Smart electronic skin having gesture recognition function by LSTM neural network

G. Y. Liu,[1] D. Y. Kong,[1] S. G. Hu,[1] Q. Yu,[1] Z. Liu,[2] T. P. Chen,[3] Y. Yin,[4] Sumio Hosaka,[4] and Y. Liu[1,a)]

[1]*State Key Laboratory of Electronic Thin Films and Integrated Devices, University of Electronic Science and Technology of China, Chengdu 610054, People's Republic of China*
[2]*School of Materials and Energy, Guangdong University of Technology, Guangzhou 510006, People's Republic of China*
[3]*School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798*
[4]*Graduate School of Engineering, Gunma University, 1-5-1 Tenjin, Kiryu, Gunma 376-8515, Japan*

(Received 17 May 2018; accepted 6 August 2018; published online 22 August 2018)

Rapid growth of soft electronics has enabled various approaches for developing artificial skin. However, currently existing electronic skin is still facing some problems such as high fabrication complexity, high production cost, and smartness of recognizing the stimulus automatically. In this work, we report a simple, low-cost Polydimethylsiloxane (PDMS)-based smart electronic skin system, consisting of a sensor array and a data processing system. The sensor array can be easily mounted on the human body or robot hand as a result of excellent softness, stretchability, and bendability of PDMS. Signals from the sensor array are processed by a Long and Short Term Memory neural network algorithm in the data processing system. The trained data processing system can recognize four types of gestures at an accuracy of 85 ± 5%, even taking into account environmental variations including folding, curvature, tensile strength, temperature, and endurance cycles. This work proves that this type of skin can be endowed with intelligence with a proper neural network algorithm and fabricated at low cost and reduced complexity. *Published by AIP Publishing.*
https://doi.org/10.1063/1.5040413

The human skin plays a vital role in our interactions with the world around, and it allows us to sense and distinguish external stimuli, such as external temperature and humidity changes, as well as surface roughness and object softness.[1–3] Much effort has been made to design and manufacture artificial electronic skin (e-skin),[4–7] which can mimic the functions of biological skin. The tactile sensing function of e-skin plays an essential role in supplying action related information such as slipping and pressing. The e-skin with tactile sensing holds great potential for applications in health medical equipment and bionic devices, e.g., continuous examination of blood pressure in health monitoring and bionic skin in the bionic robot. Thus, e-skin with the tactile sensing function has drawn much attention recently.[8–11] For example, Sohn *et al.* reported an e-skin using a single-layered piezoresistive Multi-Wall Carbon Nanotube-Polydimethylsiloxane (MWCNT-PDMS) composite film as a sensing component,[12] which can detect the level and position of the pressure; Yang *et al.* reported an e-skin system based on arrays of Graphene Woven Fabrics (GWFs),[13] which is applicable to acoustic signal acquisition and human motion detection. Wang reported an e-skin based on a sandwich structure, in which the bottom and top electrodes were PDMS embedded with Ag Nanowire (AgNW) strips.[14] It was demonstrated that the e-skin can capture tactile images.

Although e-skin with a tactile sensing function has made some progress, there are still many obstacles on the way to practical application. Some of the sensing components in the e-skin have the problems of high cost/complexity. For example, Lou *et al.* synthesized "rGO" nanosheet wrapped nanofibers as a 3D network conductive film.[15] Some others realized a complicated sensor of PDMS pyramids, which involves complex pattern processes and film assembly processes.[11] These sensors with complex materials or complicated device structures may have a high sensitivity. However, in some application scenarios such as bionic robots,[16–20] cost and production ability are more concerned than sensitivity. Furthermore, although extensive studies have been carried out on artificial skin sensors and materials, there are still few works focusing on the smartness of e-skin, i.e., how to use artificial neural networks to realize the recognition of touching, gestures, pinning, etc. Such operation is similar to the working of real human skin.

In this work, we report an e-skin system which consists of both hardware and software, including a sensor array fabricated on hand-made PDMS with an intelligent data processing system based on the neural network algorithm. The neural network algorithm can identify gestures based on the signals from the sensor array. The e-skin system can recognize four types of gestures after training on a Graphical Processing Unit (GPU) platform. Folding and stretching of the sensor array and temperature variation show little impact on the recognition accuracy of the e-skin.

