



# Multimodal Surface Sensing based on Hybrid Flexible Triboelectric and Piezoresistive Sensor

Zenan Lin\*  
Kai Chong Lei\*  
{lzn21,liqc21}@mails.tsinghua.edu.cn  
Tsinghua-Berkeley Shenzhen Institute  
Shenzhen International Graduate  
School  
Tsinghua University  
China

Shilong Mu  
Ziwu Song  
{msl22,song-  
zw20}@mails.tsinghua.edu.cn  
Tsinghua-Berkeley Shenzhen Institute  
Shenzhen International Graduate  
School  
Tsinghua University  
China

Yuan Dai  
jessiedai@tencent.com  
Tencent Robotics X Lab  
China

Wenbo Ding†  
ding.wenbo@sz.tsinghua.edu.cn  
Tsinghua-Berkeley Shenzhen Institute  
Shenzhen International Graduate  
School  
Tsinghua University  
RISC-V International Open Source  
Laboratory  
China

Xiao-Ping Zhang  
xiaoping.zhang@sz.tsinghua.edu.cn  
Shenzhen International Graduate  
School  
Tsinghua University  
Department of Electrical, Computer  
and Biomedical Engineering  
Ryerson University

## ABSTRACT

Sensing the surface properties through touch, as the most natural perceptual way of humans, has become an important and practical method for human-machine interactions (HMI) and robot manipulations. In this paper, we design a fingertip hybrid flexible tactile sensor for multimodal surface sensing, based on the triboelectric and piezoresistive mechanisms. A real-time tactile sensing system is implemented on a 3D-printed robot finger together with a wireless data acquisition board. A virtual data generation method is proposed to expand the model adaptability under different compression force levels. Moreover, considering the characteristics of data generated by our sensors, a novel deep learning model with a residual structure is developed, named parallel residual convolutional neural network (PR-CNN). Our model outperforms the state-of-the-art models, i.e., Res-CNN, LSTM-FCN and InceptionTime, with over 96% accuracy, on three classification tasks, including textures (13 types), materials (10 types), and combinations of textures and materials (18 types). The proposed system has broad applications in service robots, industrial sorting robots, and HMI.

\*Both authors contributed equally to this research.

†Wenbo Ding is the corresponding author.



This work is licensed under a [Creative Commons Attribution International 4.0 License](https://creativecommons.org/licenses/by/4.0/).

UbiComp/ISWC '22 Adjunct, September 11–15, 2022, Cambridge, United Kingdom  
© 2022 Copyright held by the owner/author(s).  
ACM ISBN 978-1-4503-9423-9/22/09.  
<https://doi.org/10.1145/3544793.3560404>

## CCS CONCEPTS

• **Human-centered computing** → Ubiquitous and mobile computing; • **Computing methodologies** → Machine learning; • **Computer systems organization** → Embedded and cyber-physical systems.

## KEYWORDS

Multimodal Tactile Sensing, Texture and Material Recognition, Flexible Sensors, Deep Learning Methods.

## ACM Reference Format:

Zenan Lin, Kai Chong Lei, Shilong Mu, Ziwu Song, Yuan Dai, Wenbo Ding, and Xiao-Ping Zhang. 2022. Multimodal Surface Sensing based on Hybrid Flexible Triboelectric and Piezoresistive Sensor. In *Proceedings of the 2022 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp/ISWC '22 Adjunct)*, September 11–15, 2022, Cambridge, United Kingdom. ACM, New York, NY, USA, 6 pages. <https://doi.org/10.1145/3544793.3560404>

## 1 INTRODUCTION

Inferring the surface properties of the target objects and realizing recognition through touch are of both practical meaning and irreplaceable importance for human-machine interactions (HMI) and precise robot manipulations [3]. Especially, in the last decade, the realization of the emerging electronic-skin concept has called for the advances of flexible tactile sensors towards better sensing capabilities and complex surface sensing tasks [2].

To this end, different types of flexible tactile sensors have been developed for surface property sensing, including pressure, strain, texture, materials and so on [15]. Usually, the sensors that detect the forces could be utilized for pressure, strain, and texture perception, such as the piezoelectric sensor for roughness sensing [11], the

piezoresistive sensor for texture recognition [13], the capacitive [1] or optical [19] sensor for strain detection, etc. Besides, by deploying the sensors in an array, the sensing capabilities including the range and resolution could be dramatically enhanced [18]. Nevertheless, the material recognition task still remains a challenge for such force-oriented tactile sensors due to its nature. The emerging triboelectric sensor has shown its potential for material recognition thanks to its unique working mechanism [8]. Recently, we also proposed a wavelet based signal decoupling method to make the single triboelectric sensor sense more than one surface property simultaneously [14]. Unfortunately, it only works well at a single compression force level which limits its applications.

