SCIENCE CHINA
Information Sciences

• Supplementary File •

Data-driven Electrical Resistance Tomography for Robotic Large-Area Tactile Sensing

Wendong Zheng[1,2], Huaping Liu[2*], Xiaofeng Liu[3] & Fuchun Sun[2]

[1] *School of Electrical Engineering and Automation, Tianjin University of Technology, Tianjin 300382, China;*
[2] *Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China;*
[3] *Key Laboratory of Maritime Intelligent Cyberspace Technology (Hohai University), Ministry of Education, Changzhou 213022, China*

All files will be submitted through our online electronic submission system at HERE.

Appendix A ERT-based Tactile Sensor

In order to understand the principles of tactile sensing based on ERT technology, the theoretical background is briefly described. Then, we will introduce the fabrication process of the ERT-based tactile sensor and its data acquisition system.

Appendix A.1 Theoretical Backgrounds of ERT

To comprehend the working principle of the ERT-based tactile sensor, it is essential to understand the forward and inverse problems of ERT. They describe the relationship between the measured voltage and the conductivity distribution within the sensor. Therefore, we briefly explain the forward and inverse problems of the ERT-based tactile sensor. For more detailed information, see [1].

The ERT method utilizes Maxwell's equations to establish the relationship between conductivity σ and voltage potential u within a closed bounded domain Ω.

$$\sigma \nabla^2 u = 0 \quad \text{in} \quad \Omega \tag{A1}$$

When a current is injected into the boundary domain $\partial\sigma$, the boundary condition can be expressed as,

$$-\partial \nabla u \cdot n = j_l \tag{A2}$$

where j_l is the current density at electrode e_l, and n is unit vector normal to $\partial\Omega$.

By combining equations (A1) and (A2), the voltage at the boundary V_l can be calculated based on a given conductivity distribution σ and the injection current density j_l. This is the forward problem of ERT, where the goal is to predict the resulting voltage measurements on the boundary given the known conductivity distribution and injected current. In practical applications, In practical applications, the finite element method (FEM) is commonly employed to solve this forward problem.

In ERT, the primary concern is how to determine the conductivity distribution σ based on the measured boundary voltage V_l. For ERT-based tactile sensing, the fundamental objective is to reconstruct the changes in conductivity distribution $\delta\sigma$ corresponding to the change in boundary voltage δV. This inverse problem is nonlinear and ill-posed, making it particularly challenging to solve.

To address this problem, a linear approximation can be employed to describe the relationship between $\delta\sigma$ and δV. This linear approximation assumes that the changes in conductivity and voltage are small enough to be considered linearly related. Mathematically, this relationship can be expressed as:

$$\delta V = J\delta\sigma + w, \tag{A3}$$

where J is the Jacobian matrix, which can be calculated using the finite element model, and w is measurement noise. By computing the pseudo-inverse with a regularization term, the solution of $\delta\sigma$ can be written as:

$$\delta\sigma = (J^T J + \lambda^2 R^T R)^{-1} J^T \delta V, \tag{A4}$$

where R is the regularization matrix, and λ is the regularization hyperparameter.

Appendix A.2 Fabrication Process

Considering the advantage of multi-layer structures in sensing distributed forces over a large area, our tactile sensor adopts a multi-layer structure, as illustrated in Fig.A1(b). In this sensor, the base layer is a $20 \times 20cm^2$ flexible printed circuit, where 16 electrodes are evenly arranged on the boundary. Following our previous work [2], we fabricate a sensing domain on the base layer by spray-coating it with carbon black CRAMOLIN 1281411). The base conductivity is controlled to be around 0.06S/m.

To enhance the sensor's ability to detect multiple contact points, the sensor's conductive layer is composed of discrete highly conductive fabric patches (Silver Fiber, YSILVER82, China) instead of using a monolithic conductive fabric. This is due to that the lattice structure is implemented to effectively prevent current flow between different touch points through the sensing layer, enhancing the sensor's ability to accurately detect and distinguish multiple pressure points. In particular, each conductive fabric patch is set as $7.5 \times 7.5mm^2$ and cut with laser cutters. This conductive layer is constructed by affixing a series of 24x24 patches of conductive fabric onto the neoprene foam using adhesive. Subsequently, the base layer and the conductive layer are carefully assembled and firmly secured using tape, as shown in Fig.A1(a). With the aforementioned fabrication process, the tactile sensor exhibits softness and flexibility, making it highly suitable for deployment on robot surfaces.

