Electronic Supplementary Information

Machine learning-guided design and development of multifunctional flexible Ag/poly (amic acid) composites using differential evolution algorithm

Mengyao Zhang, [†a] Jia Li, [†a] Ling Kang, [a] Nan Zhang, [a] Chun Huang, [a] Yaqin He, [a] Menghan Hu, [a] Xiaofeng Zhou [a] and Jian Zhang *[a, b]

[a] Shanghai Key Laboratory of Multidimensional Information Processing, East China Normal University, 500 Dongchuan Road, 200241 Shanghai, China

[b] Shanghai Institute of Intelligent Electronics & Systems, Fudan University, Shanghai 200433, China

*E-mail: jzhang@ee.ecnu.edu.cn; jzhang@ce.ecnu.edu.cn

S1. SMIE process of Ag/PAA composites.

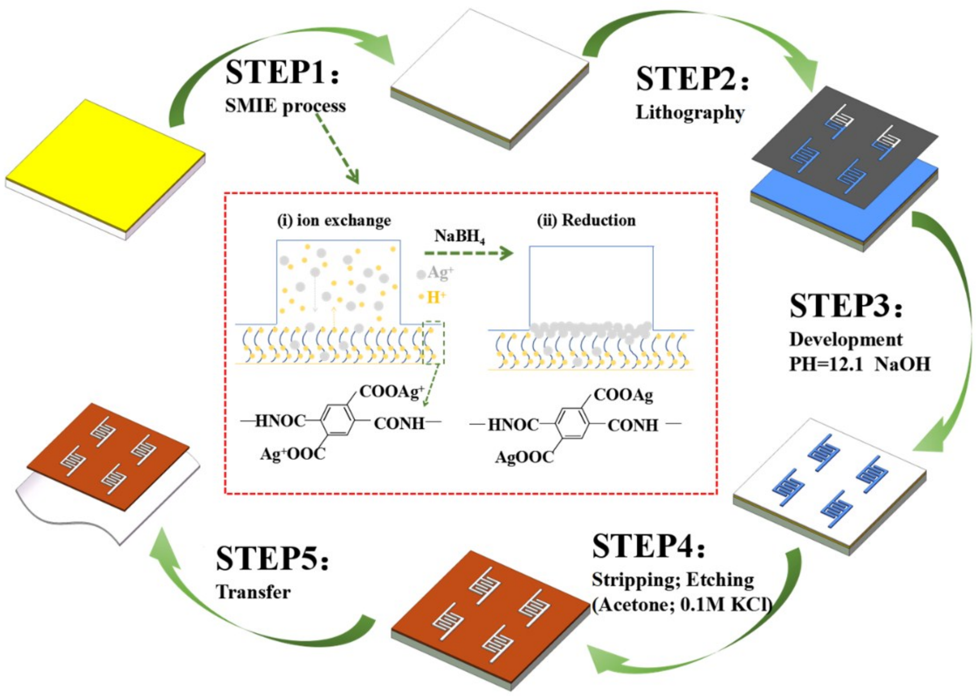


Figure S1. SMIE process of Ag/PAA composites. Step 1: SMIE for Ag deposition. Step 2: Dry film exposure. Step 3: Dry film development. Step 4: Dry film stripping & Ag etching. Step 5: Transfer to arbitrary substrate.

Figure S1 shows the schematic diagram of SMIE of Ag/PAA composite film in detailed. Step 1 is the process of SMIE. First, the PAA layer is spin-coated on PI-attached glass substrate, and then immersed in the $AgNO_3$ solution. After a certain period of time, Ag is replaced with Ag^+ by a reducing agent to obtain a metal layer, its principle is the process of in situ growth of silver nanoparticles, including the exchange between H^+ and Ag^+ in PAA, and the Ag^+ are reduced into Ag by a reducing agent. Step 2 is to attach the photosensitive dry film over the metal layer, and photolithography is performed through the mask. Step 3 is to obtain the desired pattern by the developer with NaOH and etching with 0.1 M KCl. Step 4 removes the photosensitive dry film by acetone. Step 5 is to peel the patterned composite film from the glass substrate and attach it to arbitrary substrate. The obtained Ag/PAA composites is peeled off from the glass substrate. The Ag/PAA composites is so flexible and can be attached to any substrate conformably, as shown in Figure 5.