Considering the sensing limitations of the single-mode sensor as well as the nature of human perception, the multimodal sensor-based tactile sensing systems have demonstrated superiority in complicated tasks such as object recognition as well as material and texture classifications with high accuracy and efficiency [5]. In Komeno's work, a vibration and tactile sensor were adopted for object surface classification by injecting mechanical vibrations during contact [10]. Pastor *et al.* proposed an object recognition pipeline fed with haptic images and angle values under the Bayesian fusion framework, which outperforms the approach based on neural networks [12]. There are already some great works either on the multimodal device fabrication [6] or on the data fusion algorithm development [20]. It is worth noting that designing the multimodal tactile sensor together with the customized algorithms in a quantitative and systematic way is still of great interest and significance.

In this paper, we propose a low-cost and easily-assembled tactile sensing system for multimodal surface sensing tasks. The sensing system consists of a novel fingertip flexible hybrid tactile receptor based on triboelectric and piezoresistive mechanisms and a wireless signal acquisition board collecting sensors' output. For the time-series data generated by the sensors, we develop the customized deep learning (DL) algorithm and complete three classification tasks, including textures (13 types), materials (10 types), and combinations of textures and materials (18 types), with over 96% accuracy. The contributions of this work can be trifold as follows:

- We incorporate the piezoresistive thin film into the grating-structured triboelectric sensor and form a sandwich like flexible tactile sensor for multimodal surface sensing. A real-time system is implemented on a robot manipulator.
- A virtual data generation method is proposed to expand the adaptability of the DL model and the robustness of the system under different compression force levels.
- In addition, we develop a DL model with a residual structure named parallel residual convolutional neural network (PR-CNN), which has fewer model parameters and outperforms the state-of-the-art (SOTA) models.

## 2 SYSTEM DESIGN

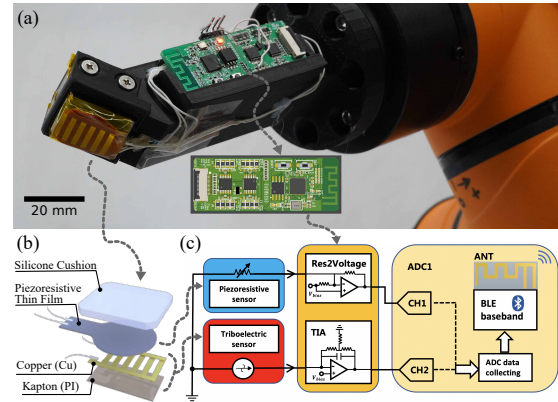
### 2.1 Working Mechanism

In this paper, we develop a multimodal tactile sensor mainly based on two working mechanisms, i.e., the triboelectric and piezoresistive working mechanisms, briefly introduced as follows.

The *triboelectric sensor* is formed by the emerging triboelectric nanogenerator (TENG), of which the electric output relies on

the coupling effect of contact electrification and electrostatic induction [17]. It can distinguish materials from the difference in electronegativity. The triboelectric sensor can distinguish materials under a small compression force range [14]. However, when applying multiple levels of compression force, it's not easy to distinguish materials of similar electronegativity with only triboelectric sensing data. Due to its low-cost and flexibility as well as the broad availability of materials, the TENG based tactile sensor has demonstrated great potential in realizing robot sensing and HMI [4].

The *piezoresistive sensor* is based on the piezoresistive effect which is the change in electrical resistance of materials caused by stretching or compression.



**Figure 1: (a) Prototype of the multimodal tactile sensing system; (b) Layered illustration of the proposed fingertip hybrid sensor; (c) Schematic diagram of the wireless data acquisition board.**

### 2.2 Multimodal Tactile Sensor

As shown in Fig. 1a, our multimodal tactile sensing system consists of two modules, i.e., the fingertip flexible tactile sensor and the wireless data acquisition board, which are equipped on a 3D printed finger attached to the robot arm.

The proposed flexible tactile sensor has a sandwich-like structure of four layers with a size of 20 mm by 23 mm and a thickness of 3 mm, as depicted in Fig. 1b. The top layer is a soft silicone cushion as the compression force buffer. The second layer is a commercial piezoresistive thin film (FSR402) which can detect the compression force. The third layer is the six interdigitated (comb-like) copper electrodes connected to two bared pads fabricated via the flexible printed circuit board process. The bottom layer is Kapton (PI) film and works as both the triboelectric layer and the protective encapsulation for the whole sensor. The comb-like copper electrodes together with the Kapton triboelectric layer formed a freestanding mode triboelectric tactile sensor. The electrode parameters have been properly chosen to enhance the signal-to-noise ratio of the output signal as well as the texture resolution.

In this context, by introducing the piezoresistive layer, the pressure could be sensed and monitored independently, which overcomes the limitation of controlling a single compression force level

when only applying free-standing triboelectric sensor and improves the sensing robustness of the system.