* Corresponding author (email: hpliu@tsinghua.edu.cn)

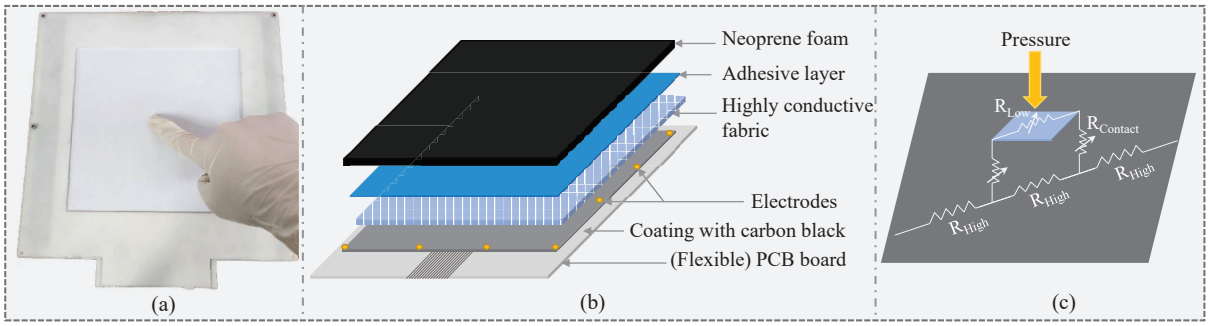

Figure A1 Schematic diagram of the structure of the tactile sensor

Appendix A.3 Data Acquisition System

To use the ERT sensor, a data acquisition system has been developed. This system incorporates the data acquisition card (HK_NET6024-S), along with a customized multiplexing board. The electrodes of the sensor sensing substrate are connected to two multiplexers (MAX306) on the multiplexer board. During sensor operation, the multiplexer is controlled by the STM32F107 microcontroller to select specific electrodes for current excitation or voltage measurement. Please refer to our previous work [2] for details on sensor fabrication.

A pairwise injection pattern is used. To facilitate the deployment of sensors in robotic applications, we use a battery and a voltage-current module to provide current excitation for the sensor rather than a constant current source. For voltage measurement, the parallel acquisition method is adopted, which not only effectively improves the acquisition efficiency, but also reduces the time difference between the data of all channels.

Appendix B Data-driven ERT Sensing Method

The overall architecture of DD-ERT is shown in Fig.B1. This is an end-to-end model, and the modules are jointly trained. For ease of description, the architecture is structurally divided into two cascaded modules: the initial imaging module and the image reconstruction network. Firstly, the initial conductivity image is reconstructed from the measured boundary voltages by the initial imaging module. After initial reconstruction, the image is fed into a reconstruction network to generate an enhanced conductivity image.

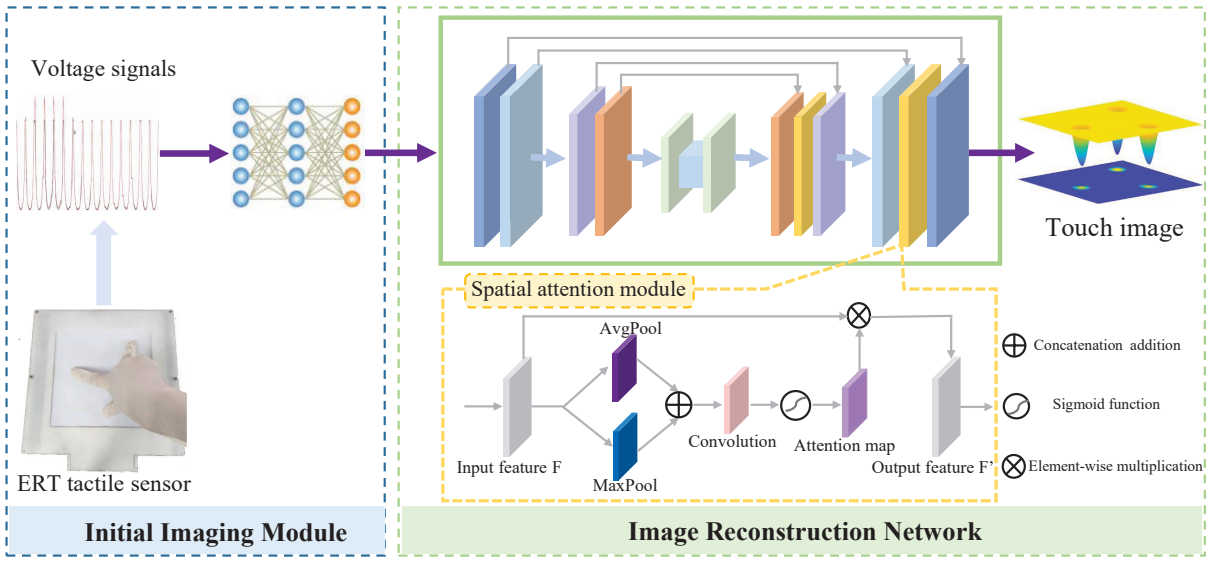

Figure B1 Framework of the proposed method DD-ERT.