The sensor array has the crossbar architecture shown in Fig. 1(a), and each sensor has a capacitor-like structure with a PDMS layer as the capacitor dielectric sandwiched between two conductive silver layers. The detailed manufacturing process of the PDMS film is presented in supplementary material Note 1. A brief description of the sensor array fabrication process is described as follows. Conductive silver paste was

a)E-mail: yliu1975@uestc.edu.cn

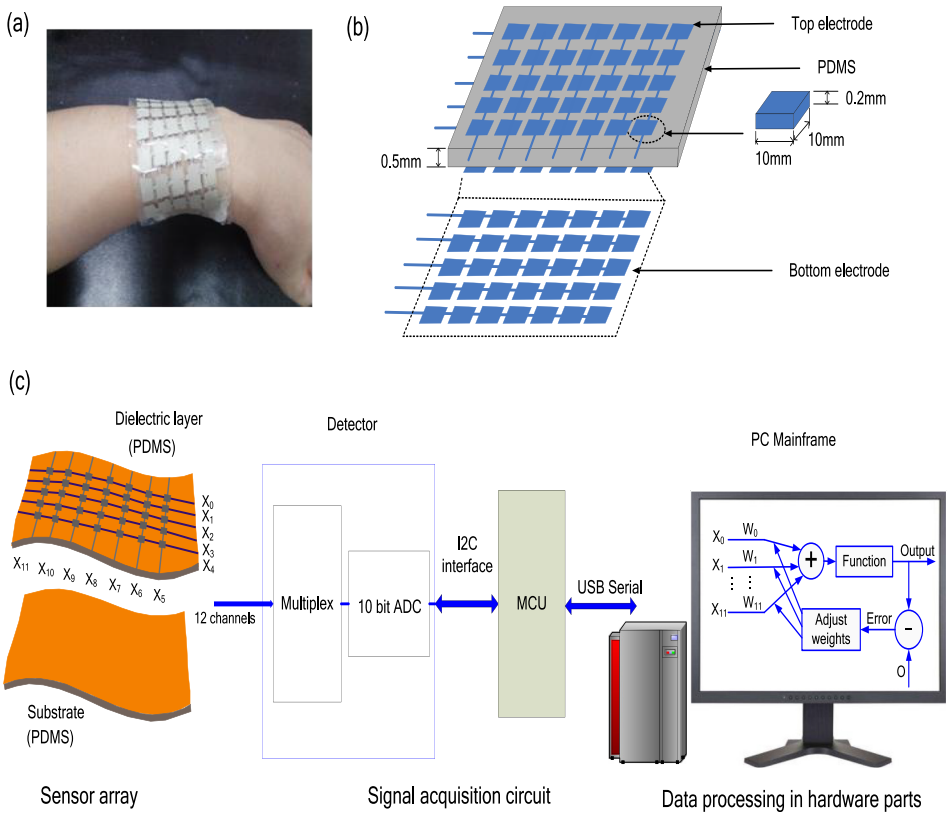

FIG. 1. Architecture diagram of the intelligent electronic skin system. (a) Photographs of the e-skin sensor array with the 5×7 crossbar architecture. Each sensor unit on the cross point has a capacitor-like structure with a PDMS layer sandwiched between two conductive silver electrodes; (b) Cross-sectional image about the structure of the sensor array; (c) the e-skin system consisting of the sensor array component and the data processing unit.

first printed on the top and bottom sides of the PDMS layer to construct seven silver-strip top electrodes and five silver-stripe bottom electrodes for the capacitor. Finally, the as-fabricated sensor array was formed by experiencing a natural drying. The fabricated flexible e-skin contains 5×7 capacitive touch cells, as shown in Fig. 1(a). Figure 1(b) shows the cross-sectional view of the as-fabricated e-skin. Electrodes (Sensor array) were grown and patterned on both top and bottom of PDMS by a conductive silver paint pen which has a pen point of around 400 μm. The thickness of the electrode was around 200 μm as indicated in Fig. 1(b).