### 2.3 Wireless Data Acquisition Board

The schematic diagram of the wireless data acquisition board is illustrated in Fig. 1c. The signal acquisition circuit is designed to convert the current signal of the triboelectric sensor and the resistance value of the piezoresistive sensor to analog voltages (0–3.3 V). It includes a trans-impedance amplifier and a non-inverting proportional amplifier corresponding to the two sensors' detection circuits. Considering that the current signal generated by the triboelectric sensor is weak and the output impedance is large, we select ADA4505 with high impedance and low bias current (typical value 0.5 pA) in our design. In the resistance variation detecting circuit, we choose a high-precision operational amplifier RS8524 and place a 1 k $\Omega$  charging resistor in the inverting input to have a quick response to the changing resistance.

To convert two channels of analog voltages from the signal acquisition circuit into digital form and send real-time signals with corresponding timestamps to the server, the microcontroller unit (ESP32-D0WD) is selected, which has an internal 12-bit successive approximation analog-to-digital converter peripherals and communication ability via Bluetooth. Besides, the signals from the triboelectric and piezoresistive sensors are filtered through both hardware and software methods.

## 3 METHODS

In this section, we introduce three features from the signal of our multimodal tactile sensor and propose virtual data generation as an efficient data argumentation method based on the physical meaning of the features as well as a DL model of our system on textures and materials recognition.

### 3.1 Multimodal Feature Selection

In Wei's work, they select four features from hybrid e-skin's sensing data including voltage amplitude, polarity, signal duration of triboelectric sensor and current amplitude of piezoresistive sensor [16]. Considering our tasks and system design, we choose three representative features ( $f_1, f_2, f_3$ ) among the output of our tactile sensors. These features of the sample represent the current intensity, the quantity of transferred charge and the compression force in the fingertip sliding process, respectively. For every sliding process movement, the proposed system generates a sample of length  $m$ .

Every sample can be described as a multivariate time-series  $\mathbf{X} = \langle \mathbf{X}_{TE}, \mathbf{X}_{PR} \rangle$ , where  $\mathbf{X}_{TE} \in \mathbb{R}^m$  and  $\mathbf{X}_{PR} \in \mathbb{R}^m$  represent the triboelectric and piezoresistive sensing data, respectively. Each sample has a label  $Y$  which can represent the object's material or texture. The three interested features of sample  $\mathbf{X}$  are calculated as:

$$f_1(\mathbf{X}) = \frac{1}{m} \cdot \sum_{t=1}^m |\mathbf{X}_{TE}(t)|, \quad (1)$$

$$f_2(\mathbf{X}) = \max_t \left( \sum_{i=1}^t \mathbf{X}_{TE}(i) \right) - \min_t \left( \sum_{i=1}^t \mathbf{X}_{TE}(i) \right), \quad (2)$$

$$f_3(\mathbf{X}) = \max_t (\mathbf{X}_{PR}(t)). \quad (3)$$

### 3.2 Virtual Data Generation Method

To enhance the robustness of the system, we implement a virtual data generation method in the data pre-processing. Our method is aimed to improve the generalization ability of the DL model by generating the datasets of unknown compression force levels from the datasets of known force levels. We collect data of compression force levels  $\Delta_{\text{real}} = \{1 \text{ N}, 10 \text{ N}, 15 \text{ N}, 25 \text{ N}\}$  and generate data of target compression force levels  $\Delta_{\text{tar}} = \{5 \text{ N}, 20 \text{ N}\}$ . The elements of  $\Delta_{\text{real}}$  and  $\Delta_{\text{tar}}$  are denoted as  $\delta_{\text{real}}$  and  $\delta_{\text{tar}}$ , respectively. The dataset with  $n$  samples of specific compression force level is defined as  $D_{Y,\delta} = \{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n\}$ , where  $\delta \in \Delta_{\text{real}} \cup \Delta_{\text{tar}}$ . The interval of sliding on the surface is positively correlated to the compression force levels. Thus, every target compression force level needs two adjacent compression force levels  $\delta_i, \delta_j$  as:

$$\delta_i = \max(\{\delta \mid \delta \in \Delta_{\text{real}} \cap \delta < \delta_{\text{tar}}\}) \quad (4)$$

$$\delta_j = \min(\{\delta \mid \delta \in \Delta_{\text{real}} \cap \delta > \delta_{\text{tar}}\}) \quad (5)$$

$D_{Y,\Delta_{\text{adj}}}$ , two datasets of  $\Delta_{\text{adj}} = \{\delta_i, \delta_j\}$ , contribute to  $D_{Y,\delta_{\text{tar}}}$  equally. Therefore,  $D_{Y,\delta_{\text{tar}}}$  is generated from  $D_{Y,\delta_i}$  and  $D_{Y,\delta_j}$  with the same amounts of samples.

