[a]The improvement from physical resolution to super-resolution for the super-resolution tactile perception systems; the spatial resolution for the nonsuper-resolution tactile perception systems; [b]The physical resolution was not reported in the article.
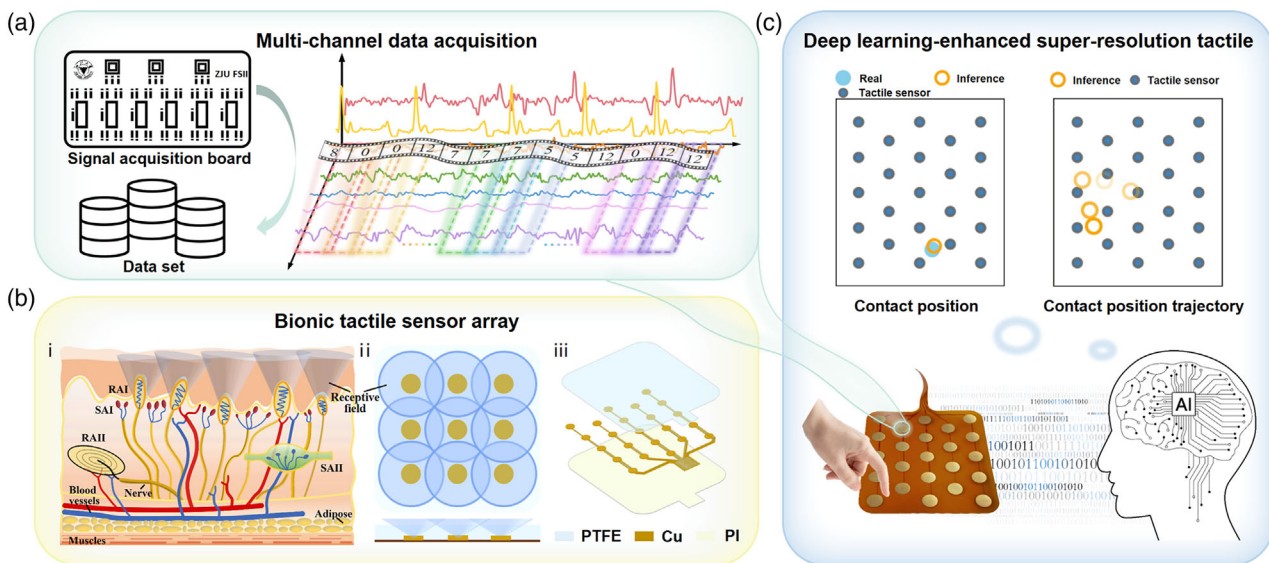
Figure 1. Schematic of the bionic deep learning-enhanced super-resolution tactile perception system. a) Schematic of multichannel data acquisition. b) The bionic design of the tactile sensor array. (i) Illustration of the human skin structure. (ii) Illustration of the overlapping receptive field structures in a tactile sensor array. (iii) Explosive view of the bionic tactile sensor array. c) The prospective applications of the deep learning-enhanced super-resolution tactile array in interaction tasks.

tactile sensor array with the optimized topological structure and a super-resolution algorithm customized for the hardware of the array, called multistage attention-based adaptive spatial–temporal graph convolutional networks (MS-AASTGCN). The sensing mechanism of SETENG endows it with an inherent advantage as an artificial tactile sensor for rapidly detecting dynamic subtle stimuli. The optimized topological structure has a skin-inspired overlapping receptive field structure that effectively combines the receptive field of individual tactile sensors, facilitating super-resolution performance. Additionally, the MS-AASTGCN is also used to enhance the positioning super-resolution by inferring multichannel signals and revealing the hidden relationships between individual taxels. Consequently, our proposed tactile sensor array, along with the MS-AASTGCN, simultaneously achieved a rapid dynamic response of ≈30 ms and a super-resolution factor of 75.19 (1.50 mm root mean square error [RMSE] was achieved for localization inference). Furthermore, we demonstrated that our tactile sensor array possesses the capability for rapid response and precise localization of dynamic subtle stimuli from the external environment by tracking the bouncing trajectory of a table tennis ball, proving that this super-resolution sensor array holds great promise for complex and time-sensitive interaction tasks in the future.

2. Results

2.1. SETENG-Based Bionic Tactile Sensor Array Design

2.1.1. Electrical Output Characteristics of SETENG

Four types of mechanoreceptors, distinguished by their functional properties, exist in the human skin. According to their sensitivity to static and dynamic events, they are categorized into slowly adapting (SA) and rapidly adapting (RA), of which the SA units respond to sustained indentations while the RA units mainly encode stimuli transients. Besides, they are classified into type I (surface-located) and type II (deeply located) depending on the size and boundary conditions of their receptive fields, leading to four types of mechanoreceptors,[25] that is, RA-I, RA-II, SA-I, and SA-II (Figure 1b i).

Among these, RA-I responses are related to smaller receptive fields with distinct borders and dynamic low-frequency vibrations (such as tapping and flutter in the range of 1–10 Hz),[26] RA-II responses are related to larger receptive fields with diffuse borders and dynamic high-frequency vibrations (such as machine vibrations in the range of 200–300 Hz).[5] In contrast, SA-I responses are related to smaller receptive fields with distinct borders and sustained indentations on the skin (i.e., perception of superficial light pressure), SA-II responses are related to larger receptive fields with diffuse borders and perception of skin stretch (i.e., perception of deep pressure).[3] Based on the characteristics of these four types of mechanoreceptors and the requirements of interaction tasks mentioned above, RA-I meets our needs for artificial tactile sensors. Additionally, the overlapping receptive fields of RA-I promote more refined and acute tactile perception and are well suited for analysis of spatial stimulus at the skin surface using a minimal number of mechanoreceptors[27] (Figure 1b i).

Inspired by the structure and working mechanism of RA-I mechanoreceptors in human skin, we designed the sparse SETENG-based tactile sensor array, where each SETENG-based tactile sensor mimics an individual RA-I mechanoreceptor in the skin, aiming to achieve high spatial resolution and rapid response in tactile perception. Figure 1b ii illustrates the overlapping receptive field structure in a tactile sensor array. Figure 1b iii shows a schematic of this flexible SETENG sensor array. A square piece of polyimide (PI) film was utilized as the flexible dielectric substrate; the Cu electrode layer consisted of two parts:

The electrode array with a predesigned special topology (the topological structure will be introduced in Section 2.1.2) on the top side served as the charge-sensing element and the circuit configuration electrodes on the bottom side that connect to external measuring equipment; the electrode array and the configuration electrodes interconnect with each other via through-holes (Figure S1, Supporting Information). We adopted commercial flexible printed circuit board (FPCB) processing methods to fabricate the substrate and electrode layers. Then, a polytetrafluoroethylene (PTFE) film tape, cut to the same size as the substrate layer, was covered onto the top; this film, serving as the electrification layer, generates triboelectric charges when being touched.