Please note that the traditional data-to-image method achieves ERT image reconstruction by training a network model to learn the nonlinear relationship between boundary voltage and conductivity distribution. In these approaches, the physical model of the ERT is not explicitly considered, and the neural network needs to learn the basic physical model from scratch. This learning task is often difficult or even impossible for ERT-based sensors. Unlike existing data-to-image methods, we initialize the initial imaging module using the sensitivity matrix obtained by traditional reconstruction algorithms in this work. Due to the introduction of the physical model, the network is easier to converge during training and the performance of the model can be improved. Compared with methods that use a fixed initial reconstructed image, our proposed method can effectively eliminate errors caused by ill-posedness, thereby improving the reconstruction quality of tactile images.

Appendix B.1 Initial Imaging Module

The initial imaging module comprises a fully connected layer, which is responsible for processing the input voltage data and generating the corresponding initial conductivity image. The input to this module is a 1×208 voltage data obtained from the ERT-based sensor, and the output is an initial conductivity image represented by a 1×576 vector. Considering the positional relationship of the conductivity vector within the measurement domain, the conductivity vector is projected onto a 24×24 initial conductivity image that aligns with the spatial layout of the sensor. The resulting conductivity image is used as the input to the subsequent image reconstruction network. The mathematical model of the initial imaging module can be expressed as follows:

$$\sigma_i{=}f_\theta(v_\Delta), v^{1\times 208} \rightarrow \sigma^{24\times 24} \tag{B1}$$

where f_θ denotes the network of the initial imaging module with parameters θ. $v_\Delta = v_i - v_0$ is the difference between the voltage v_i at time t and the voltage v_0 at initial time t_0, and σ_i is touch conductivity image corresponding to time t.

Appendix B.2 Image Reconstruction Module

The main goal of this module is to generate high-quality distribution images from low-resolution initial images to effectively describe tactile information about physical interactions.

Appendix B.2.1 *Network Structures*

(1) Generator: U-Net [3] has been widely recognized for its ability to preserve detailed spatial information through the use of skip connections, making it suitable for tasks such as image-to-image translation. However, in our experiments, we observed that the standard encoder-decoder U-Net networks do not explicitly focus on the regions that represent the effective contact area, which greatly limits the performance of the model. To this end, we propose a generative adversarial model for the ERT-based tactile sensor, which follows the basic structure of U-Net, as shown in Fig.B1.

In the encoding stage, a series of multi-layer convolutional blocks is employed to learn and extract high-level features from the initial conductivity image. In the decoding stage, the encoded features are progressively upsampled and merged with CNN features from different levels that were skipped in the encoding path. By combining features from different levels of abstraction, the network can leverage both local and global contextual information, resulting in more accurate and detailed reconstructions. To further enhance the performance of the model, a self-attention module is integrated into the network to explicitly capture the effective touching area. In particular, a self-attention module is added to the last two feature maps of size 6×6 and 4×4, respectively.

This network design is primarily motivated by the fact that the attention mechanism exhibits a superior capability in capturing global information. This enables the model to have the ability to dynamically allocate its attention to different regions based on their saliency. This adaptability is critical for the network to accurately capture and understand tactile data, especially when dealing with complex or dynamic tactile interactions. Moreover, by integrating the spatial attention module, the network can selectively emphasize regions where both high-level semantic features and fine-grained details are present. In addition, deploying this module at a high level can help the network converge during training.

As shown in Fig.B1, the spatial attention feature map S is mainly obtained by the operation of global average pooling (GAP), global maximum pooling (GMP), and convolution. Given a feature map $F \in R^{C\times H\times W}$, we first apply GAP and GMP operations on a feature map F, generating two 2D maps $F^s_{avg} = AvgPool(F) \in R^{1\times H\times W}$ and $F^s_{\max} = MaxPool(F) \in R^{1\times H\times W}$. They are then concatenated and convolved through standard convolutional layers, resulting in a 2D spatial attention map. In summary, the computation of the spatial attention can be expressed as:

$$M_s(F) = sig(Con([F^s_{avg}; F^s_{\max}]) \tag{B2}$$

where sig represents the sigmoid function and Con denotes a convolution operation.

(2) Discriminator: In our model, the discriminator follows the PatchGAN [4] design. In a regular GAN, the discriminator typically outputs a scalar value representing its confidence in the input being real or fake. In PatchGAN, the discriminator instead outputs a matrix of size $N \times N$. Each element in the matrix represents whether a patch of the input image is classified as real or fake. In particular, the input image is divided into smaller patches, and the discriminator evaluates each patch individually. By considering patches instead of the entire image, the discriminator can capture fine-grained details and textures from different perspectives.

In the ERT imaging problem, the PatchGAN-based discriminator penalizes the generator to make the output conductivity image more discriminative. In particular, PatchGAN enables the generator to capture and evaluate image details at different scales and locations by analyzing patches of input images. With this discriminator constraint, the generator is able to generate more realistic and accurate conductivity images.