The diagram of the e-skin system is shown in Fig. 1(c). In order to enable the e-skin system to smartly respond to the external stimulus, a data processing system was implemented. In this work, Arduino uno (Atmega328P-PU) was used as the data acquisition platform. MPR121 (proximity capacitive touch sensor controller, Freescale Semiconductor) was used as the capacitive touch sensor controller. The Serial Interface was realized using the Serial Clock Line (SCL) and Serial Data Line (SDL) to achieve bi-direction commutation between MPR121 and Arduino uno. The design details of the signal acquisition module are given in supplementary material Note 2. The acquired data were sent to the host computer via USB. The flow chart of the MCU program is shown in Fig. S3(a-i). Signal acquisition module with 12 input channels was used to collect the electric signal from the sensor array. In the experiment of touch sensing, a digital push-pull meter was used to apply the force using Digital Force Gauge Model SL-5500 to each sensor, as shown in Fig. 2(a). Figure 2(b) shows that capacitances for all touch sensors in the array increase with applied force in the range of 0–25 N. This indicates that all the touch sensors work properly and reliably.

The neural network algorithm can enable the sensor array to smartly recognize the external stimulus. In this work, a Long and Short Term Memory (LSTM) neural network algorithm is adopted to process the data collected from the sensor array. The neural network works according to a flowchart as shown in Fig. S3(a-ii). Four types of gestures as shown in Fig. S3(b) can be recognized by the e-skin system. To prepare the training data of the neural network model, the data of 1250 times of operation were collected for each of the four types of gestures. To strengthen the generalization of the neural network model, noise in the range of $(-10, 10)$ was added on the total 5000 samples (i.e., 1250×4) for the four types of gestures. The neural network model was trained on the Platform of Tensorflow offline running on a host computer equipped with an Nvidia Graphic Processing Unit (GPU) 1080Ti.

Figure 3 shows the neural network model consisting of three components, i.e., a single-layer perception (SLP) as the input, a Long and Short Term Memory (LSTM) model as the

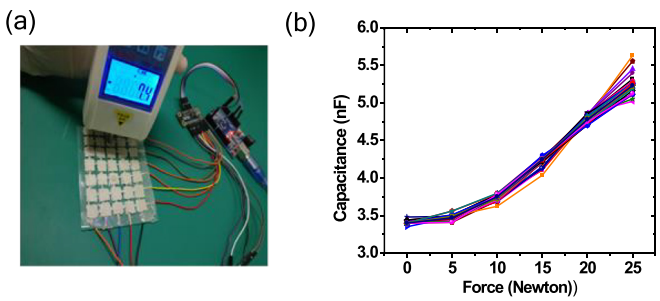

FIG. 2. Relationship between the force and capacitance of the sensor array. (a) Photograph of the push-pull meter for applying force to the sensors; (b) relationship between the applied force and capacitance for the 35 sensors.