The flexible bionic SETENG-based tactile sensor array comprises 23 SETENG-based tactile sensors, each with a radius of 5 mm (**Figure 2**a), functioning as the original physical taxels (the smallest sensing unit in the tactile sensor array, i.e., each tactile sensor, that can detect tactile information, named a taxel, is responsible for sensing interactions near its location), whose working principle can be explained by the coupling effect of contact electrification and electrostatic induction (Figure 2b). The electrode is connected in SETENG mode with resistor loads. Due to the differing electron affinities of PTFE and the finger (bare finger or finger wearing a nitrile butadiene rubber (NBR) glove), the finger surface will accumulate positive charges, while the PTFE surface will accumulate negative charges when they are in close contact with each other according to contact electrification. When the two charged surfaces are separated, free electrons are impelled from the Cu electrode to the ground to
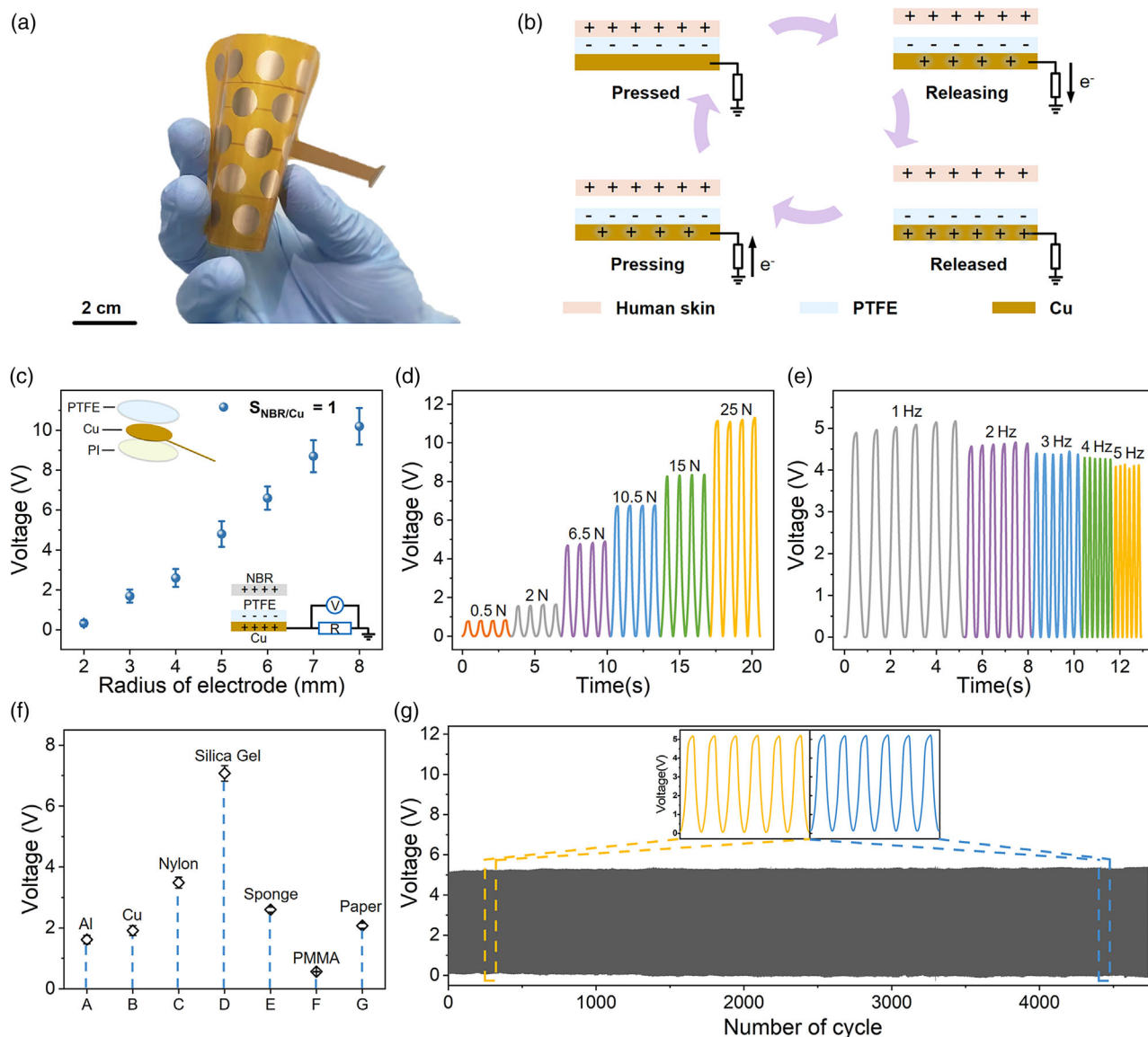

Figure 2. Characterizations of the SETENG-based tactile sensor. a) Photograph of a fabricated SETENG-based tactile sensor array with high flexibility. b) Working mechanism of the sensor. The output voltage of the SETENG-based tactile sensor under stimuli of c) different contact areas (the size ratio of Cu electrode and NBR foil was maintained at 1), d) contact forces, and e) frequencies. f) The output voltage induced by various contact materials including Al, Cu, Nylon, silica gel, sponge, PMMA, and paper. g) The wear resisting of the proposed triboelectric sensor under over 4500 cycles of friction.

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
INTELLIGENT
SYSTEMS**
Open Access

www.advintellsyst.com

balance the electric field caused by the negative charges on the PTFE film, inducing a current flow and an output voltage in this electrostatic induction process. Subsequently, when the finger is approaching and eventually contacts the electrode area, electrons will flow back to the Cu electrode due to the potential difference, resulting in the reversal of the electrical signals. This working mechanism effectively converts the dynamic touch activities occurring on this tactile sensor into electrical signals. The device structure and the schematic of the measuring circuit are shown in the inset of Figure 2c.

The basic characterization of the tactile sensor was conducted with respect to key parameters that influenced the triboelectric voltage output, such as contact areas, forces, frequencies, and materials. For contact area, the output voltage increased with increasing contact area between the circular Cu electrode and the NBR sample, with the radius ranging from 2 to 8 mm (Figure 2c). As the contact force increased from 0.5 to 25 N, the output voltage also increased in gradient (Figure 2d). Notably, the output voltage changed negligibly with increasing contact frequencies (Figure 2e), which proved that the output voltage of the proposed SETENG was independent of contact frequencies, contributing to stable responses under various contact speeds. Additionally, the various output voltages of the SETENG-based tactile sensor, in response to diverse ranges of contact objects made from different materials with varied electronegativity (Figure 2f), indicated that the proposed SETENG was suitable for application in complex surroundings. Moreover, to validate the durability of this sensor, over 4500 cycles of repetitive contact tests at a frequency of 1 Hz were conducted; the output voltage exhibited no apparent fluctuation during this prolonged test, reflecting the stability and signal repeatability of SETENG. The excellent mechanical and electrical properties of SETENG will be beneficial for rapid dynamic tactile perception.

2.1.2. Topological Structure of the Tactile Sensor Array

In contrast to the typical aligned rectangular arrangement observed in the majority of sensor arrays, we designed the topological structure of the tactile sensor array by rationally combining the receptive field size of individual tactile sensors, based on the empirical knowledge that overlapping multiple taxels' receptive fields trigger super-resolution behavior.[9] By attempting to transfer overlapping receptive field structures from the biological domain to the design of artificial tactile sensor arrays' topological structure, we aim to achieve large-scale perception with fewer tactile sensors and lower wiring density, that is, higher-quality tactile perception at a lower cost.

The design of the sensor array's topological structure was abstracted into a mathematical problem (**Figure 3**a). Three tactile sensors form a basic optimization unit, with the receptive field radius of each tactile sensor set as r and the distances between the centers of the three tactile sensors as d_1, d_2, d_3. Find the values of $[d_1, d_2, d_3]$ that contribute to the largest union of the receptive fields of three tactile sensors, under the premise of $d_1, d_2, d_3 \leq r$. Utilizing the particle swarm optimization algorithm combined with the Monte Carlo algorithm for area calculation, the optimal arrangement of $[d_1, d_2, d_3]$ is determined to be $d_1 = d_2 = d_3 = r$, forming an equilateral triangular arrangement (the optimization

process is shown in Figure 3b). Figure 3c illustrates the receptive field test of a SETENG sensing unit and its output voltage when the stimuli are applied at varying distances from the sensor (as shown in the inset of Figure 3c), demonstrating a notable disparity in output voltage within a distance of 15 mm, which can be considered the receptive field radius of the tactile sensor. Therefore, 15 mm was identified as the sensor spacing, with the overall size of the tactile sensor array being $80 \times 72 \, mm^2$. The topological structure of the tactile sensor array is shown in Figure 3d. To be noted, each tactile sensor within this tactile sensor array was fabricated using the same materials and processes, ensuring that they all possess almost identical electrical characteristics. Therefore, for the tactile sensor array, the relative differences in output voltage values among the 23 tactile sensors are attributed to the contact position, which helps MS-AASTGCN learn the patterns of voltage variations among the 23 tactile sensors under different contact positions. The entire tactile sensor array consists of multiple basic optimization units, ensuring optimal utilization of each tactile sensor's receptive field to promote super-resolution. Furthermore, this tactile sensor array's topological structure enables the activation of at least three sensors in response to an external stimulus, resulting in the generation of richer tactile signals and a finer, sharper tactile perception.