S2. Conventional Process Optimization Difficulty

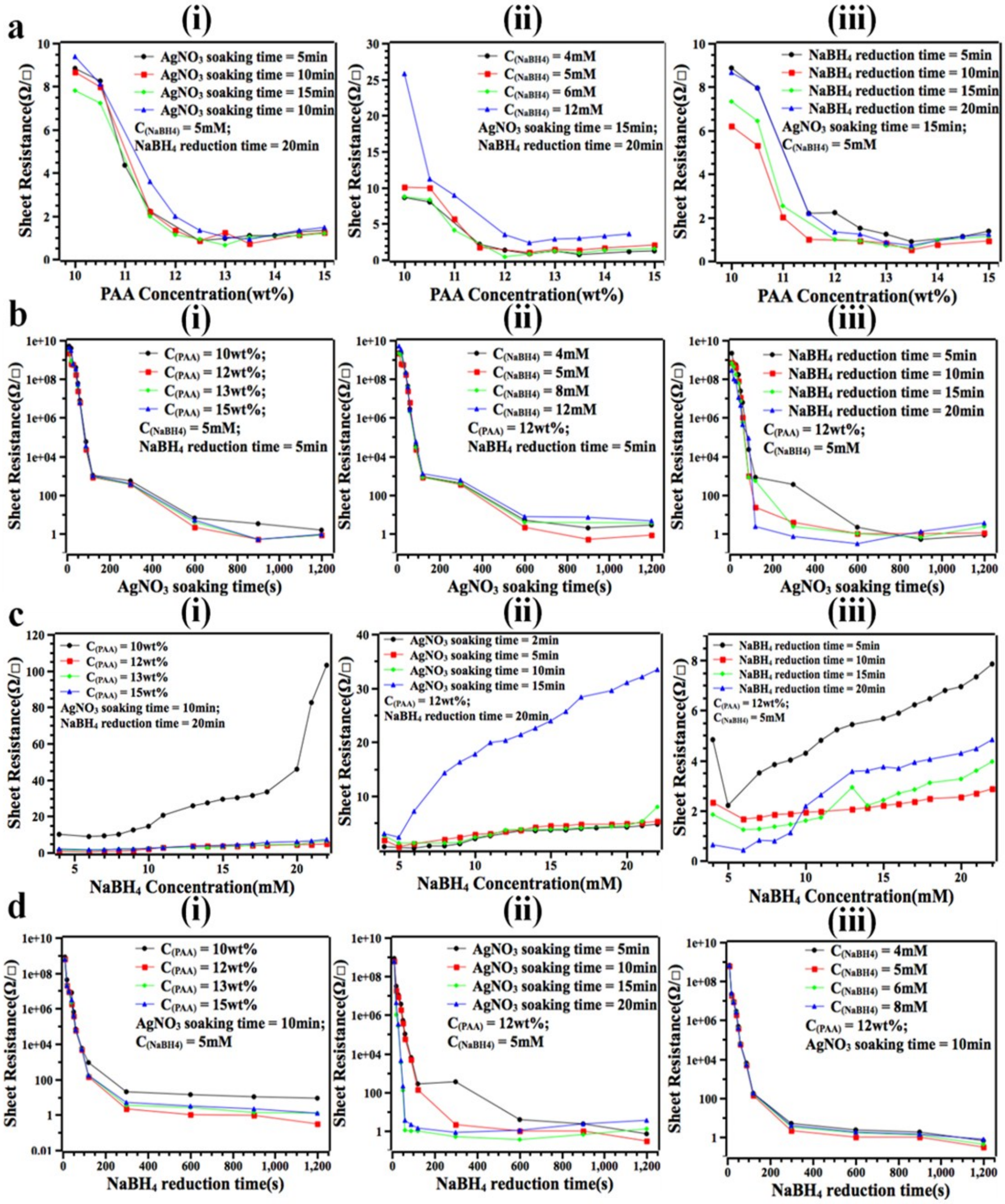


Figure S2. Comparison of factors affecting Ag/PAA composites by orthogonal analysis. a(i)-a(iii) Change $AgNO_3$ soaking time, $C_{(NaBH4)}$ and $NaBH_4$ reduction time in turn, $C_{(PAA)}$ and sheet resistance change relationship. b(i)-b(iii) Change $C_{(PAA)}$, $C_{(NaBH4)}$ and $NaBH_4$ reduction time in turn, $AgNO_3$ soaking time and sheet resistance change relationship. c(i)-c(iii) Change $C_{(PAA)}$, $AgNO_3$ soaking time and $NaBH_4$ reduction time in turn, $C_{(NaBH4)}$ and sheet resistance change relationship. d(i)-d(iii) Change $C_{(PAA)}$, $AgNO_3$ soaking time, $C_{(NaBH4)}$ in turn, $NaBH_4$ reduction time and sheet resistance relationship.

We first use the orthogonal analysis to analyze the influence of single variable. For example, from Figure S2a(i), with the $NaBH_4$ concentration(5 mM) and the reduction time(20 min) fixed, under different the $AgNO_3$ soaking time(i.e. $AgNO_3$ ion exchange time), the relationship of the PAA concentration vs the sheet resistance of Ag/PAA composites was studied. It observed that as the PAA concentration