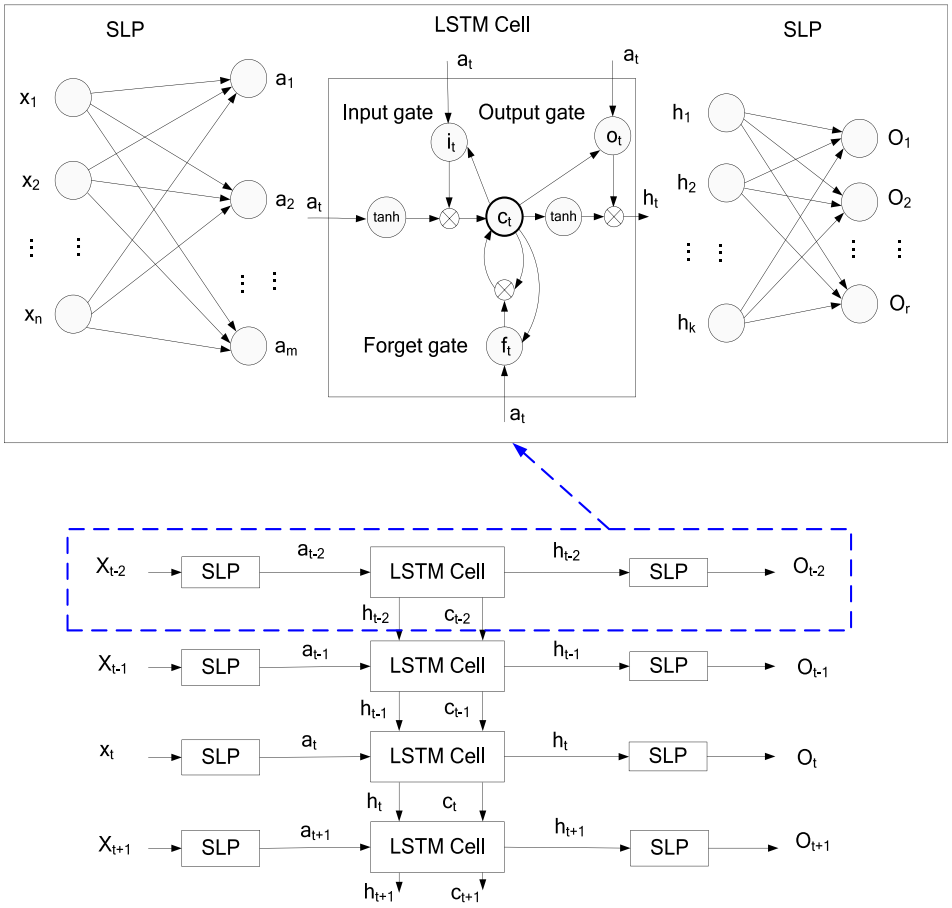

FIG. 3. Framework of the LSTM neural network algorithm. The neural network consists of three components: a SLP layer as an input component, a LSTM model as a core processing component, and a SLP layer as an output component.

core processing component, and a SLP as the output. The LSTM model has been widely adopted to solve the long-term dependency problem.[21–24] The LSTM cell has three gates i.e., the forget gate, the input gate, and the output gate. Equation (1) realizes the sigmoid function. Equation (2) represents the forget gate which determines the information to be discarded from the state of the LSTM Cell. The output range of the sigmoid function locates in [0, 1]. Equation (3) presents the tanh function of the equation. Equation (4) represents the input gate that determines the new information to be added. The output gate function of Eq. (5) determines the magnitude of the current cell state to be filtered out. The output of the overall LSTM is defined as h_t based on Eq. (6)

$$sigmoid(x) = \frac{1}{1+e^{x}}, \tag{1}$$

$$f(t) = sigmoid(W_{af} \times a_t + W_{af} \times h_{t-1} + b_f), \tag{2}$$

$$\tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}, \tag{3}$$

$$i(t) = sigmoid(W_{ai} \times a_t + W_{hi} \times h_{t-1} + W_{ci} \times c_{t-1} + b_i), \tag{4}$$

$$o(t) = sigmoid(W_{ao} \times a_t + W_{ho} \times h_{t-1} + b_o), \tag{5}$$

$$h(t) = o_t \times \tanh(c_t). \tag{6}$$

In the above equations, W terms denote the weight matrices (e.g., W_{af} is the input of the forget gate weight matrices) and b terms denote the bias vectors (e.g., b_f is the vector of the forget gate). Supplementary material Note 4 describes the LSTM network in detail. A single gesture involves four sensors, and the collected data from each sensor unit are defined as x^{t-2}, x^{t-1}, x^{t}, and x^{t+1}, respectively. The amount of sensor units contained in a gesture determines the time step of the LSTM neural network model as four. x^{t-2}, x^{t-1}, x^{t}, and x^{t+1} are the 12-dimentional data in 12 channels. Taking x^{t-2} as an example, $x^{t-2} = (x_0^{t-2} x_1^{t-2} \ldots x_{11}^{t-2})$. The shape of input data of the neural network is marked as S(B, T, D), where B denotes the batch size, T denotes the time step of the neural network, and D denotes the dimension of data vector. The input SLP layer nonlinearly maps input data from S(250, 4, 12) to S(250, 4, 30). The weight matrix of the SLP is $W^{12\times 30}$, which is randomly initialized according to the standard normal distribution function. The bias applied to neurons of the SLP second layer is a vector of 30 elements, and the value of each element is −1. The output of the SLP layer is fed to the LSTM Cell, and the shape of LSTM input data is S(250, 4, 30). The amount of hidden neurons is forty in the LSTM Cell, and the shape of output data is S(250, 4, 40). In each batch, the output of LSTM is defined as h^{t-2}, h^{t-1}, h^{t}, and h^{t+1}. The dimension of the bias applied to neurons of the second SLP layer is four, and the element value of each dimension is −1. In each batch, the softmax function converts output of the neural network model into probability, and the shape of softmax layer's output data is S(1, 250, 4). The softmax function can be expressed as