2.2. Deep Learning-Enhanced Super-Resolution Tactile Perception

Having identified a reasonable topological structure for the tactile sensor array, its spatial resolution can be effectively enhanced by a deep learning model that generates high-resolution virtual taxels. By synthesizing signals from multiple sensing units, the deep learning model estimates external contacts at a high spatial resolution, surpassing the tactile sensor array's physical resolution. As illustrated in Figure 3e, when an external stimulus approaches the tactile sensor array and finally contacts its surface, a basic optimization unit (i.e., the three sensors closest to the stimulus contact position) will be activated. The acquisition circuit then inputs the obtained 23-channel voltage signals into the MS-AASTGCN, which outputs the inferred coordinates of the contact position for display on the graphical user interface (GUI). In this super-resolution tactile perception process, the overlapping receptive fields, which fully utilize information from individual taxels, and the MS-AASTGCN, which correlates voltage variations with contact positions, both contribute significantly to the spatial resolution enhancement.

The optimized topological structure of the tactile sensor array establishes the distance-based physical connectivity between individual taxels, that is, the original adjacency matrix that will be mentioned in the following paragraphs. The localization of external dynamic stimulation within the tactile sensor array relies on each taxel, where the location of external stimulation corresponds to the output voltage signals of these sensors (Figure S2, Supporting Information). Additionally, the working mechanism of the SETENG-based tactile sensor reveals that the generation process of the tactile sensor's output voltage aligns with the approach and retreat processes of external stimuli. By combining these two phenomena, this issue can be defined as a spatial-temporal problem.
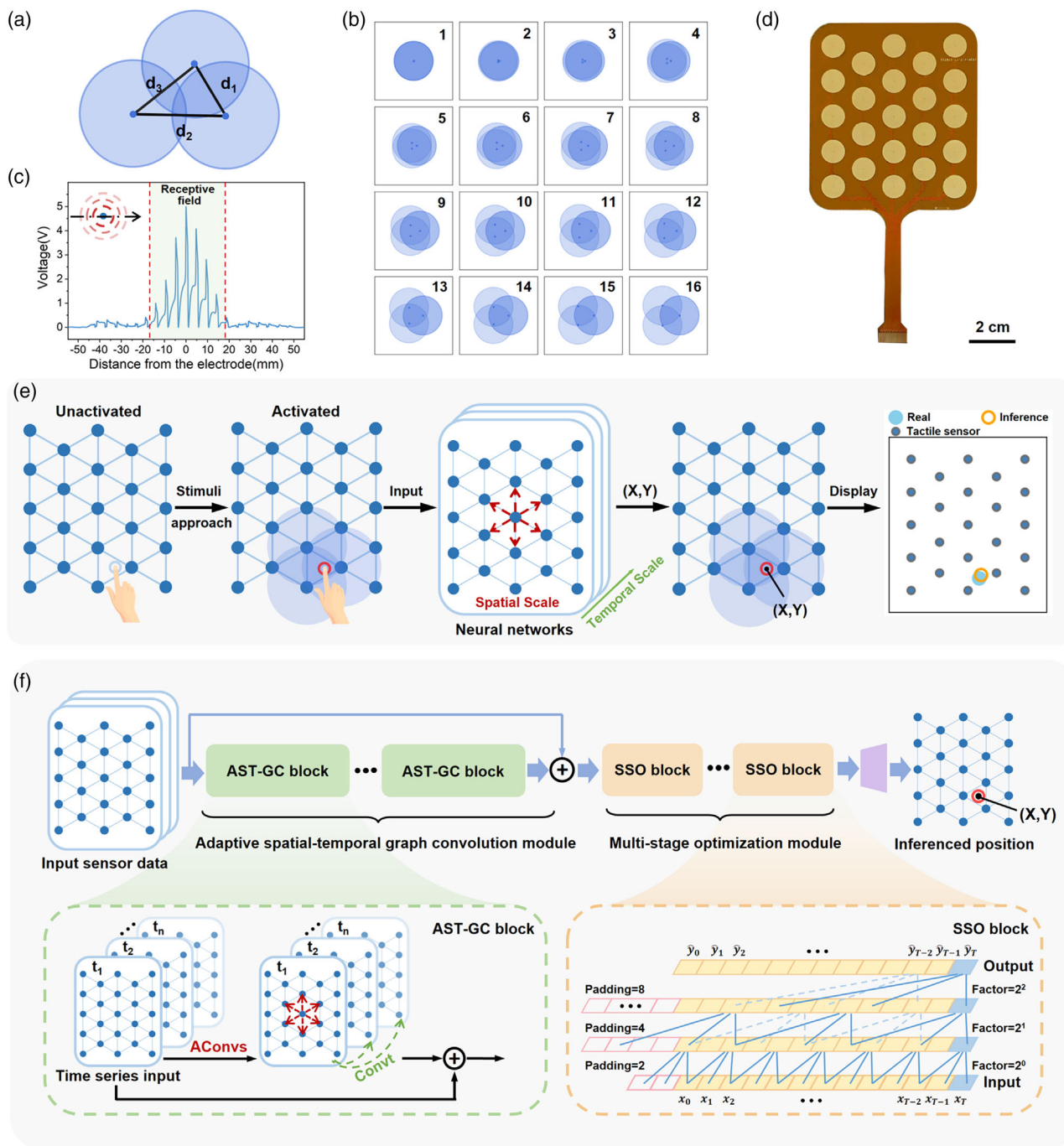
ADVANCED
SCIENCE NEWS
www.advancedsciencenews.com

ADVANCED
INTELLIGENT
SYSTEMS
Open Access
www.advintellsyst.com

Figure 3. Tactile super-resolution method. a) Abstract the basic optimization unit layout optimization problem into a mathematical problem. Given the receptive field of each sensor is r and constraints $d_1, d_2, d_3 \leq r$, find the values of $[d_1, d_2, d_3]$ that maximize the total area of their union. b) The optimization process for the layout of a basic unit comprising three tactile sensors. The optimal layout is the equilateral triangular arrangement. c) Test of the receptive field of a SETENG sensing unit. d) The topological structure of the SETENG-based tactile sensor array. The overall size of the tactile sensor array is 80×72 mm². e) The overall process of super-resolution tactile perception. f) Illustration of the overall architecture of MS-AASTGCN, which consists of an AST-GC module and a multistage optimization module.