increased, the sheet resistance decreased sharply first, then tended to be flat. Figure S2a(ii) and S2a(iii) showed the influence of $NaBH_4$ concentration and reduction time, respectively. Although these curves share the similar trend, due to changes in other manufacturing factors, the differences of the curves (slope, extreme point, etc.) is obvious. Following the similar procedure, we can analyzed Figure S2b(i)-S2b(iii), S2c(i)- S2c(iii), and S2d(i)-S2d(iii). Therefore, in case of four variable condition, it is difficult to obtain an optimal judgment condition in a continuous domain since 12 kinds of variable conditions existing.

S3. The implementation of the DE algorithm

The DE algorithm is essentially a multi-objective optimization algorithm, which is used to solve the global optimal solution in multi-dimensional space. Figure S3 is a whole flow chart of BP neural network based on DE algorithm. The input data of the model in this study is (x_1, x_2, x_3, x_4), and the details of implementation of the DE algorithm are as following[1]:

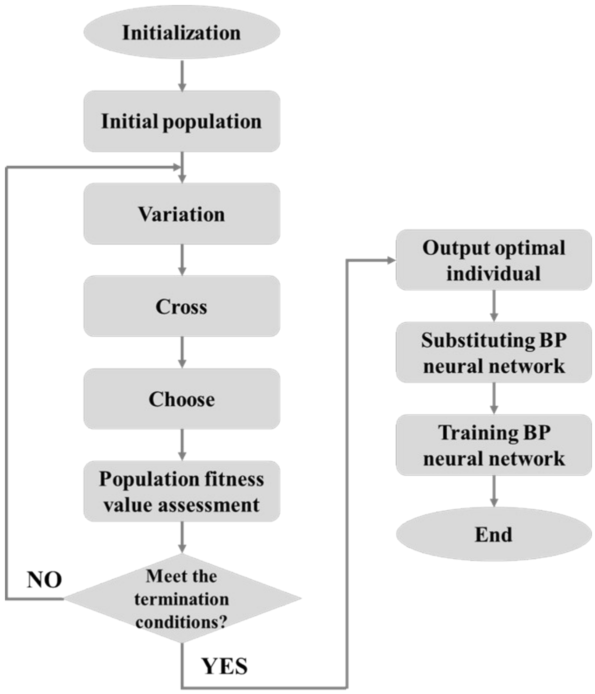


Figure S3. BP neural network flow based on DE algorithm.