---

#### Algorithm 1 Virtual Data Generation Method

---

**Input:** Target force levels  $\Delta_{\text{tar}}$ , known datasets  $D_{Y,\Delta_{\text{real}}}$ ;

**Output:** Generated dataset  $D_{Y,\Delta_{\text{tar}}}$ ;

```

1: for all  $\delta_{\text{tar}} \in \Delta_{\text{tar}}$  do
2:    $\Delta_{\text{adj}} \leftarrow \text{get\_adjacent}(\delta_{\text{tar}}, \Delta_{\text{real}})$  by Eqn. (4) and (5);
3:   for all  $\delta_{\text{adj}} \in \Delta_{\text{adj}}$  do
4:     for all  $\mathbf{X}^{\delta_{\text{adj}}} \subset D_{Y,\delta_{\text{adj}}}$  do
5:       • Calculate  $f_1(\mathbf{X}^{\delta_{\text{adj}}})$  and  $f_3(\mathbf{X}^{\delta_{\text{adj}}})$  from  $\mathbf{X}^{\delta_{\text{adj}}}$  based on
         Eqn. (1) and (3);
6:     end for
7:     • Fit  $F$  to satisfy  $f_1 = F(f_3)$ ;
8:     for all  $\mathbf{X}^{\delta_{\text{adj}}} \subset D_{Y,\delta_{\text{adj}}}$  do
9:       • Calculate two scaling ratio as
          $\eta_{TE}(Y, \delta_{\text{adj}}, \delta_{\text{tar}}) = F(\delta_{\text{tar}})/f_1(\mathbf{X}^{\delta_{\text{adj}}})$ , (6)
          $\eta_{PR}(Y, \delta_{\text{adj}}, \delta_{\text{tar}}) = \delta_{\text{tar}}/f_3(\mathbf{X}^{\delta_{\text{adj}}})$ ; (7)
10:      • Calculate  $\mathbf{X}^{\delta_{\text{tar}}} = \langle \mathbf{X}_{TE}^{\delta_{\text{tar}}}, \mathbf{X}_{PR}^{\delta_{\text{tar}}} \rangle$  with
          $\mathbf{X}_{TE}^{\delta_{\text{tar}}} = \mathbf{X}_{TE}^{\delta_{\text{adj}}} \cdot \eta_{TE}(Y, \delta_{\text{adj}}, \delta_{\text{tar}})$ , (8)
          $\mathbf{X}_{PR}^{\delta_{\text{tar}}} = \mathbf{X}_{PR}^{\delta_{\text{adj}}} \cdot \eta_{PR}(Y, \delta_{\text{adj}}, \delta_{\text{tar}})$ ; (9)
11:      •  $D_{Y,\delta_{\text{tar}}} \leftarrow \text{append}(\mathbf{X}^{\delta_{\text{tar}}})$ ;
12:    end for
13:  end for
14:   $D_{Y,\Delta_{\text{tar}}} \leftarrow \text{append}(D_{Y,\delta_{\text{tar}}})$ ;
15: end for
```

---

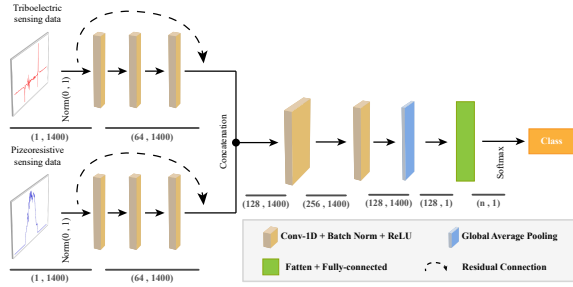
Algorithm 1 describes the virtual data generation process. The known datasets  $D_{Y,\Delta_{\text{real}}}$  and  $\Delta_{\text{tar}}$  are the input, and the output is  $D_{Y,\Delta_{\text{tar}}}$ , the generated datasets of  $\Delta_{\text{tar}}$ .

For each  $\delta_{\text{tar}}$ , we have  $\Delta_{\text{adj}}$  which is calculated by equation (4) and (5). To every sample  $\mathbf{X}^{\delta_{\text{adj}}}$  from different compression force

levels among  $\Lambda_{\text{adj}}$ , we get  $f_1$  and  $f_3$ , which are two representative features of  $\mathbf{X}^{\delta_{\text{adj}}}$ , using Eqn. (1) and (3). The next step is to fit the relationship between  $f_1(\mathbf{X}^{\delta_{\text{adj}}})$  and  $f_3(\mathbf{X}^{\delta_{\text{adj}}})$  by a cubic function  $F$ . Afterwards, we calculate the scaling ratios of triboelectric and piezoresistive sensing data denoted as  $\eta_{\text{TE}}(Y, \delta_{\text{adj}}, \delta_{\text{tar}})$  and  $\eta_{\text{PR}}(Y, \delta_{\text{adj}}, \delta_{\text{tar}})$ . Every generated sample  $\mathbf{X}^{\delta_{\text{tar}}}$  equals to the sample from  $D_{Y, \delta_{\text{adj}}}$  multiplied by two scaling ratios and is collected as an element of virtual dataset  $D_{Y, \delta_{\text{tar}}}$ . It is worth noting that the interval between target compression force level and its adjacent compression force level should be higher than the resolution of piezoresistive thin film.