The key to solving this spatial-temporal problem lies in addressing inconsistencies (which will be discussed in Section 2.3) among each tactile sensor in the sensor array, as well as in exploring spatial-temporal relationships and aggregating spatial-temporal information among taxels. Due to the outstanding capability of graph neural networks in aggregating information from adjacent nodes, we introduce a graph structure[28] into the inference of external stimulus location. The adaptive

spatial-temporal graph convolution (AST-GC) module utilizes an adaptive adjacency matrix to perform convolution operations in both spatial and temporal dimensions, effectively extracting spatial-temporal features from the original input sensing signals. The unique structure of the adaptive adjacency matrix significantly enhances the robustness of the model (which will be explained in detail), establishing it as the backbone structure for this model. Furthermore, the input to this model is time-series data, and a simple convolution operation usually focuses on local information while neglecting the temporal information of preceding and succeeding segments. In comparison, the multistage optimization module, which integrates stacked dilated convolutional layers, features an expanded receptive field that enables it to capture temporal information over a broader range. Consequently, the multistage optimization module is selected to refine the inferred contact location for enhanced accuracy. The proposed MS-AASTGCN for super-resolution localization of dynamic tactile stimulations, with the overall architecture shown in Figure 3f, consists of two main components: an AST-GC module and a multistage optimization module. The 23-channel voltage time-series signals from the tactile sensor array first perform convolution operations in both spatial and temporal dimensions via the AST-GC module; subsequently, the data passes through the multistage optimization module, which extracts deeper temporal information to refine the inference results; eventually, the data is processed through dimensionality reduction, yielding the inferred contact position as 2D coordinates: (X, Y). The MS-AASTGCN automatically learns both spatial and temporal patterns from data, further promoting super-resolution based on the optimized tactile sensor array.

In the MS-AASTGCN, the construction of the tactile sensor array's undirected temporal-spatial graph follows the work of spatial-temporal graph convolutional networks (ST-GCN).[29] In detail, each sensor in the array corresponding to a vertex of the graph is denoted as $V = \{v_{ti} | t = 1, \ldots, T, i = 1, \ldots, N\}$, where N is the number of tactile sensors (in this article N is 23) and T is the data sampling time (T frames). Additionally, there are two types of edges in the graph which are spatial edges and temporal edges, respectively, the spatial edges represent the intratopology connection of the tactile sensor array at each frame, denoted as $E_s = \{v_{ti} v_{tj} | (i, j) \in H\}$, where H is the set of topologically connected tactile sensors. The temporal edge is the interframe edge connecting the same tactile sensors in consecutive frames, denoted as $E_T = \{v_{ti} v_{(t+1)i}\}$. And a description of the graph structure in the form of an $N \times N$ adjacency matrix A where each weight represents the relationship between two vertexes A can be written as

$$A_{i,j} = \begin{cases} 1, & v_i \text{ and } v_j \text{ are connected} \\ 0, & \text{others} \end{cases} \tag{1}$$

This method of spatial-temporal graph construction enables the network architecture to work on diverse scales and topological structure sensor arrays with excellent scalability and transferability. The intratopology connections of the tactile sensor array at each frame are represented as $(A + I)$ by an adjacency matrix A and an identity matrix I representing self-connection.

As shown in Figure 3f, an AST-GC block consists of an adaptive spatial graph convolution (AConvs) and a temporal graph convolution (Convt). The implementation of the Aconvs within a single frame is formulated as[30]

$$f_{\text{out}} = (A_k + B_k + C_k)f_{\text{in}}W \tag{2}$$

where f_{in} is the measured data $x = (x_1, x_2, \ldots, x_{22}, x_{23})^T$ processed to obtain the input 23-channel features $f_{\text{in}} = (f_1, f_2, \ldots, f_{22}, f_{23})^T$, w represents related parameters of the AConvs operation. The adaptive adjacency matrix is divided into three $N \times N$ adjacency matrices: A_k, B_k, and C_k. The first part A_k is the original normalized adjacency matrix which represents the physical topological structure of the tactile sensor array. The second part B_k is a parameterized adjacency matrix whose elements can be arbitrary values and are parameterized and optimized together with the other parameters in the training process. B_k indicates the existence and strength of the connections between two taxels beyond the information in the predefined adjacency matrix A_k. The third part C_k is a data-dependent adjacency matrix expressing the connection between each pair of vertexes which learns a typical adjacency matrix for each sample. Notably, unlike the common graph convolution,[31] the adaptive adjacency matrix was introduced instead of a predefined adjacency matrix to perform graph convolution operations. This is because the values within the predefined adjacency matrix are relatively simplistic, with a large number of zero values, which means it can merely represent the physical connections among tactile sensors, as illustrated in Equation (1). When applied directly for convolution, this would result in a lack of valid information because multiplying by zero eliminates original data (multiplying zero by any number yields zero). As a result, during the convolution process, the focus tends to be on the nonzero positions within the matrix. In contrast, the values within the adaptive adjacency matrix are more diverse, allowing it to capture both explicit physical connections and hidden relationships among tactile sensors, through predefined and data-driven approaches, respectively. Introducing the adaptive adjacency matrix into spatial graph convolution can effectively address performance inconsistencies among tactile sensors caused by fabrication defects or external conditions, while also facilitating the capture of spatial features. Additionally, as shown in Equation (2), the operations between these three $N \times N$ adjacency matrices are additive, preventing any loss of information. The AConvs layer is shown in Figure S3, Supporting Information; in addition to the adaptive adjacency matrix there is a residual connection[32] to keep its initial features when the layer is inserted into any other modules. For the Convt, since the number of neighbors for each vertex is fixed as 2 (corresponding tactile sensors in two consecutive frames), a 1D convolution is applied on the output feature map calculated by AConvs.

The AST-GC module is a stack of nine basic blocks (AST-GC blocks), as shown in Figure 3f. After that, a spatial pooling is performed at the last block to aggregate the spatial features among the N tactile sensors.

After the AST-GC module, the multistage optimization module[33] is integrated into the network to optimize the inference of touch locations. The multistage optimization module is composed of several single-stage optimization (SSO) blocks stacked,

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
INTELLIGENT
SYSTEMS**

www.advintellsyst.com

and the output of each stage module is improved by the next stage module, as follows.

$$Y^0 = X_{out} \tag{3}$$

$$Y^s = F(Y^{s-1}) \tag{4}$$

where Y^0 is the input of the multistage optimization module, X_{out} is the output of the AST-GC module which can be viewed as the initial inference, Y^s is the output at stage s, Y^{s-1} is the output of the previous stage, and $F(\cdot)$ represents an SSO block which is shown in Figure 3f. The SSO block consists of stacked three dilated 1D residual convolutional layers with doubled dilation factors The multistage optimization module is a stack of four SSO blocks. In this way, we can obtain expanded receptive fields which help capture long-range dependencies and improve the quality of inferences. After the last SSO block, we used the global average pooling layer and a 1D convolutional layer to better retain the temporal and spatial structure of data. The resulting 2D output inference represents the coordinate values on the X-axis and Y-axis respectively, indicating the contact location on the surface of the tactile sensor array.

To be noted, the proposed MS-AASTGCN architecture shares some similarities with STGCN. However, in order to infer more accurately and efficiently the contact position of external stimuli for tactile super-resolution localization, a pruning operation was conducted on the STGCN model structure, reducing the number of model parameters to match the requirements of this task.

2.3. Contact Data Acquisition

A manipulator (robotic arm) equipped with a force sensor was used to exert external stimuli on the tactile sensor array and simultaneously record the coordinates of the contact position. To mimic human finger touch, a 3D printed finger-like indenter of a hemispherical silicone head (diameter, 20 mm) was connected to the force sensor. The tactile sensor array (72 mm by 80 mm) was uniformly sampled at 1-mm spacing in both X-axis and Y-axis directions, generating 5913 sampling points (73 by 81) in total. These sampling points' positions were defined as the center of the bottom of the indenter contacting the sensor array's surface and were encoded in Cartesian coordinates. While contacting each position to generate the datasets, an Arduino device with a specifically designed multichannel voltage acquisition external board (Figure S4, Supporting Information) was used to measure and record the voltage signals, which included the 23-channel tactile sensor output voltage signals. The data acquisition setup is shown in **Figure 4**a. This single-point acquisition assigns uniformly and individually distributed virtual taxels to cover the entire tactile sensor array, facilitating the MS-AASTGCN in inferring the contact position through virtual taxels.