Population initialization. Initializing the population operation means generating NP D-dimensional individuals to form the original population, and the j-th element of the i-th individual of the original population is generated as shown in equation (1).

$$x_{i,0}^{j} = x_{min}^{j} + rand(0,1) \cdot (x_{max}^{j} - x_{min}^{j}) \tag{1}$$

The random number generation function $rand(0,1)$ indicates that a uniformly distributed random number is generated between (0, 1), x_{max} and x_{min} respectively represent the maximum and minimum values of the constraint search space. Where $i = 1,2,\cdots NP$, $j = 1,2,\cdots D$.

Population fitness assessment. All individuals are sequentially substituted into the test function, and the fitness value $f[X_i(G)]$ of each target vector $X_i(G)$ is calculated.

Mutation operation. The mutation operation is the core of differential evolution. The mutation vector is composed of the vector scaling difference between individuals in the population and other different individuals. It is mainly divided into two parts: the base vector and the difference vector. According to different methods of generating the variation vector, multiple mutation strategies can be obtained. The most common mutation strategy is "DE/rand/1":

$$V_{i,G} = X_{r_1^i,G} + F \cdot (X_{r_2^i,G} - X_{r_3^i,G}) \tag{2}$$

The random indices r_1^i, r_2^i, and r_3^i are integers ranging from $[1,NP]$, which are different from each other. The scaling factor F is a positive real number that controls the rate at which the population develops, with no upper limit.

Cross operation. The crossover operation is to cross-mix the mutated vector with the target vector to obtain a test vector. The parameter CR is used as the crossover probability to control the probability that the test vector consists of elements of the mutated vector. The specific execution method is as follows:

$$u_{i,j,G} = \begin{cases} v_{i,j,G} & rand(0,1) \le CR \ \ or \ \ j = j_{rand} \\ x_{i,j,G} & others \end{cases} \tag{3}$$

Where j_{rand} is a random integer in [1, D], and condition $j = j_{rand}$ is to ensure that there must be elements from the mutation vector in the test vector, to avoid the stagnation of the algorithm while ensuring that the crossover operation is effective.

Select an action. The DE algorithm uses a "greedy" choice strategy. First, the fitness values of all newly generated test vectors are calculated one by one, and then the fitness values of the test vectors are compared with the fitness values of the target vectors one by one. If the test vectors have lower objective function values, the test is performed. The vector replaces this target vector into the next generation, otherwise the target vector remains. The specific operation formula is as follows:

$$X_{i,G+1} = \begin{cases} U_{i,G} & f(U_{i,G}) \le f(X_{i,G}) \\ X_{i,G} & others \end{cases} \tag{4}$$

The initial weight and the threshold of neural network are adjusted by the DE algorithm. Therefore, the decision variable of DE algorithm is the initial weight and the threshold of BP, and the range of each element is [-1,1].

In Matlab R2013a, the initial populations, the new individuals through mutation, the crossover and the selection operations are randomly selected to generate NP individuals. These new individuals are the weights and the thresholds of neural network, which can assign the optimal weights and the thresholds to neural network.

And the final result was calculated following the model. The sum of squared errors (MSE) between the predicted value and the true value is used as the fitness function of the population.[2] The expression is:

$$MSE = \frac{\sum_{t=1}^{k} (y_{tc} - y_{tt})^2}{k} \tag{5}$$

Here, $y_{tt}$is the actual value, y_{tc} is the predicted value obtained by the BP neural network, and k is the number of data in the training set, respectively.

S4. The choice of DE-BP parameters

(a)

Parameters of BP neural network	**Value**
Transfer function	tan sig
Maximum number of trainings	1000
Training target error	1e-5
Learning step	0.001
Maximum number of iterations G	50

(b)

Parameters of DE algorithm	**Value**
Population size NP	60
Dimension D	31
Scaling factor F	0.5
Cross probability CR	0.9

Figure S4. The choice of (a) BP neural network parameters and (b) DE algorithm parameters.