### 3.3 Parallel Residual CNN Model

DL, which provides an efficient solution to automatically learn the temporal dependencies in time-series data, has successfully been applied to address time series data processing and forecasting problems. There are some SOTA DL models for time-series data processing, such as LSTM-FCN [9], InceptionTime [7], and Res-CNN [21]. Considering the dataset feature of our multimodal sensor, we propose the PR-CNN model, which has a smaller model size but higher accuracy than the SOTA models.



**Figure 2: Multimodal sensing signals and customized residual network based recognition model for extracting and classifying materials and textures. (n: number of classes)**

The input end of the PR-CNN consists of two parallel residual blocks fed with the triboelectric signal and the piezoresistive signal respectively. Two one-dimensional convolutional layers serve as the data fusion block to fuse the features from the two residual blocks. The output layer following the data fusion block consists of a global pooling layer and a softmax. The schematic illustration of the PR-CNN is depicted in Fig. 2. Since the electrostatic interference from the external environment can cause sufficient noise to our sensors, the residual structure is necessary to suppress noise in the sensing signal and enhance the gradient propagation. The two parallel residual blocks extract features independently to avoid mutual interference between the two modalities, which improves training efficiency.

## 4 EXPERIMENTS AND RESULTS

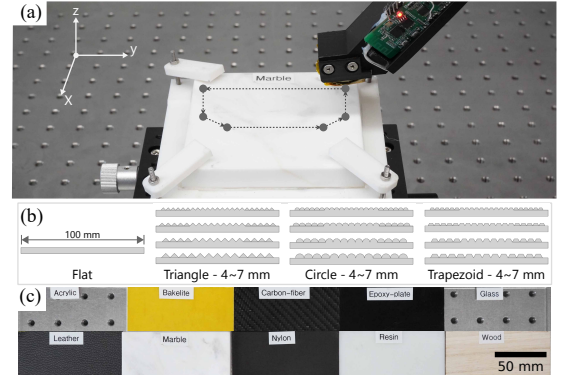
### 4.1 Calibration for the Piezoresistive Thin Film

A commercial integrated pressing sensor (DS2-50N-XD) is selected as the ground truth to calibrate the piezoresistive thin-film sensor. We set our multimodal tactile sensor in the vertical direction of

the pressing sensor and collect 20 sets of piezoresistive thin-film sensing voltage data corresponding to precise compression force levels by slowly adjusting the height of the fine-tuning platform.

### 4.2 Experimental Setup

Fig. 3a shows the whole data collecting system. A 6-axis manipulator, a 3D-printing finger and a two-degree-of-freedom fine-tuning platform are adopted to perform the data collection in the tactile sensing system by mimicking the sliding process of human interactions. The fine-tuning platform works as fixing experimental boards as well as setting six different compression force levels  $\Delta = \{1 \text{ N}, 5 \text{ N}, 10 \text{ N}, 15 \text{ N}, 20 \text{ N}, 25 \text{ N}\}$ . With the help of the robot arm, the fingertip can slide on the board surface along a trapezoidal track. The base angle is set to 30 degrees to extend the contact-separation phase. The wireless acquisition board is used to synchronously acquire the signals of the triboelectric and piezoresistive sensors at the sampling frequency of 250 Hz. The sliding movement is repeated  $n = 15$  times for every compression force level.



**Figure 3: Data collecting system and experiment setup. (a) Experimental platform for data collection. (b) 13 experiment specimens of different textures. (c) 10 experimental specimens made of different materials.**

In this paper, we validate the performance of our fingertip multimodal tactile sensing system, our model with multi-channel residual structure and virtual data generation method through three classification tasks as follows.

**4.2.1 Texture Classification.** As the Fig. 3b shows, we design 13 different texture specimens with four different types of shapes and four different periodic lengths (4, 5, 6, 7 mm). All the texture boards are in the size of 10 cm by 10 cm, made from resin and manufactured by stereo lithography appearance technique in WeNext Technology Co., Ltd. They are named Flat, Cir4~7, Tri4~7, Tra4~7, respectively.

**4.2.2 Material Classification.** As Fig. 3c shows, we select 10 different materials in the size of 10 cm by 10 cm, including Acrylic, Bakelite, Carbon fiber, Epoxy plate, Marble, Leather, Nylon, Glass, Resin and Wood, which are commonly seen in our daily life.