Appendix B.2.2 *Objective Function*

(1) Content Consistency Loss: Existing research has demonstrated that the use of the L2 norm in generative networks can lead to the generation of blurry images and an inability to capture low-frequency components that are present in tactile images. As suggested by [5], the L1 norm can effectively capture low-frequency components and does not require a new framework. This motivated us to integrate the L1 norm into the objective function to ensure low-frequency correctness. Simultaneously, the discriminator network can focus primarily on modeling high-frequency components. In particular, the parameters of the generator G can be learned using the L1 norm between the generated image and the target image as a loss function, which can be expressed as

$$L_{L1}(\theta_G) = \frac{1}{N}\sum_{i=1}^{N} \|G(\sigma_0) - \sigma\|_1 \tag{B3}$$

(2) Texture Adversarial Loss: During training, the discriminator D learns to distinguish between real data and generated data by assigning higher probabilities to real samples and lower probabilities to samples obtained from the generator G. This process drives adversarial learning and enables the generator to learn from the feedback provided by the discriminator, ultimately enabling the generation of higher-quality tactile images. The objective function of the discriminator D is to minimize the texture adversarial loss $L_{adv}(\theta_G, \theta_D)$, which can be expressed as

$$L_{adv}(\theta_G, \theta_D) = \log D(\sigma) + \log(1 - D(G(\sigma_0)) \tag{B4}$$

where σ denotes the true conductivity image, σ_0 represents the initial conductivity image from the initial imaging module, and σ_g is the image generated by the generator of the image reconstruction module.

(3) Mask-based Constraint Loss: In an ERT-based tactile sensor, a tactile image can be partitioned into two main regions: the background region and the foreground region. The background region of a tactile image refers to the untouched area, which usually does not contain any tactile information related to object contact. The foreground region represents the object-sensor interaction area, which contains tactile signals and provides information about the interaction between the sensor and the object.

For the background region, our focus is primarily on ensuring profile similarity between the generated image and the reference tactile image, without necessarily capturing its specific structural details. In contrast, when it comes to the foreground region, our attention shifts towards the contact information directly reflected by the pixel values. These pixel values and edge structures within the foreground area are crucial for ERT-based tactile sensing because they contain contact information required for tactile perception and analysis.

Based on the above analysis and discussion, mask-based constraint loss is proposed for the generative model. Firstly, the application of a threshold to the pixel values of the tactile image σ enables the division of the image into foreground and background regions, resulting in the binary mask image σ_{mask}. Specifically, any pixel value below a certain threshold is set to 0, while pixel values equal to or above the threshold are set to 1. Then, the target image σ_{mask} is multiplied by

the mask image to obtain an image $\sigma_{rm} = \sigma_r \times \sigma_{mask}$ that contains only the region of interest (ROI). The same operation is performed on the reconstructed image to get an image $\sigma_{gm} = \sigma_g \times \sigma_{mask}$. The mask loss is defined as the cross entropy loss between σ_{rm} and σ_{gm}, which is expressed as

$$L_{\text{mask}}(\theta_{\text{G}}) = -\frac{1}{N}\sum_{i=1}^{N} \sigma_{rm}^{i} \log \sigma_{gm}^{i} + (1-\sigma_{rm}^{i}) \log(1-\sigma_{gm}^{i}) \tag{B5}$$

where N is the number of conductivity values in the ERT tactile image.

(4) Overall Loss and Algorithm Optimization: With the above discussions, the formally objective function of the DD-ERT model consists of three loss functions, which can be expressed as

$$L(\theta_G, \theta_D) = L_{mse} + \alpha L_{adv} + \beta L_{mask} \tag{B6}$$

where α and β denote the weight parameter for mean square error loss L_{mse} and mask loss L_{mask}, respectively.

With the above discussions, the training goal of the discriminator D is to maximize the probability of the training data and minimize the probability of the sampled data obtained from the generator G, and the training goal of the generator G is to generate images that are as similar as possible to real images. This is a minimax optimization problem, which can be expressed as,

$$L(\theta_G, \theta_D) = \arg\min_{G}\max_{D}(L_{L1} + \alpha L_{adv} + \beta L_{mask}) \tag{B7}$$

The discriminator D and the generator G can be trained alternately using the gradient descent method.

Appendix B.3 Dataset Generation Strategy

When training the model, there is a challenge faced by the lack of specific data sets available for ERT-based tactile sensors since there are no instruments that can accurately measure the conductivity distribution. Synthetic data generation is a useful alternative. In this work, EIDORS was ultimately chosen for data generation because it has proven suitable for modeling ERT-based tactile sensing in the literature.

The simulated conductivity distribution is an important factor in generating the data set, as it should reflect real-world scenarios. To achieve this, a homogeneous background reference conductivity was set to $c_0 = 0.06S/m$. In practical applications, when a tactile sensor is subjected to external force or pressure, the local conductivity of the contact area will decrease. Therefore, the conductivity value at the contact location is set within a range of $(0.1-0.9) * c_0$ S/m. The relatively wide conductivity variation range is intended to make the model have better generalization and robustness in real sensors. To ensure the simulation data reflects real-world conditions accurately, random noise was added to the voltage measurements obtained from the forward model of ERT. By incorporating such noise into the simulated data, the model can be trained and evaluated under more realistic conditions, improving its ability to handle noise and generalize to real-world scenarios.