In general, the control parameters of BP neural network include the transfer function, the maximum number of trainings, the training target error, the learning step, the maximum number of iterations, etc. And the control parameters of the DE algorithm[3] include the population size (NP), the dimension (D), the scaling factor (F), the cross probability (CR), etc. In this paper, the transfer function of the hidden layer neurons selects the tan sig function : $f_h(x) = \frac{2}{1 + e^{-2x}} - 1$.[4] The maximum number of training is generally set to 1000 times[5] and the maximum number of iterations G is set to 50. The neural network calculation cannot always guarantee the convergence of the iteration results under various parameter configuration. When the iteration result does not converge, the maximum number of iterations is fixed at 50. The training target error is generally set to 1e-3~1e-5, which is set to 1e-5 in this work. When the error of two adjacent iteration results is less than this value, the system ends the iterative calculation and gives the results. The learning step is set to 0.001. When the learning step is too large, it maybe causes the oscillation and non-convergence. On the contrary, if this parameter is too small, the convergence speed will be lowered. The population size reflects the amount of population information in the algorithm, which is generally set to 30~160. And it is set to 60 in our work. The larger the population size, the richer the population information. However, the excessive population size will result in complicated calculation and low efficiency. The dimension is set to 31, which could guarantee the relative error within 2% in our work, and further increasing the complexity of the model will make the calculation time longer. The scaling factor represents the global optimization capability of the algorithm, which is set to 0.5 in this work. The smaller scaling factor, the better the local search ability. The larger scaling factor will make the algorithm escape the local extreme point easily, but the rate of convergence is very slow. The cross probability indicates the magnitude of the information exchanged between the parent, the offspring, and the intermediate mutant during the cross process. The greater cross probability, the greater the information exchanged. The cross probability is generally set to 0.4~0.97. [6,7] In our work, it is set to 0.9.

S5. The block diagram of BP neural network learning algorithm

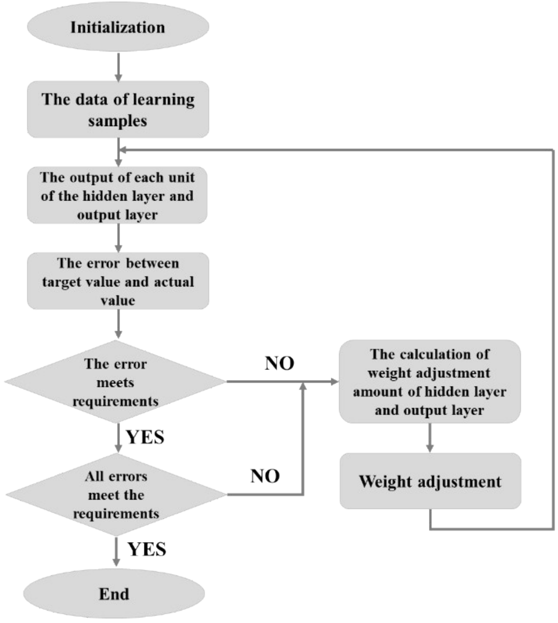


Figure S5. The block diagram of BP neural network learning algorithm.