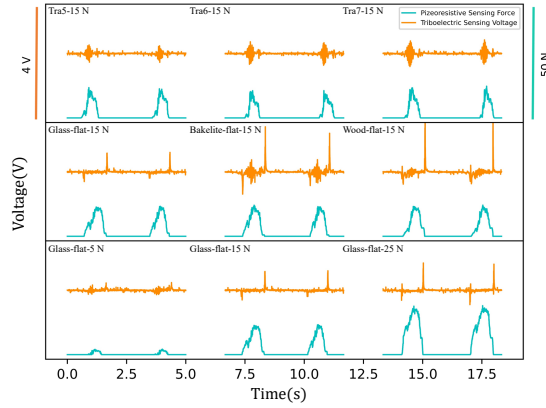
**4.2.3 Texture and Material Classification.** We prepare 18 boards with 10 materials (Silk, Paper, Acrylic, Bakelite, Marble, Leather, Nylon, Glass, Resin, Wood) and 5 textures (Flat, Tri4, Tri6, Cir4,



Cir6). The 18 boards are named Silk-flat, Silk-cir4, Silk-cir6, Silk-tri4, Silk-tri6, Paper-flat, Paper-cir4, Paper-cir6, Paper-tri4, Paper-tri6, Acrylic-flat, Bakelite-flat, Marble-flat, Leather-flat, Nylon-flat, Glass-flat, Resin-flat, Wood-flat.

### 4.3 Data Preprocessing & Training Process

After finishing tactile sensing data collection, we follow these three procedures to complete data splitting and augmentation: (1) Discard redundant data points at the beginning and end of every time sequence data file. (2) Generate samples through a sliding window with a length of  $m = 1400$  data points and a step of 10 data points. (3) Select the qualified samples by comparing their variance with a preset threshold. In the experiment, every time sequence data file can generate around 150 samples. Fig. 4 shows the waveforms of samples and compares them horizontally by controlling different textures, materials and compression force levels.



**Figure 4: Multimodal tactile sensing waveforms of sliding on different materials and textures under three compression force levels.**

After finishing the data preprocessing, we separate our samples into training set and validation set according to the compression force levels. The training set includes samples of compression force levels  $\Delta_{\text{real}}$ . And the rest belong to the validation set. We generate target dataset  $D_{Y, \Delta_{\text{tar}}}$  of compression force levels  $\Delta_{\text{tar}}$  from  $D_{Y, \Delta_{\text{real}}}$  through virtual data generation method, and combine them with the training set as the final training set.

In the training process, we select cross-entropy as the loss function, and keep the same batch size, epoch, and initial learning rate in the experiment. The specific parameters are 128, 40, 0.05, separately. To speed up convergence, the learning rate decreases to a tenth in the epochs of 5, 15, 25. This can help model with more delicate training performance and overcome the local optimum situation. The code is based on the Pytorch library, and models are trained on a GeForce RTX 3090 GPU.

### 4.4 Validation Results

To prove the advancement of multi-channel residual structure, we select Res-CNN [21], InceptionTime [7], LSTM-FCN [9], three SOTA models, as the baseline model in the training process and make a

performance comparison between our model and the other three models. Table. 1 shows the performance of three DL methods in material classification task. Four models can distinguish ten materials with high accuracy. Our model shows excellent classification ability which has the highest validation accuracy, smallest model size as well as lowest validation loss among the four models.

**Table 1: Performance Comparison in Material Classification.**

| Method            | Model Size     | Val Loss     | Val Acc       |
|-------------------|----------------|--------------|---------------|
| <b>Our model</b>  | <b>266,251</b> | <b>0.047</b> | <b>99.06%</b> |
| InceptionTime [7] | 456,074        | 0.060        | 98.14%        |
| Res-CNN [21]      | 257,803        | 0.141        | 96.30%        |
| LSTM-FCN [9]      | 786,276        | 0.253        | 91.79%        |

In addition, we validate our virtual data generation method by comparing the training process with and without the method. Without the virtual data generation method, the training accuracy is close to 1, when the validation accuracy is around 10% which is the same as the random probability among 10 materials. It seems that the virtual data generation method enhances the adaptability of the DL model under compression force levels  $\Delta_{\text{tar}}$  and helps the model alleviate the overfitting phenomenon.

Finally, we deploy the proposed model for three classification tasks. Fig. 5 shows the results in the form of confusion matrices. And the average recognition accuracies of the three tasks reach 96.47%, 99.52% and 97.57%, respectively. The lowest recognition accuracy of all classes based on the combined dataset of materials and textures is 90.90% which appears on the classification of silk board with triangle texture and 6 mm periodic length.

## 5 CONCLUSIONS AND DISCUSSIONS

In this work, we design a multimodal tactile sensing system for texture and material classification. A fingertip flexible tactile sensor based on a sandwich-like four-layered structure that couples the triboelectric and piezoresistive effects is proposed. We also propose a virtual data generation method to enhance model generalization ability and increase the data availability. Furthermore, the PR-CNN model with a multi-channel residual structure is developed and outperforms the SOTA models on three texture and material classification tasks with over 96% accuracy. A real-time system is also implemented on the robot manipulator, indicating the potential applications in service robots and HMI.