Rather than complete random synthesis, a specific strategy for generating simulation data is proposed to ensure the generalization and robustness of the proposed model. We apply a perturbation of conductivity values to the elements of an arbitrary finite element mesh and solve a forward problem to obtain the corresponding voltage measurements. FEM discretizes the continuous sensing domain into N elements. In this experiment, the sensor's resistive network consists of a total of 24×24elements, where each square element is subdivided into two right-angled sub-units to improve simulation accuracy. Since our electrode shape can be effectively approximated by point contacts, the point electrode model is adopted.

Appendix C Effectiveness analysis of the method

Appendix C.1 Comparison Methods

To validate the effectiveness of our proposed method, different reconstruction algorithms are compared with the proposed method in the simulation data. One-step Gauss-Newton (one-step GN), Truncated Singular Value Decomposition (TSVD), and primal-dual interior point method (PDIPM) have been widely applied to some ERT imaging tasks. In addition, some deep-learning-based algorithms are applied to reconstruct ERT images. Therefore, the three traditional algorithms of GN, TSVD, PDIPM, and deep-learning-based algorithms of FCNN, CNN, and EIT-NN [6] are used as comparison methods in this experiment.

In this simulation, a two-dimensional square sensing membrane with normalized side length $l = 1$ experienced a circular stimulus with radius $r = 0.2$. The baseline conductivity of the sensing film was set at $\sigma_0 = 0.06S/m$, and target conductivity change is $\sigma_\Delta = 0.4\sigma_0$. In this experiment, a constant SNR of 50dB is applied to the simulation data. This is mainly because 50dB SNR is comparable to the noise level in real-world scenarios. The adjacent driving and measuring method is used in this experiment. An excitation current of 25 mA is injected into adjacent electrode pairs and voltages are measured between pairs of the remaining electrodes.

Appendix C.2 Evaluation Metrics

In order to quantitatively evaluate the quality of reconstructed tactile images by different methods, the following four metrics are used for evaluation.

(1) **AR**: The amplitude in the reconstructed image represents the magnitude of the conductivity change caused by the added stimulus. Therefore, the amplitude of the image is used to quantify the intensity of the tactile stimulus. The amplitude response (AR) is the measure of the ratio between the pixel intensities of the true conductivity to that of the predicted conductivity, which is given by

$$AR = \frac{\sum_{i=1}^{N} \sigma_i^*}{\sum_{i=1}^{N} \sigma_i} \tag{C1}$$

where σ_i^* is the predicted conductivity of the ith element, and σ_i is the true conductivity of the ith element. Ideally, the AR should have a value of 1 at any location based on the same input stimulus, as any value other than 1 indicates an intensity different from the true conductivity.

(2) **RES**: For quantitative examination of spatial resolution of reconstructed tactile images, the metric of RES [6] is adopted. It is the ratio of the area higher than 50% of the maximum value of the conductivity change $A_{50\%}$ to the whole conductive domain area σ_{total}, which can be represented as follows:

$$RES = (1 - \sqrt{\frac{A_{50\%} - A_{single}}{A_{total}}}) \times 100 \tag{C2}$$

where A_{single} is the area of a single FEM element. By subtracting the σ_{single} value from the square root operation, the metric is bound to 100%. This is because the smallest area contains the single largest element. As the RES_{50} value increases, so does the tactile sensor's ability to differentiate between two simultaneous touch stimuli rather than reconstructing them as one. A value of 100 represents the best performance for this metric.

(3) **PE**: In order to evaluate the localization performance of the sensor, the centroid localization error [7] is utilized to measure. It can be calculated from the Euclidean distance between the center of target stimuli (x_0, y_0) and the centroid of reconstructed conductivity change (x_r, y_r), which is given as:

$$PE = \|(x_r, y_r) - (x_0, y_0)\| \tag{C3}$$

Here, (x_r, y_r) is defined as the weighted center of the reconstructed stimulus, which is calculated from

$$(x_r, y_r) = \left(\frac{\sum_{i=1}^{N} \sigma_i x_i}{\sum_{i=1}^{N} \sigma_i}, \frac{\sum_{i=1}^{N} \sigma_i y_i}{\sum_{i=1}^{N} \sigma_i}\right) \tag{C4}$$

(x_i, y_i) is the coordinate of the i-th FEM element, σ_i is the reconstructed conductivity value and N is the total number of elements of the reconstructed image. A zero error value indicates that the sensor has achieved accurate positioning performance.