The acquired output voltage data was processed by low-pass filtering to eliminate power frequency noise. Then, since the 23-channel voltage signals were time series data, a sliding window (Figure S5, Supporting Information) was applied to extract the voltage change sequences of the sampling points contacted states and uncontacted states. Eventually, min-max normalization was used to normalize the 23-channel voltage signals,

making it easier to compare the relative importance among the 23 tactile sensors by scaling them to the same range. After processing, the data comprised over 180 000 samples, which were split into datasets of training, validation, and testing with a ratio of around 3:1:1 and denoted by S_{train}, $S_{validation}$, S_{test}.

Aiming to gain deeper insight into the performance of the tactile sensor array, we visualized the receptive fields of each tactile sensor using the complete dataset obtained. Subsequently, the receptive fields of 23 tactile sensors were superimposed to generate the receptive field map of the entire tactile sensor array. The details were as follows: the receptive field of each tactile sensor was evaluated based on whether it was activated and the degree of activation at different contact positions, with the activation threshold set to 0.2 V. First, the activation count for each position in the tactile sensor array was initialized to 0. Each time a contact position activated the tactile sensor, the activation count for that position was incremented by 1. Ultimately, a receptive field map was generated, as shown in Figure 4b, representing the ease of activation for each position, with darker colors indicating easier activation. The receptive field map revealed notable inconsistencies among the tactile sensors, primarily demonstrated in the varying sizes of each sensor's receptive field and their differing sensitivities to stimuli. This observation further indicates that a solely topologically structure-based predefined original adjacency matrix cannot fully capture the connections among the tactile sensors. This limitation arises from the construction of the original adjacency matrix, which relies on the assumption that the performance of the vertexes within the topological structure is entirely uniform, implying that each tactile sensor within this array should exhibit identical, which is contrary to the actual situation. To address this issue, we introduced the adaptive adjacency matrix in the MS-AASTGCN. As shown in Figure 4c, both the original adjacency matrix and the adaptive adjacency matrix are illustrated. It is evident that the original adjacency matrix has a significant amount of missing information (i.e., many values within the matrix are zero), whereas the adaptive adjacency matrix reveals more complex hidden connections among the tactile sensors. This distinction is primarily reflected in two aspects: first, the adaptive adjacency matrix is devoid of numerous zero values, avoiding the loss of valid information; second, the values it encompasses span a wider range, including both positive and negative values, which express more diverse and detailed connections between the tactile sensors. The connections among the tactile sensors are represented by the adaptive adjacency matrix, which includes fixed connections based on the tactile sensor array's topological structure and hidden connections learnt from the acquired data, facilitating a more effective resolution of the inherent inconsistencies in the tactile sensor array.

2.4. Experiments and Validation

2.4.1. Evaluation of the MS-AASTGCN Model

In this section, we first verified whether the MS-AASTGCN model can accurately estimate the contact location based on the original sensor measurements. The training set S_{train} was used to determine the hyperparameters of the training process, resulting in a batch size of 256 and a learning rate of 0.003.
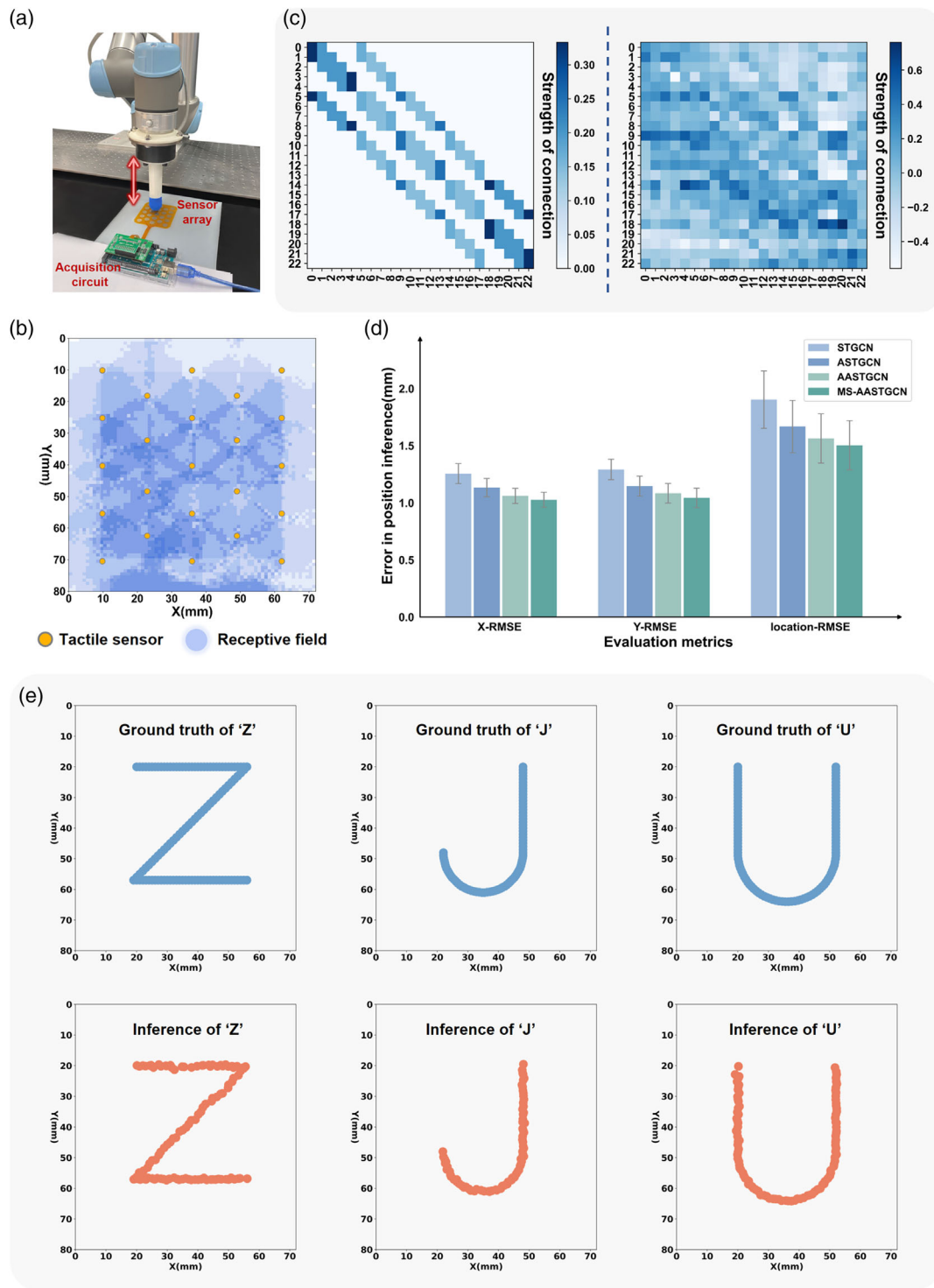
Figure 4. Experimental results validating the effectiveness of the MS-AASTGCN. a) Illustration of real-time data acquisition of 5913 different contact positions achieved by a manipulator. b) A receptive field map of the tactile sensor array generated by superimposing 23 tactile sensors' receptive fields. c) The contrast between the original adjacency matrix (left) and the adaptive adjacency matrix (right). d) Evaluation results of the ablation experiments across three evaluation metrics. The X-RMSE represents RMSEs in the X-axis, the Y-RMSE represents RMSEs in the Y-axis, and the location-RMSE represents RMSEs in the overall position. Error bars indicate standard deviations. e) Validation results of the contact localization with super-resolution. The top: The ground truth position of the letters "Z," "J," and "U". The bottom: The inferences of the MS-AASTGCN for the ground truth positions of "Z," "J," and "U".