$$p(x_i) = \frac{e^{x_i}}{\sum_{i=1}^{4} e^{x_i}}, \tag{7}$$

where x_i denotes the ith elements in the vector. In this work, One-Hot code is used as the sample label of the neural network. The cross entropy was used as the cost function as given below[25]

$$C(t) = -\frac{1}{n}\sum_{i=1}^{n} y_i \ln a_i + (1 - y_i)\ln(1 - a_i), \tag{8}$$

where y_i denotes the probability vector of a gesture sample, ai denotes the sample label, and n denotes the amount of samples contained in the batch. The gradient descent algorithm was used to minimize the cost function. The output of the neural network is compared with the sample label. If the output matches the sample label, the prediction result is marked as "True;" otherwise, it is marked as "False." By constantly feeding data to the neural network, the cost function eventually reaches a minimum, and weight parameters of the neural network are optimized. Supplementary material Note 5 shows that the error rate decreases with training epochs. The neural network was trained with the training data set containing 5000 samples, and the training accuracy of 100% was obtained. After training, the network can conduct the gesture recognition.

The finger moves along four types of different trajectories on the sensor array, and each trajectory consists of four different sensing units, as shown in Fig. 4(a). To examine the classification/recognition reliability of the e-skin, the PDMS sensor array was flatten on the surface of a desktop, and then, the skin was examined for 50 000 times by randomly selecting the four gestures (each cycle contains 100-time examinations), as shown in Fig. 4(b-i). The recognition accuracy remains above 80%, which means that the classification/recognition of the e-skin system is reliable. To examine the folding effect on the recognition accuracy of the e-skin, the electronic skin was