S6. Ag/PAA morphology under SEM

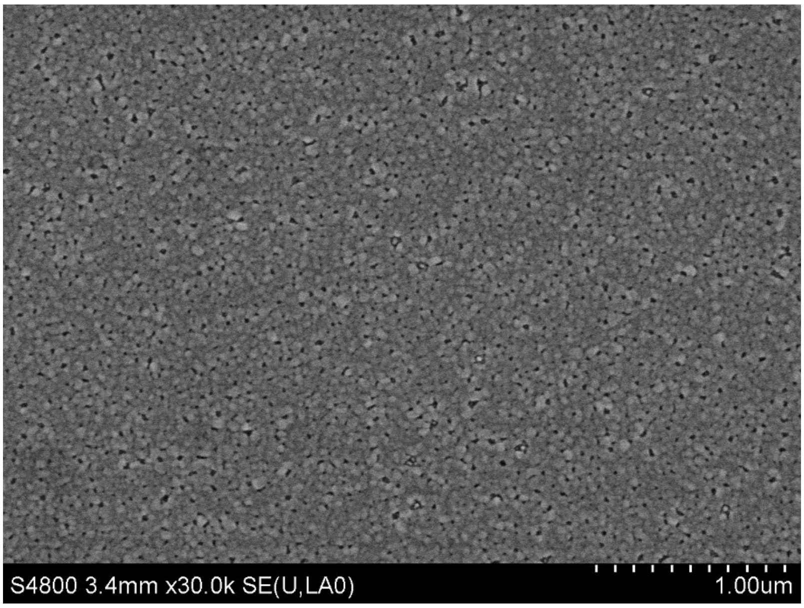


Figure S6. Ag/PAA morphology under SEM.

Concentration of PAA is 11wt %, reaction time of $AgNO_3$ is 10 min, concentration of $NaBH_4$ is 5 mM, reaction time of $NaBH_4$ is 20 min. Obtained by SEM, the Ag grown is more uniform and dense.

S7. Gauge Factor calculation

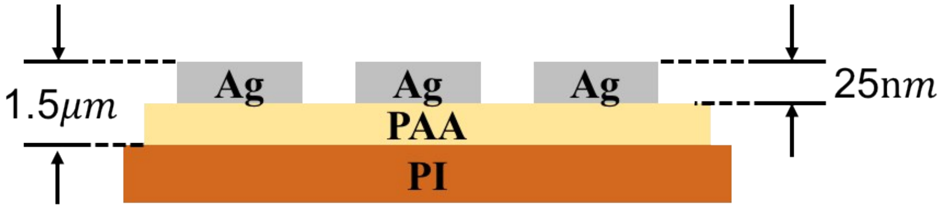
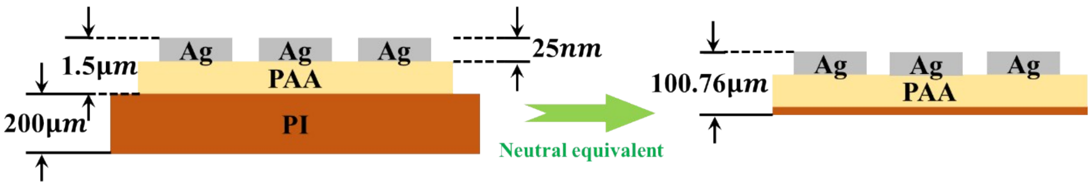


Figure S7. Cross-sectional schematic view of preparing patterned Ag/PAA composite film on PI substrate. A schematic cross-sectional view of a patterned PAA-Ag composite film prepared on a PI substrate, wherein SEM measured Ag/PAA is about 1.5 um thick, and upper layer Ag is about 25 nm.

(a)



(b)

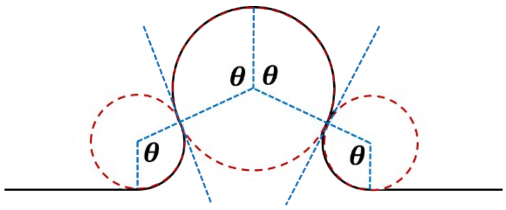


Figure S7. (a) The neutral equivalent diagram. (b) Schematic diagram of the deformation relationship of the film in the process of bending deformation of Ag/PAA composites.

In order to judge the sensitivity of the Ag/PAA composite film used as a pressure sensor, we use the neutral plane[8,9] of the composite film as a reference, and its deformation equivalent diagram is shown in Figure 7a during bending. Therefore, we can establish the following geometric relationships.

$$\frac{2\theta\pi R}{180°}+\frac{2\theta\pi r}{180°}=\varepsilon_0 \tag{1}$$

where, set the original length L_0, deformation length of Ag/PAA composite film after deformation $\varepsilon^{'}$

$$\frac{2\theta\pi(R+d)}{180°}+\frac{2\theta\pi(r+d)}{180°}=\varepsilon^{'} \tag{2}$$

where

$$GF = \frac{\Delta R}{R} \Big/ \frac{\Delta\varepsilon}{\varepsilon},$$ Available from (1)(2),

$$GF = \frac{45°\Delta R}{\theta\pi d R} \tag{3}$$

Therefore, we have calculated that the sensitivity Gauge Factor GF of the strain sensor is calculated to be about 8.77, which is larger than that of the ordinary bulk metal material, which further confirms our idea that the obtained film is the Ag/PAA chimeric structure.