## ACKNOWLEDGMENTS

This work is supported in part by Tencent Robotics X Lab, and by the grant from the Institute for Guo Qiang of Tsinghua University 2020QG1004.

## REFERENCES

- [1] Asli Atalay, Vanessa Sanchez, Ozgur Atalay, Daniel M. Vogt, Florian Haufe, Robert J. Wood, and Conor J. Walsh. 2017. Batch Fabrication of Customizable Silicone-Textile Composite Capacitive Strain Sensors for Human Motion Tracking. *Advanced Materials Technologies* 2, 9 (2017), 1700136. <https://doi.org/10.1002/admt.201700136>
- [2] Ravinder Dahiya. 2019. E-Skin: From Humanoids to Humans. *Proc. IEEE* 107, 2 (2019), 247–252.

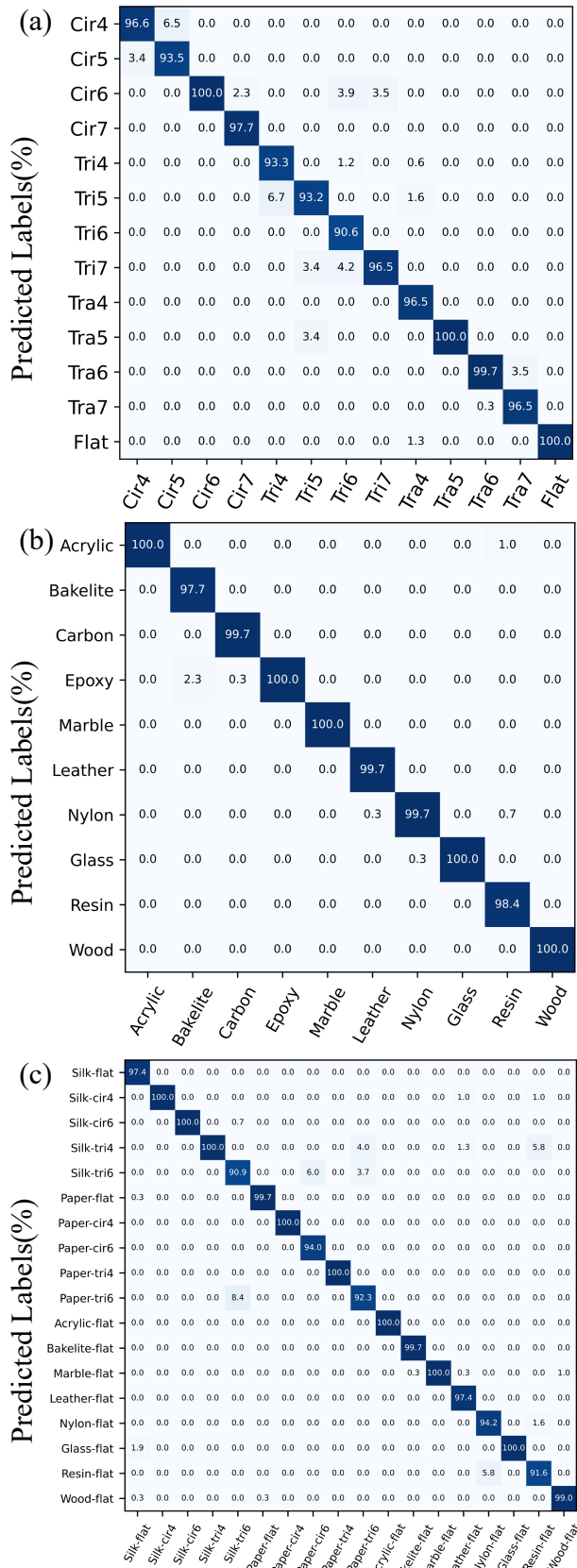


Figure 5: Confusion matrices of the (a) texture classification task, (b) material classification task, and (c) texture and material classification task.