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
INTELLIGENT
SYSTEMS**
Open Access

www.advintellsyst.com

Afterward, this model was trained on S_{train} with the purpose of minimizing the RMSEs between its inputs and the ground truth labels. The loss function of one output vector $o^i = (o_1^i, o_2^i)^T$ and its corresponding label $l^i = (l_1^i, l_2^i)^T$ is defined as follows.

$$L^i(o^i, l^i) = \sqrt{\frac{1}{2}\sum_{j=1}^{2}\left(o_j^i - l_j^i\right)^2} \qquad (5)$$

For the optimizer, we used Adam with a learning rate of 0.003 which was divided by 10 at the 35th and 45th epoch. The training process was ended at the 100th epoch.

The testing set S_{test} was used to evaluate the trained model, which reported RMSEs of 1.50, 1.02, and 1.04 mm in overall positioning, X-axis positioning, and Y-axis positioning, respectively. Combining the aforementioned experimental results with the super-resolution factor formula[24] as follows

$$\Omega = \frac{A}{n \cdot \pi \cdot RMSE_x \cdot RMSE_y} = \frac{80 \cdot 72}{23 \cdot \pi \cdot RMSE_x \cdot RMSE_y} \qquad (6)$$

where A is the sensor array's area, n is the number of tactile sensors. Our 80 mm by 72 mm tactile sensor array achieved a super-resolution factor of 75.19 with the MS-AASTGCN. Additionally, all experiments were conducted on the PyTorch deep learning framework with the Nvidia GeForce RTX 4090 GPU.

2.4.2. Ablation Experiment of MS-AASTGCN Model

The MS-AASTGCN model was based on the STGCN model, to which the additional modules were added. We took the pruned STGCN model as the baseline model for the ablation experiment, using the data collected by ourselves as the experimental data, and gradually built and verified the components. **Table 2** and Figure 4d show the evaluation results of the ablation experiments across three evaluation metrics (RMSEs in X-axis positioning, Y-axis positioning, and overall positioning), including the AST-GC module and the multistage optimization module. Each model was trained using the same training set S_{train} and tested on the same testing set S_{test}.

The AST-GC module in the MS-AASTGCN model was responsible for obtaining the real connections between each tactile sensor expressed through the adjacency matrix, capturing not only fixed connections due to topological structure but also flexible connections influenced by non-negligible human factors such as variations in the fabrication of the tactile sensor array. To verify the practical benefits of adopting the AST-GC module, the ablation experiments compared STGCN (the baseline model without the AST-GC module), ASTGCNs (ASTGCN, the model with the AST-GC module), and attention-based ASTGCNs (AASTGCN, the model with the attention-based AST-GC module). According to the results, the model with the AST-GC module, especially with an attention-based AST-GC module, outperformed the model without it in all three aspects. Specifically, the RMSEs were reduced by 15.20% in X-axis positioning, 16.28% in Y-axis positioning, and 17.89% in overall positioning. Moreover, the variations in all aspects were smaller when using the AST-GC module, illustrating its better stability. Thus, the effectiveness of the AST-GC module has been validated.

Then, the multistage optimization module was added to the AASTGCN, resulting in a new model called MS-AASTGCN. The results indicated that the performance of the MS-AASTGCN model was further improved in all three aspects. Specifically, compared to the baseline model (STGCN), the RMSEs were reduced by 18.79% in X-axis positioning, 19.38% in Y-axis positioning, and 21.05% in overall positioning. These results indicate that the MS-AASTGCN's performance was significantly enhanced through the addition of modules. The convergence of the training of MS-AASTGCN and the sample statistics of the MS-AASTGCN testing results can be found in Figure S6, Supporting Information.

Taking a step further, to validate the performance of the MS-AASTGCN in real-world applications, we utilized the same data set acquisition approach as described previously to contact different positions of the tactile sensor array point by point in "Z," "J," and "U" patterns, respectively. The results are shown in Figure 4, indicating that the three letters were accurately fitted. Moreover, the localization errors at most contact points were found to be less than 1.01 mm, demonstrating that the proposed localization method is both valid and practical for real-world applications.

2.4.3. Experiment of Rapid Response of the Tactile Sensor Array to Dynamic Subtle Tactile Stimuli

Considering practical interaction scenarios, dynamic tactile perception is also of great significance for robots to achieve fast responses and deft manipulation, similar to those of the human body.[34] In this article, we utilized the SETENG-based tactile sensor's sufficiently fast response to mimic the function of RA-I, which is most responsive to changing contact conditions, such as when an external stimulus makes or breaks contact. Owing to SETENG's working mechanism and the effective topological structure of the tactile sensor array, the robotic skin can rapidly detect dynamic stimuli over an area of low sensor density.

The voltage signal strength of each tactile sensor is determined by its distance from the stimulus source, which can be

Table 2. The ablation experiment results regarding RMSEs in X-axis positioning, Y-axis positioning, and overall positioning.

Model	RMSEs in X-axis positioning [mm]	RMSEs in Y-axis positioning [mm]	RMSEs in overall positioning [mm]
STGCN	1.26 ± 0.08	1.29 ± 0.09	1.90 ± 0.25
ASTGCN	1.13 ± 0.08	1.15 ± 0.08	1.67 ± 0.23
AASTGCN	1.06 ± 0.07	1.08 ± 0.08	1.56 ± 0.21
MS-AASTGCN	1.02 ± 0.06	1.04 ± 0.07	1.50 ± 0.20

observed in Video S1, Supporting Information, where the corresponding voltage signal of the 23-channel undergoes distinct changes when the stimulus is applied to different positions. Besides this, the video also demonstrates that our tactile sensor array exhibits a rapid response time to external stimuli with minimal time delay.

Furthermore, to demonstrate the rapid response of the tactile sensor array to dynamic subtle tactile stimuli, we facilitated the free fall of a 2.715 g table tennis ball (Figure S7, Supporting Information) above the tactile sensor array. The MS-AASTGCN was used to collect and process signals to localize each contact during the table tennis ball bounce (**Figure 5**a