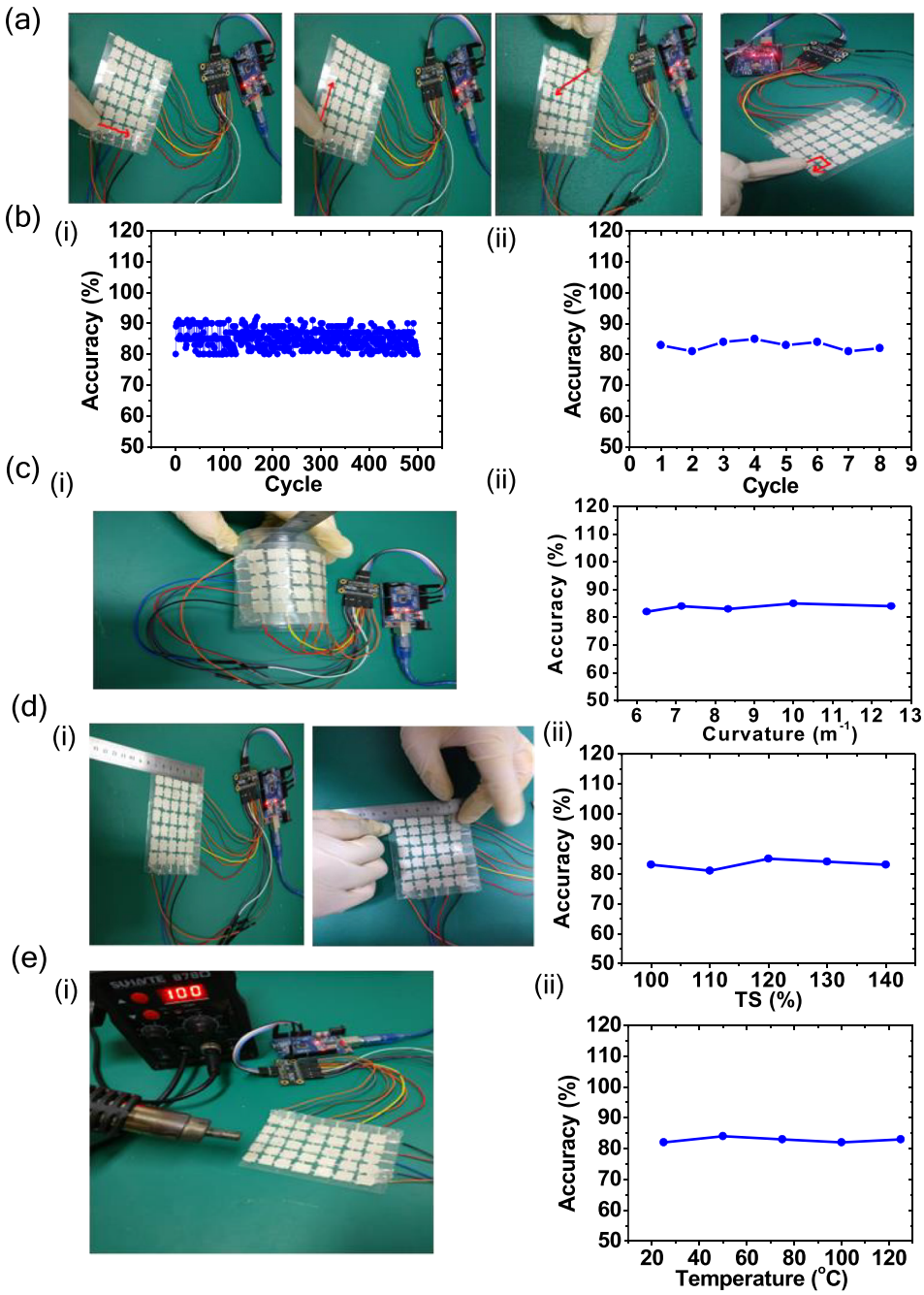

FIG. 4. Effects of cycles, curvature, tensile strength, and temperature on the gesture recognition accuracy rate. (a) The four types of gestures; (b-i) recognition reliability, and (b-ii) effect of folding on the recognition accuracy (iii); (c) effect of curvature on the recognition accuracy; (d) tensile strength effect on the gesture recognition accuracy rate of the e-skin system; and (e) temperature effects on the recognition accuracy rate. Variation of temperature is realized with a hot air gun (the heat source is a Soldering Station Rework 878).

folded for 800 times. After each 100-time folding [i.e., one cycle as shown in Fig. 4(b-ii)], 100-time examinations of recognition by randomly selecting the four gestures were conducted. As shown in Fig. 4(b-ii), the sensor array has a good folding endurance, and the recognition accuracy remains above 80% after 8 cycles, i.e., folding 800 times.

The curvature of the holding substrate of the sensor array shows little impact on the recognition accuracy of the e-skin system. The experimental setup for the curvature test is shown in Fig. 4(c-i). The skin was mounted on the surface of a cylinder with different curvatures of 6.3 m^{-1}, 7.1 m^{-1}, 8.3 m^{-1}, 10.0 m^{-1}, and 12.5 m^{-1}, respectively. At each curvature, the sensor array is examined for 100 times for randomly selected one of the four types of gestures. As shown in Fig. 4(c-ii), the recognition accuracy remains above 80% for all curvature conditions.

The tensile effect on the recognition of the e-skin was also examined as shown in Fig. 4(d-i). Here, the ratio of the stretchable length to the initial length is defined as the tensile strength (TS). The tensile effect on the recognition accuracy was carried out for TS of 100%, 110%, 120%, 130%, and 140%, respectively. The tensile also has little impact on the recognition accuracy, as shown in Fig. 4(d-ii). Note that if the e-skin is further stretched for more than 150%, it is possible to break. The thermal effect on the recognition accuracy was also examined by heating the e-skin with a thermal blower equipped with a temperature indicator as shown in Fig. 4(e-i). Temperature was varied from room temperature of around 23 °C to 120 °C. As can be observed in Fig. 4(e-ii), temperature below 120 °C has little impact on the accuracy of the gesture recognition by the neural network.

This work reports a low-cost e-skin system fabricated with a simple process. The e-skin system consists of the sensor array and data processing system. The sensor array was fabricated with a crossbar architecture, and each sensor unit has a capacitor-like structure with a PDMS layer sandwiched between two conductive silver electrodes. The signal acquisition module collects the data of the capacitance variations of the sensor array. The LSTM neural network algorithm is used to realize gesture recognition. This work provides a possibility to realize a low-cost e-skin with the artificial intelligence capability similar to that of a biological skin.

See supplementary material for the preparation of the PDMS film, the design of signal acquisition circuit, diagram of four types of gestures, the description of the principle of the LSTM neural network, and the relationship between the neural network training error and the epochs.

This work was supported by NSFC under Project Nos. 61774028 and 61771097 and the Fundamental Research Funds for the China Central Universities under Project No. ZYGX2016Z007.