S8. The comparison of GF

way	Strain (%)								
	0.95	2.76	2.90	3.71	5.16	5.65	6.36	7.42	8.30
CF_1	1.63	1.84	2.12	2.54	3.24	3.94	4.64	6.46	6.95
CD_1	1.67	1.45	2.48	3.23	3.87	4.94	5.64	7.63	7.49
CF_2	1.20	3.14	4.54	5.96	7.29	8.64	9.56	9.96	10.46
CD_2	0.98	4.15	6.47	7.63	8.74	9.19	9.30	9.80	9.90

Figure S8. The comparison of GF. CF_1: Curve fitting of sample 1. CD_1: Calculation derivation of sample 1. CF_2: Curve fitting of sample 2. CD_2: Calculation derivation of sample 2.

To verify the plausibility of the fitting curve, we compared the GF values calculated from the fitting curves with the ones from the mechanic theory of plate and shell. It is easy to see from the Table that the GF is roughly equivalent, and the deviation is mainly due to the systematic error and instrumentation error in the measurement. The difference in GF between sample 1 and sample 2 is due to the difference in initial resistance and measurement error during subsequent testing.

S9. Strain gauge sensor process parameters

(a)

1% strain		
Independent Variable	DE-BP optimization	Orthogonal analysis
Concentration of PAA (wt%)	10.5	12
Reaction time of $AgNO_3$ (min)	10	10
Concentration of $NaBH_4$ (mM)	5	5
Reaction time of $NaBH_4$ (min)	15	15

(b)

3% strain		
Independent Variable	DE-BP optimization	Orthogonal analysis
Concentration of PAA (wt%)	10.5	11.5
Reaction time of $AgNO_3$ (min)	7	10
Concentration of $NaBH_4$ (mM)	5	5
Reaction time of $NaBH_4$ (min)	18	18

(c)

5% strain		
Independent Variable	DE-BP optimization	Orthogonal analysis
Concentration of PAA (wt%)	10.5	11.5
Reaction time of $AgNO_3$ (min)	8	3
Concentration of $NaBH_4$ (mM)	8	12

Reaction time of $NaBH_4$ (min)	20	20

(d)

7% strain		
Independent Variable	DE-BP optimization	Orthogonal analysis
Concentration of PAA (wt%)	11	11.5
Reaction time of $AgNO_3$ (min)	10	5
Concentration of $NaBH_4$ (mM)	5	18
Reaction time of $NaBH_4$ (min)	20	10

(e)

Stability test		
Independent Variable	DE-BP optimization	Orthogonal analysis
Concentration of PAA (wt%)	11	12
Reaction time of $AgNO_3$ (min)	10	5
Concentration of $NaBH_4$ (mM)	6	20
Reaction time of $NaBH_4$ (min)	20	10

Figure S9. The process parameters of strain gauge sensors with (a) 1% strain, (b) 3% strain, (c) 5% strain, (d) 7% strain and (e) stability test.

Figure S9 compares two different process parameters for DE-BP optimization and orthogonal analysis. The un-optimized parameters achieved from 1077 learning samples, are the result of orthogonal analysis. It can be found that the BP can predict the characteristics of Ag/PAA composite effectively and precisely. The experimental data of the stability test is consistent with the source of the bending deformation test.