and Video S2, Supporting Information). The corresponding contact positions were displayed on the screen; Figure 5b illustrates the deployment of the experiment. As results show (Figure 5c), this records the contact position of the table tennis ball at 1, 2, 3, and 4 s during the experiment. Each image recorded 5 contact positions, including the current contact position and four previous contact positions, indicated by varying colors based on contact time (with earlier positions appearing lighter) to illustrate the trajectory of the stimuli. This allows for the acquisition of all contact positions of the table tennis ball throughout its entire movement. The results showed that our SETENG-based tactile sensor array, in conjunction with the MS-AASTGCN, enabled the
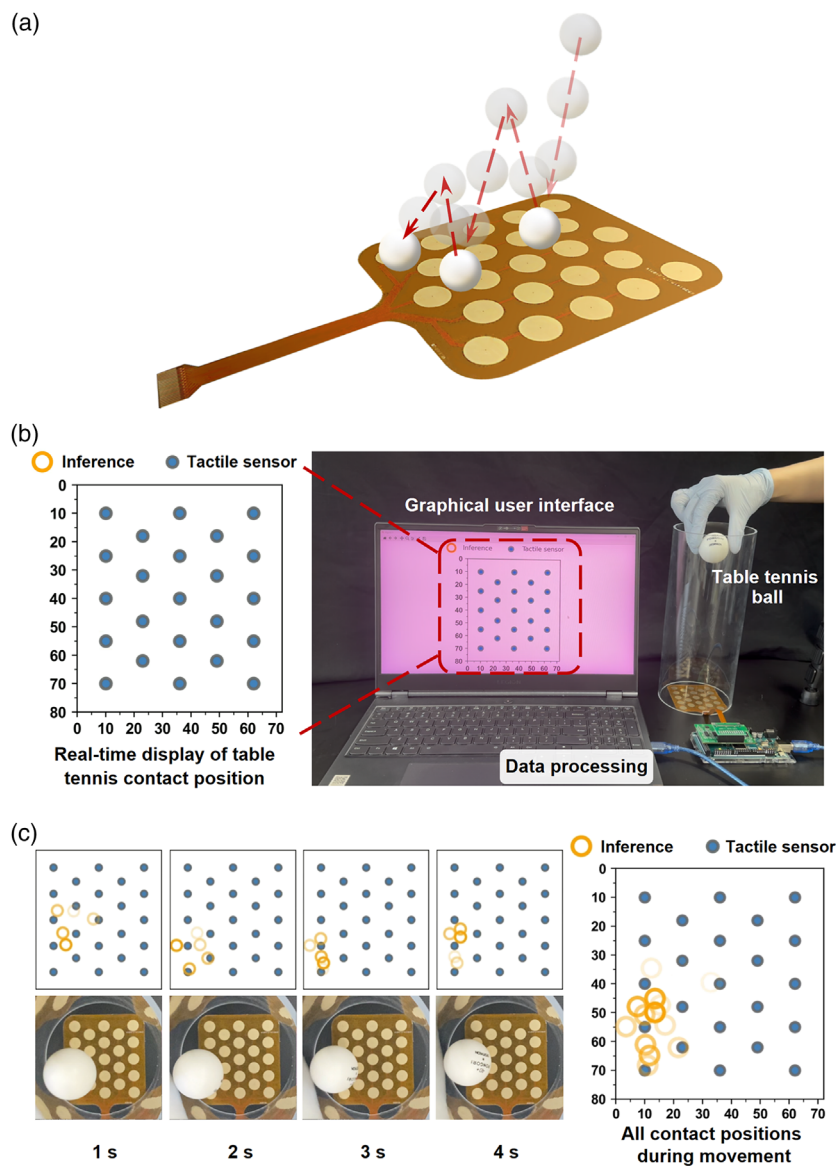

Figure 5. Application of SETENG-based tactile sensor array on tracking the bouncing trajectory of a table tennis ball. a) Schematic of a table tennis ball bouncing on the surface of the tactile sensor array. b) The deployment of an experiment to track the contact positions of table tennis ball bounces. c) An experiment tracking the contact positions of a table tennis ball's bounces. Left top: The display status of the GUI during the experiment. Left bottom: A top-down view recording the contact positions of a table tennis ball during the experiment. Right: All contact positions of a table tennis ball recorded during its movement.

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
INTELLIGENT
SYSTEMS**
Open Access

www.advintellsyst.com

robotic skin to successfully and rapidly perceive and localize dynamic subtle stimuli, such as the bounce of a table tennis ball.

To the best of our knowledge, this kind of rapid-response super-resolution localization for dynamic subtle stimuli represents a notable achievement that has not been achieved in previous related research. To be noted, in the experiment for the trajectory tracking of the table tennis ball, the voltage signal generated by the contact between the table tennis ball and the tactile sensor array was never seen at any stage of model training. The experimental result further proved that the MS-AASTGCN exhibits its excellent generalization ability.

3. Conclusion

In this article, a tactile sensor array that simultaneously integrates super-resolution and rapid response capabilities for the tactile perception of dynamic and subtle stimuli was achieved by mimicking the location-sensing mechanism of human skin. The effective combination of the topological structure of the tactile sensor array and the graph structure-based MS-AASTGCN endows the tactile sensor array with super-resolution, even at a low taxel density, and the introduced adaptive adjacency matrix reveals hidden connections among the taxels based on the acquired signals, significantly enhancing the generalization ability of the MS-AASTGCN. Corresponding experiments and tasks were conducted to validate the performance of this super-resolution algorithm-sensor system. As a result, the MS-AASTGCN was validated to be effective through ablation experiments and demonstrated excellent generalization ability in real-world applications. In the task of inferring contact positions, the tactile sensor array achieved RMSEs of 1.50, 1.02, and 1.04 mm in overall positioning, X-axis positioning, and Y-axis positioning, eventually reaching a super-resolution factor of 75.19. Moreover, the tactile sensor array exhibited a rapid response capability to external dynamic subtle stimuli due to the inherent performance attributes of SETENG. The results from tracking the contact positions of table tennis ball bounces suggested that the tactile sensor array is capable of perceiving subtle and dynamic stimuli with rapid responsiveness, demonstrating its potential for more challenging interaction tasks.

This work offers new insights into robotic tactile skin by effectively combining hardware design and super-resolution algorithms to create a rapid-response tactile sensing system for detecting dynamic subtle stimuli. It also addresses hardware-level sensor inconsistencies from a software perspective. So far, our training data for contact position inference has focused solely on single-point touches, leaving the force perception capabilities of tactile sensors unexplored. Future research will aim to achieve multipoint perception in combination with force perception, utilizing the dataset acquired through single-point acquisition.

4. Experimental Section

Materials: Polytetrafluoroethylene (PTFE) film tapes (50 μm) were purchased from a company named Chenguang Plastic (China).

Fabrication of SETENG-Based Tactile Sensor Array: The substrate and electrode layers of SETENG-based tactile sensor array were fabricated by a company named Shenzhen JLC Electronics (China) adopting the commercial flexible PCB (FPCB) processing methods. Next, the PTFE film tape was cut to the same size as the substrate layer and applied to the top to obtain the tactile sensor array.

Characterizations: A linear motor (DA60-B1-T60-C010-0.2, Dynamikwell Technology, China) was used to apply different stimuli for the characterization of triboelectric performance. The open-circuit voltage was collected by an electrometer (Keithley 6514, Tektronix, USA). For the data acquisition in all applications, the open-circuit voltage was obtained via a portable Arduino Mega 2560 board with a multichannel voltage acquisition external board which was fabricated by a company named Shenzhen JLC Electronics (China).

Data Acquisition of Datasets: The ATmega328P microcontroller (Microchip Technology Inc., USA) was used as the processor for the data acquisition module. 23 channels were configured on the voltage acquisition external board through AD7689 analog-to-digital converters (Analog Devices, Inc., USA) to record voltage signals of the tactile sensor at a sampling rate of 75 Hz. For all datasets used for contact position inference, the UR5 manipulator (Universal Robots, Inc., Denmark) was used to exert an external stimulus on the tactile sensor array and record the coordinates of the contact position. The tactile sensor array (72 mm by 80 mm) was uniformly sampled at 1 mm spacing in both X-axis and Y-axis directions, generating 5913 sampling points (73 by 81) in total, each of which was sampled once. The acquired data was processed into time series data, consisting of over 180 000 samples, which were split into datasets of training, validation, and testing with a ratio of around 3:1:1. The same data acquisition module was used to collect the dataset of the three characters "Z," "J," "U" verifying the validity of the MS-AASTGCN in practical applications.

Segmentation Approach of Datasets: The sliding window (Figure S5, Supporting Information) was used to divide the whole dataset into numerous samples of voltage change sequences including contacted states of each sampling position and uncontacted states. The size of the sliding window should be precisely sized to encompass the voltage rising process of each contact action, that is, a sliding window can encompass the voltage variation features of the 23 tactile sensors at each sampling position. The size of the sliding window was treated as a hyperparameter of the deep learning model and the optimal window size was found through hyperparameter random search experiments. Then, using the divided data samples, the deep learning model learnt the voltage variation patterns among the 23 tactile sensors under different contact positions and uncontacted states to infer the future contact states and contact positions.