(4) **RIE**: Relative image Error quantifies the difference between the reconstructed conductivity image and the real conductivity image, which is defined as

$$RIE = \frac{\|\sigma_R - \sigma_0\|_2}{\|\sigma_0\|_2} \tag{C5}$$

where σ_R is the reconstructed conductivity image, and σ_0 denotes the target conductivity image. As stated in the definition, the lower the value of RIE, the higher the quality of the reconstructed image.

Appendix C.3 Experimental results and analysis

The tactile images reconstructed by these seven methods are quantitatively evaluated using the four evaluation metrics mentioned above. The results are shown in Table C1.

Table C1 Performance comparison of seven different methods

Methods	AR	RES	PE	RIE
One-step GN	1.231	79.81	0.1204	0.0664
TSVD	1.235	80.12	0.1431	0.0681
PDIPM	1.218	81.17	0.1391	0.0679
FCNN	1.206	82.81	0.1452	0.0591
DCNN	1.181	84.9	0.1429	0.0588
EIT-NN	1.096	85.26	0.1369	0.0581
Our method	**1.061**	**86.69**	**0.0843**	**0.0501**

As shown in Table C1, it can be observed that our proposed method achieves the best performance across all four metrics. This means that the proposed model is better than other methods. In particular, the proposed method can obtain tactile information more accurately during interaction. It should also noted that the deep learning methods perform significantly better than the classical linear approach. These results demonstrate the effectiveness of our method in ERT tactile image reconstruction.

Appendix C.4 Effectiveness of Different Components

To investigate the role of each component in the proposed model, additional experiments are conducted on three different variant combinations of components under the same tactile contact conditions. The average AR, RES, PE, and RIE of reconstructed images of the proposed method and its three different variant combinations are shown in Fig.C1.

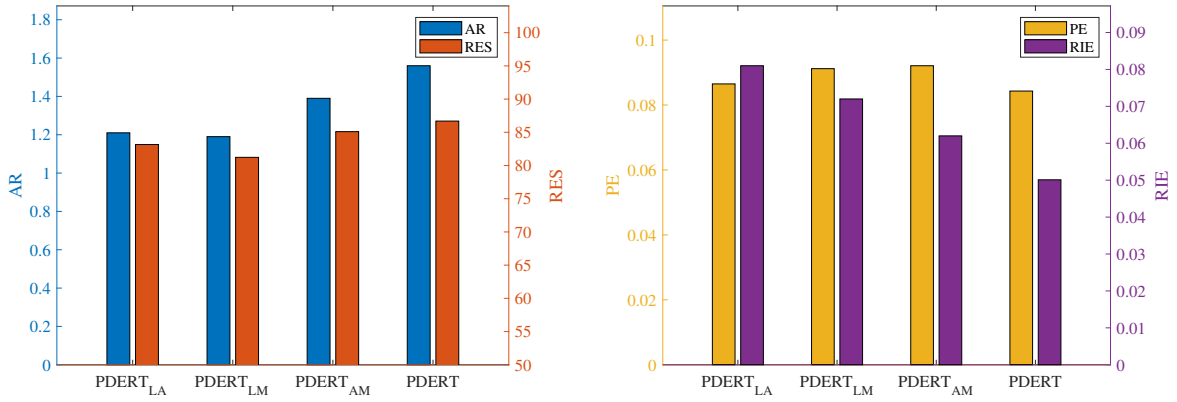

Figure C1 The experimental results of the proposed method and its three component variations. (a) Average AR and RES; (b) Average PE and RIE

From Fig.C1, It can be found that the performance of the proposed model is better than any of the other three variant combinations. It shows that optimizing all components in the model simultaneously achieves superior performance compared to optimizing only a subset of components in the objective function. It verifies that all three components play an important role in the image generation of the ERT-based tactile sensor.

Appendix C.5 Parameter Sensitivity Analysis

The parameters α and β play a crucial role in the regularization process of the proposed method, and the value of each parameter significantly influences the performance of the method. The parameter α controls the strength of the adversarial term in relation to the data fidelity term and the mask-based prior constraint. A higher value of α emphasizes the adversarial term, leading to a more faithful representation of the input data, but may also allow more noise and inconsistencies to be present in the reconstructed image. Conversely, a lower value of α can lead to a smoother reconstruction with reduced noise and artifacts. Similarly, the parameter β influences the balance between the adversarial term and the mask-based prior constraint incorporated into the model. A larger value of β increases the influence of the prior constraint, potentially biasing the reconstruction towards the imposed constraints or prior assumptions. Conversely, a smaller value of β gives less weight to the imposed constraints, allowing the adversarial term to have a stronger impact on the reconstruction.

To determine the optimal combination of the parameters α and β, we conduct extensive experiments using a grid search method. We explored a range of candidate values for α and β, specifically within the set $0.0001, 0.001, 0.01, 0.1, 1, 10, 100$. During the grid search process, the evaluation metrics AR and PE are selected to assess the performance of each combined method. The relations of evaluation metrics of AR and PE with respect to the parameters α and β are shown in Fig.C2. The imaging performance of the proposed method demonstrates stability across a wide range of parameter values for α and β. Through extensive experimentation, we find that the optimal performance of the method is achieved when α is set to 0.001 and β is set to 0.1.