[1]A. Chortos, J. Liu, and Z. N. Bao, Nat. Mater. **15**, 937 (2016).
[2]S. H. Daebritz, J. S. Sachweh, B. Hermanns, B. Fausten, A. Franke, J. Groetzner, B. Klosterhalfen, and B. J. Messmer, Circulation **108**, 134 (2003).
[3]C. Antfolk, M. Alonzo, B. Rosen, G. Lundborg, F. Sebelius, and C. Cipriani, Exp. Rev. Med. Devices **10**, 45 (2013).
[4]L. Gao, D. D. Dong, J. G. He, K. K. Qiao, F. Cao, M. Li, H. Liu, Y. B. Cheng, J. Tang, and H. Song, Appl. Phys. Lett. **105**, 153702 (2014).
[5]K. Takei, T. Takahashi, J. C. Ho, H. H. Ko, A. G. Gillies, P. W. Leu, R. S. Fearing, and A. Javey, Nat. Mater. **9**, 821 (2010).
[6]H. J. Xian, C. R. Cao, J. A. Shi, X. S. Zhu, Y. F. Huang, S. Meng, L. Gu, Y. H. Liu, H. Y. Bai, and W. H. Wang, Appl. Phys. Lett. **111**, 121906 (2017).
[7]H. S. Shin and S. Bergbreiter, Appl. Phys. Lett. **112**, 044101 (2018).
[8]B. C.-K. Tee, C. Wang, R. Allen, and Z. N. Bao, Nat. Nanotechnol. **7**, 825 (2012).
[9]S. Yamada, T. Sato, and H. Toshiyoshi, Appl. Phys. Lett. **110**, 253501 (2017).
[10]C. Larson, B. Peele, S. Li, S. Robinson, M. Totaro, L. B. Mazzolai, and R. Shepherd, Science **351**, 1071 (2016).
[11]H. H. Chou, A. Nguyen, A. Chortos, J. W. F. To, C. Lu, J. Mei, T. Kurosawa, W. G. Bae, J. B.-H. Tok, and Z. N. Bao, Nat. Commun. **6**, 8011 (2015).
[12]K. S. Sohn, J. Chung, M. Y. Cho, S. Timilsina, W. B. Park, M. Pyo, N. Shin, K. Sohn, and J. S. Kim, Sci. Rep. **7**, 11061 (2017).
[13]T. Yang, W. Wang, H. Zhang, X. Li, J. D. Shi, Y. He, Q. S. Zheng, Z. H. Li, and H. Zhu, ACS Nano **9**, 10867 (2015).
[14]H. Wang, IOP Conf. Ser. **207**, 012040 (2017).
[15]Z. Lou, S. Chen, L. Wang, K. Jiang, and G. Z. Shen, Nano Energy **23**, 7 (2016).
[16]S. Tasoglu, E. Diller, S. Guven, M. Sitti, and U. Demirci, Nat. Commun. **5**, 3124 (2014).
[17]S. W. Jeong, J. W. Jeong, S. Chang, S. Y. Kang, K. L. Cho, and B. K. Ju, Appl. Phys. Lett. **97**, 253309 (2010).
[18]N. Yogeswaran, W. Dang, W. T. Navaraj, D. Shakthivel, E. O. Polat, S. Gupta, H. Heidari, M. Kaboli, L. Lorenzelli, G. Cheng *et al.*, Adv. Rob. **29**, 1359 (2015).
[19]I. Aleksander, Nature **416**, 789 (2002).
[20]Y. Cheng, R. Wang, H. Zhai, and J. Sun, Nanoscale **9**, 3834 (2017).
[21]F. Buggenthin, F. Buettner, P. S. Hopppe, M. Endele, M. Kroiss, M. Strasser, M. Schwarzfischer, D. Loeffler, K. D. Kokkaliaris, O. Hilsenbeck *et al.*, Nat. Methods **14**, 403 (2017).
[22]M. I. Jordan and T. M. Mitchell, Science **349**, 255 (2015).
[23]N. Y. Hammerla, S. Halloran, and T. Ploetz, in *IJCAI-16, New York, USA, 9-15 July 2016* (IJCAI, New York, USA, 2016), p. 1533.
[24]F. J. Ordóñez and D. Roggen, Sensors **16**, 115 (2016).
[25]H. White, Neural Comput. **1**, 425 (1989).