MS-AASTGCN Model: The MS-AASTGCN model consisted of two main components: an AST-GC module and a multistage optimization module. The AST-GC module consisted of nine stacked AST-GC blocks, with the number of output channels for each block being 8, 8, 16, 16, 16, 32, 32, 64, and 64. After that, a spatial pooling was performed at the last block to aggregate the spatial features among the N tactile sensors. The structure of the AST-GC block is shown in Figure S8, Supporting Information, which primarily consists of AConvs and Convt. More specifically, the AConvs layer was followed by a batch normalization (BN) layer, a rectified linear unit (ReLU) layer, and a dropout layer with a drop rate of 0.25. The Convt was followed by a BN layer and a ReLU layer. To stabilize the training, a residual connection was added to each spatial-temporal convolution block. The multistage optimization module was composed of four stacked SSO blocks, with each block containing three dilated convolutional layers, and the number of filters in all dilated convolutional layers being eight. The structure of the SSO block is shown in Figure S8, Supporting Information, which consisted of stacked three dilated 1D residual convolutional layers with doubled dilation factors and then followed by a BN layer, a ReLu layer, and a residual connection between the original input and the convolutional output.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Shuyao Zhou: conceptualization (equal); data curation (lead); software (lead); validation (lead); visualization (lead); writing—original draft (lead); writing—review & editing (lead). **Depeng Kong**: conceptualization (lead); methodology (supporting); software (supporting); writing—review & editing (lead). **Mengke Wang**: methodology (supporting); software (supporting); **Baocheng Wang**: methodology (supporting); writing—review & editing (supporting). **Yuyao Lu**: investigation (equal); methodology (supporting); writing—review & editing (supporting). **Honghao Lyu**: funding acquisition (equal); methodology (supporting); supervision (supporting). **Zhangli Lu**: software (supporting); writing—review & editing (supporting). **Yong Tao**: investigation (equal); project administration (equal). **Kaichen Xu**: resources (equal); supervision (supporting). **Geng Yang**: funding acquisition (lead); supervision (lead). **Shuyao Zhou** and **Depeng Kong** contributed equally to this work.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords

[1] A. Billard, D. Kragic, *Science* **2019**, *364*, 6446.
[2] Q. Li, O. Kroemer, Z. Su, F. F. Veiga, M. Kaboli, H. J. Ritter, *IEEE Trans. Rob.* **2020**, *36*, 1619.
[3] V. E. Abraira, D. D. Ginty, *Neuron* **2013**, *79*, 4.
[4] S. Maksimovic, M. Nakatani, Y. Baba, A. M. Nelson, K. L. Marshall, S. A. Wellnitz, P. Firozi, S. Woo, S. Ranade, A. Patapoutian, E. A. Lumpkin, *Nature* **2014**, *509*, 617.
[5] M. S. Fleming, W. Luo, *Front. Biol.* **2013**, *8*, 4.
[6] C. Wang, L. Dong, D. Peng, C. Pan, *Adv. Intell. Syst.* **2019**, *1*, 1900090.
[7] M. Wang, T. Wang, Y. Luo, K. He, L. Pan, Z. Li, Z. Cui, Z. Liu, J. Tu, X. Chen, *Adv. Funct. Mater.* **2021**, *31*, 2008807.
[8] M. Wang, Y. Luo, T. Wang, C. Wan, L. Pan, S. Pan, K. He, A. Neo, X. Chen, *Adv. Mater.* **2021**, *33*, 2003014.
[9] H. Lee, H. Sun, H. Park, G. Serhat, B. Javot, G. Martius, K. J. Kuchenbecker, *IEEE Trans. Autom. Sci. Eng.* **2023**, *20*, 425.
[10] K. Park, H. Yuk, M. Yang, J. Cho, H. Lee, J. Kim, *Sci. Rob.* **2022**, *7*, eabm7187.
[11] D. Kong, G. Yang, G. Pang, Z. Ye, H. Lv, Z. Yu, F. Wang, X. V. Wang, K. Xu, H. Yang, *Adv. Intell. Syst.* **2022**, *4*, 2200050.
[12] M. Kim, H. Choi, K. Cho, S. Jo, *IEEE Rob. Autom. Lett.* **2021**, *6*, 4970.
[13] X. Wang, H. Zhang, L. Dong, X. Han, W. Du, J. Zhai, C. Pan, Z. L. Wang, *Adv. Mater.* **2016**, *28*, 201503407.
[14] Z. Yan, L. Wang, Y. Xia, R. Qiu, W. Liu, M. Wu, Y. Zhu, S. Zhu, C. Jia, M. Zhu, R. Cao, Z. Li, X. Wang, *Adv. Funct. Mater.* **2021**, *31*, 2100709.
[15] X. Zhi, S. Ma, Y. Xia, B. Yang, S. Zhang, K. Liu, M. Li, S. Lia, P. Wan, X. Wang, *Nano Energy* **2024**, *125*, 109523.
[16] L. Lin, Y. Xie, S. Wang, W. Wu, S. Niu, X. Wen, Z. L. Wang, *ACS Nano* **2013**, *7*, 8266.
[17] J. Yun, N. Jayababu, D. Kim, *Nano Energy* **2020**, *78*, 105325.
[18] J. Zhu, S. Ji, J. Yu, H. Shao, H. Wen, H. Zhang, Z. Xia, Z. Zhang, C. Lee, *Nano Energy* **2022**, *103*, 107766.
[19] Y. Lu, D. Kong, G. Yang, R. Wang, G. Pang, H. Luo, H. Yang, K. Xu, *Adv. Sci.* **2023**, *10*, 2303949.
[20] Y. Yan, Z. Hu, Z. Yang, W. Yuan, C. Song, J. Pan, Y. Shen, *Sci. Rob.* **2021**, *6*, eabc8801.
[21] L. Massari, G. Fransvea, J. D'Abbraccio, M. Filosa, G. Terruso, A. Aliperta, G. D'Alesio, M. Zaltieri, E. Schena, E. Palermo, E. Sinibaldi, C. Oddo, *Nat. Mach. Intell.* **2022**, *4*, 425.
[22] B. Ward-Cherrier, N. Pestell, L. Cramphorn, B. Winstone, M. E. Giannaccini, J. Rossiter, N. F. Lepora, *Soft Rob.* **2018**, *5*, 20170050.
[23] Z. Hu, L. Lin, W. Lin, Y. Xu, X. Xia, Z. Peng, Z. Sun, Z. Wang, *Adv. Intell. Syst.* **2023**, *5*, 2200371.
[24] H. Sun, G. Martius, *Sci. Rob.* **2022**, *7*, eabm0608.
[25] Å. B. Vallbo, R. S. Johansson, *Hum. Neurobiol.* **1984**, *3*, 14.
[26] Å. B. Vallbo, K. Å. Olsson, K.-G. Westberg, F. J. Clark, *Brain* **1984**, *107*, 727.
[27] R. S. Johansson, *J. Physiol.* **1978**, *281*, 101.
[28] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, P. S. Yu, *IEEE Trans. Neural Networks Learn. Syst.* **2021**, *32*, 4.
[29] S. Yan, Y. Xiong, D. Lin, in *Proc. 2018 AAAI Conf. Artif. Intell.*, AAAI, New Orleans, LA, February **2018**.
[30] L. Shi, Y. Zhang, J. Cheng, H. Lu, in *Proc. 2019 IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, IEEE, Long Beach, CA, June **2019**, p. 12026.
[31] T. Kipf, M. Welling (Preprint), arxiv:1609.02907, v4, Submitted: February. **2017**.
[32] K. He, X. Zhang, S. Ren, J. Sun, in *Proc. 2016 IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, IEEE, Las Vegas, NV, June **2016**, p. 770.
[33] Y. A. Farha, J. Gall, in *Proc. 2019 IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, IEEE, Long Beach, CA, June **2019**, p. 3575.
[34] R. Balasubramanian, V. J. Santos, *The Human Hand as an Inspiration for Robot Hand Development*, Springer, Cham, Switzerland **2014**.