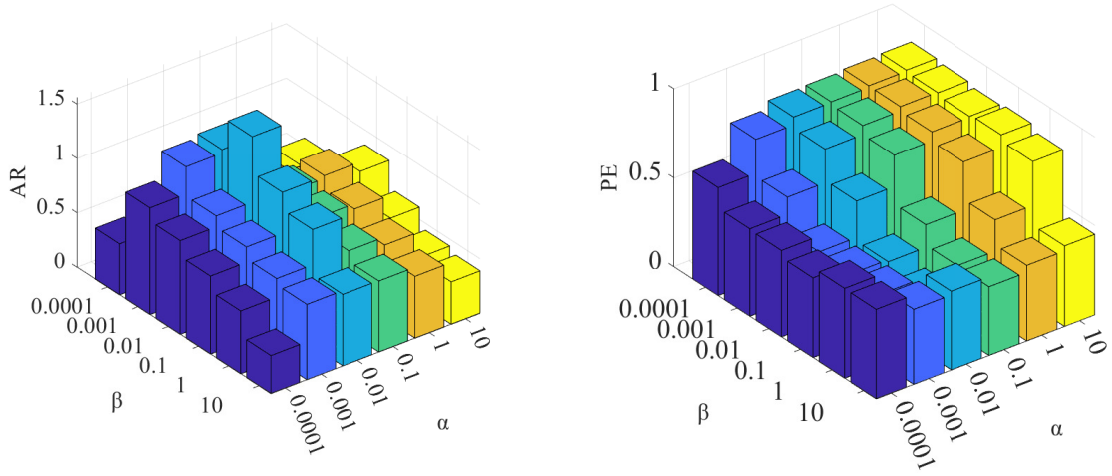

Figure C2 Relations of AR and PE with respect to parameters α and β. (a) Evolution of AR with respect to α and β; (b) Evolution of PE with respect to α and β.

Appendix D Performance Evaluation of The Sensor

To comprehensively evaluate the performance of the tactile sensor, a series of physical experiments are conducted to evaluate its sensing performance from four different perspectives.

Appendix D.1 Location Performance Analysis

In robotics applications, accurate and reliable location performance is of paramount importance for large-area tactile sensing systems. It enables robots to precisely localize tactile events, and effectively analyze and interpret sensed data, which is crucial for tasks such as object manipulation, grasping, and dexterous interaction with the environment.

To validate sensor location performance, experiments are conducted at 81 different positions on the sensor surface. Touch pressure is applied to these locations through point contact. To mitigate the effects of noise around the contact area, the thresholding technique is applied to process the conductivity image. Specifically, any values in the conductivity image below the 25% maximum value are considered noise and set to zero. Subsequently, the measured tactile position is calculated using a weighted centroid.

In the experiment, the positioning error is measured as the Euclidean distance between the target contact point and the calculated centroid. The experimental results of location error are shown in Fig.D1, where black dots represent the centroids of the target positions and blue asterisks denote the centroids of contact locations measured by the sensor. It can be observed that while there is some positioning error, the sensor is able to measure the contact positions, indicating the feasibility of the proposed sensor for touch localization.

Appendix D.2 Sensitivity Analysis

In order to further evaluate the performance of the sensor, the sensitivity of the sensor is analyzed. Due to the location dependence of the ERT-based tactile sensor, we divide the entire sensing area of the sensor into a 5×5 virtual array. Considering the symmetry of the sensing area, 25 sensing areas (1, 2, 3...., 25) are divided into six regions (I, II,..., VI), as shown in Fig.D2. Areas with the same color, that is, belong to the same region, have the same tactile sensitivity.

The relationship between the value of the conductivity change in the contact area and the contact force is shown in Fig.D3. For each sensing region, the peak conductivity values are averaged over the same sensing areas. It is evident that as the pressing force increases, the conductivity of the reconstructed map also exhibits a nonlinear increase within the range of 1-10N. This result clearly demonstrates the tactile sensor's ability to effectively detect tactile events and quantify contact force magnitude, confirming its suitability for force sensing applications.