- [3] Ravinder S Dahiya, Philipp Mittendorf, Maurizio Valle, Gordon Cheng, and Vladimir J Lumelsky. 2013. Directions toward effective utilization of tactile skin: A review. *IEEE Sensors Journal* 13, 11 (2013), 4121–4138.
- [4] Wenbo Ding, Aurelia C. Wang, Changsheng Wu, Hengyu Guo, and Zhong Lin Wang. 2019. Human–Machine Interfacing Enabled by Triboelectric Nanogenerators and Tribotronics. *Advanced Materials Technologies* 4, 1 (2019), 1800487.
- [5] Jonathan Engel, Jack Chen, Zhifang Fan, and Chang Liu. 2005. Polymer micro-machined multimodal tactile sensors. *Sensors and Actuators A: Physical* 117, 1 (2005), 50–61.
- [6] Qilin Hua, Junlu Sun, Haitao Liu, Rongrong Bao, Ruomeng Yu, Junyi Zhai, Caofeng Pan, and Zhong Lin Wang. 2018. Skin-inspired highly stretchable and conformable matrix networks for multifunctional sensing. *Nature Communications* 9, 1 (2018), 1–11.
- [7] Hassan Ismail Fawaz, Benjamin Lucas, Germain Forestier, Charlotte Pelletier, Daniel F Schmidt, Jonathan Weber, et al. 2020. Inceptiontime: Finding alexnet for time series classification. *Data Mining and Knowledge Discovery* 34, 6 (2020), 1936–1962.
- [8] Tao Jin, Zhongda Sun, Long Li, Quan Zhang, Minglu Zhu, Zixuan Zhang, et al. 2020. Triboelectric nanogenerator sensors for soft robotics aiming at digital twin applications. *Nature Communications* 11, 1 (2020), 1–12.
- [9] Fazle Karim, Somshubra Majumdar, Houshang Darabi, and Samuel Harford. 2019. Multivariate LSTM-FCNs for time series classification. *Neural Networks* 116 (2019), 237–245.
- [10] Naoto Komeno and Takamitsu Matsubara. 2021. Tactile Perception Based on Injected Vibration in Soft Sensor. *IEEE Robotics and Automation Letters* 6, 3 (2021), 5365–5372.
- [11] Weiting Liu, Ping Yu, Chunxin Gu, Xiaoying Cheng, and Xin Fu. 2017. Fingertip Piezoelectric Tactile Sensor Array for Roughness Encoding Under Varying Scanning Velocity. *IEEE Sensors Journal* 17, 21 (2017), 6867–6879. <https://doi.org/10.1109/JSEN.2017.2721740>
- [12] Francisco Pastor, Jorge García-González, Juan M Gandarias, Daniel Medina, Pau Closas, Alfonso J García-Cerezo, et al. 2020. Bayesian and neural inference on LSTM-based object recognition from tactile and kinesthetic information. *IEEE Robotics and Automation Letters* 6, 1 (2020), 231–238.
- [13] Mahdi Rasouli, Yi Chen, Arindam Basu, Sunil L. Kukreja, and Nitish V. Thakor. 2018. An Extreme Learning Machine-Based Neuromorphic Tactile Sensing System for Texture Recognition. *IEEE Transactions on Biomedical Circuits and Systems* 12, 2 (2018), 313–325. <https://doi.org/10.1109/TBCAS.2018.2805721>
- [14] Ziwu Song, Jihong Yin, Zihan Wang, Chengyue Lu, Ze Yang, Zihao Zhao, et al. 2022. A flexible triboelectric tactile sensor for simultaneous material and texture recognition. *Nano Energy* 93, 1 (2022), 106798.
- [15] Yongbiao Wan, Yan Wang, and Chuan Fei Guo. 2017. Recent progresses on flexible tactile sensors. *Materials Today Physics* 1 (2017), 61–73.
- [16] Xiao Wei, Hao Li, Wenjing Yue, Song Gao, Zhenxiang Chen, Yang Li, and Guozhen Shen. 2022. A high-accuracy, real-time, intelligent material perception system with a machine-learning-motivated pressure-sensitive electronic skin. *Matter* 5, 5 (2022), 1481–1501.
- [17] Changsheng Wu, Aurelia C. Wang, Wenbo Ding, Hengyu Guo, and Zhong Lin Wang. 2019. Triboelectric Nanogenerator: A Foundation of the Energy for the New Era. *Advanced Energy Materials* 9, 1 (2019), 1802906.
- [18] Shichao Yue and Walied A. Moussa. 2018. A Piezoresistive Tactile Sensor Array for Touchscreen Panels. *IEEE Sensors Journal* 18, 4 (2018), 1685–1693. <https://doi.org/10.1109/JSEN.2017.2776936>
- [19] Zhong Zhang, Yu Zheng, Jia Pan, Xiong Li, Kaiwei Li, and Zhengyou Zhang. 2020. A Flexible Dual-Core Optical Waveguide Sensor for Simultaneous and Continuous Measurement of Contact Force and Position. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2020*. Las Vegas, USA, 7375–7380.
- [20] Wendong Zheng, Huaping Liu, and Fuchun Sun. 2020. Lifelong visual-tactile cross-modal learning for robotic material perception. *IEEE Transactions on Neural Networks and Learning Systems* 32, 3 (2020), 1192–1203.
- [21] Xiaowu Zou, Zidong Wang, Qi Li, and Weiguo Sheng. 2019. Integration of residual network and convolutional neural network along with various activation functions and global pooling for time series classification. *Neurocomputing* 367 (2019), 39–45.