S10. The 4×4 capacitive sensor array

A 4×4 pressure capacitive sensor array (width 500 μm) was demonstrated. Figure S10a showed the cross sectional diagram of the capacitive sensor array and its corresponding dimensions. The sensor is mainly composed of two identical Ag/PAA electrodes on the PI substrate and the dry films as the insulation layer. This sandwich-like structure was encapsulated by the ecoflex0030 layers. Two Ag/PAA electrodes are fabricated with lithography process, with the linewidth of 500 μm. The dry films was bonded and patterned to form the insulation layer. The direction of the upper and lower electrodes is perpendicular to each other, forming a cross capacitor structure. The sensor size and the electrode dimension are 1 cm × 1 cm and 5 mm × 5 mm, respectively. The air gap between two plate electrodes is ~ 48 μm. The initial capacitance of a single cell has been estimated as 4.61 pF. When pressure is applied to a bump, the upper ecoflex0030 deforms and capacitance increases until the air gap is completely closed. Therefore, the total thickness of the upper electrode and bump layer determines the sensitivity. The capacitive sensor array was shown in Figure S7b.

The testing circuit will be shown in the next work, and the sensor arrays could be used to map the pressure distribution[10], which is potential for the tactile sensor, e-skin, etc.

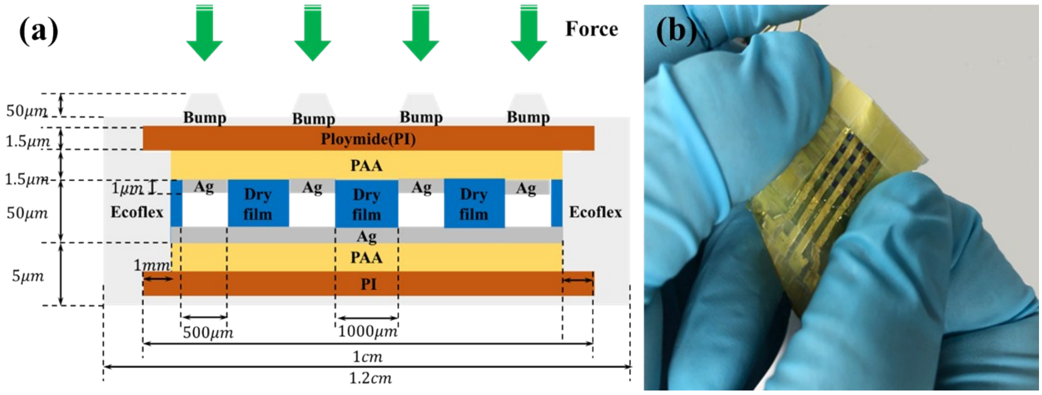


Figure S10. (a) Cross-sectional view and dimensions of the capacitive sensor array. (b) The capacitive sensor array.

References

1 R. M. Gilberto, S. Javier, B. Xavior and M. H. Juan, *Evol. Comput., 2011 IEEE Congress on IEEE*, 2011, **47(10)**:1551-1556.

2 J. I. Mwasiagi, *Fash. Text.*, 2016, **3**: 23.

3 J. Zhang and Z. Dong, *Algorithms*, 2019,12,71.

4 M. Alas and *S. I. A. Ali*, *Int. J. Technol.*, 2019, **10**:417-427.

5 S. Y. Liu, L. D. Zhang, Q. Wang and J. J. Liu, *International Conference on Wavelet Analysis & Pattern Recognition IEEE*, 2008, 216-220.

6 J. Brest, S. Greiner, B. Boskovic, M. Mernik and V. Zumer, *IEEE T. Evolut. Comput.*, 2006, **10(6)**:646-657.

7 X. Pan, W. Zhou, Y. Lu and N. Sun, *IEEE ACCESS*, 2019, 55807-55816.

8 J. Y. Zhang, W. Z. Lu, R. Huo and R. Feng, *Build. Environ.*, 2008, **43**, 1101-1108.

9 M. Sanjari, A. Karimi Taheri and A. Ghaei, *J. Mater. Process. Tech.*, 2007, **186**, 147-153.

10 H. K. Lee, S. I. Chang and E. Yoon, *J. Microelectromech. S.*, 2006, **15**, 1681-1686.