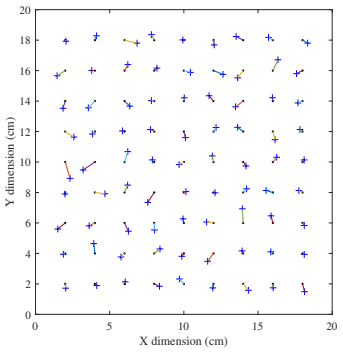

Figure D1 Localization error

Figure D2 Region division

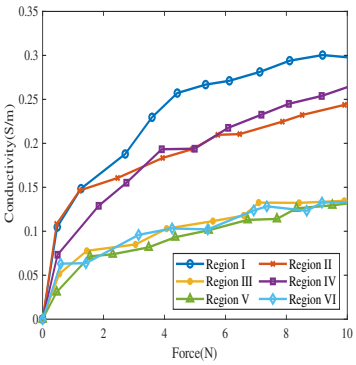

Figure D3 Sensitivity performance

Appendix D.3 Static Contact Target Detection

In order to verify the sensor's ability to effectively detect touch contacts and achieve multi-target detection, a set of experiments were conducted, including both single-target and multi-target detection scenarios. In these experiments, we use a 100g weight to apply a contact force to the sensor. The experimental results are presented in Fig.D4.

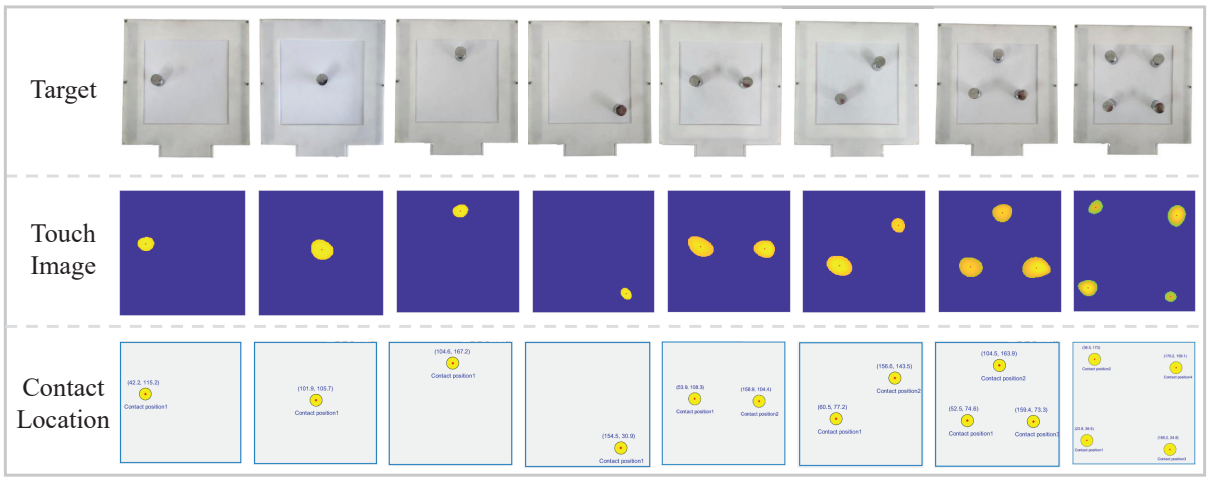

Figure D4 Single-target and multiple-target static touch detection.

For single-target contact, the touch images can clearly show the position of the pressure, indicating the effectiveness of the tactile sensor in pressure detection. In multi-target cases, although the reconstructed images have some artifacts, contact position can still be effectively located.

In fact, it is difficult for ERT-based sensors to achieve multi-point contact detection because applying pressure at multiple points usually leads to error collapse, greatly reducing the quality of image reconstruction. From the experimental results, the sensor can still effectively detect multi-target contacts, further verifying the effectiveness of the developed sensor.

Appendix D.4 Dynamic Continuity Touch Tracking

In addition to effectively detecting static contacts, a tactile sensor should also be able to track dynamic targets accurately. This ability is crucial in applications where real-time tracking of objects or tactile events is required. In robotic applications, the ability to accurately track and follow a moving object can facilitate tasks such as object manipulation, tracking human gestures, or interactive touch-based interfaces.

In order to verify the dynamic sensing performance of the sensor, experiments are conducted to track the location of a moving target on its surface. In particular, we dynamically interact with the tactile sensor by writing letters or numbers on the tactile sensor surface with fingers to explore the dynamic performance of the sensor, as illustrated in Fig.D5.

By conducting this experiment, it can be observed that the sensor is capable of accurately reproducing various characters and providing tactile information even in highly dynamic conditions. To demonstrate actual sensing performance, a set of results is shown in Fig. D5. This experiment showcases the sensor's ability to effectively track moving targets. This real-time detection and trajectory tracking capability further demonstrates the effectiveness of the sensor in dynamic scenarios.

Appendix D.5 Applications of Robotic Systems

In this section, we demonstrate the application of the proposed tactile sensor in a real robotic system. Specifically, we have integrated the developed tactile sensor onto the UR5 robotic arm to perform obstacle avoidance experiments.

As depicted in Figure D6, the robotic arm is tasked with grasping an object and attempting to place it at a target location. However, there is a possibility of randomly placed obstacles along the robot's trajectory towards the object. The robot does not have prior knowledge of the obstacle locations in the environment, but it can rely on the tactile information provided by the developed sensors when its arm comes into contact with an obstacle. By utilizing this tactile information,

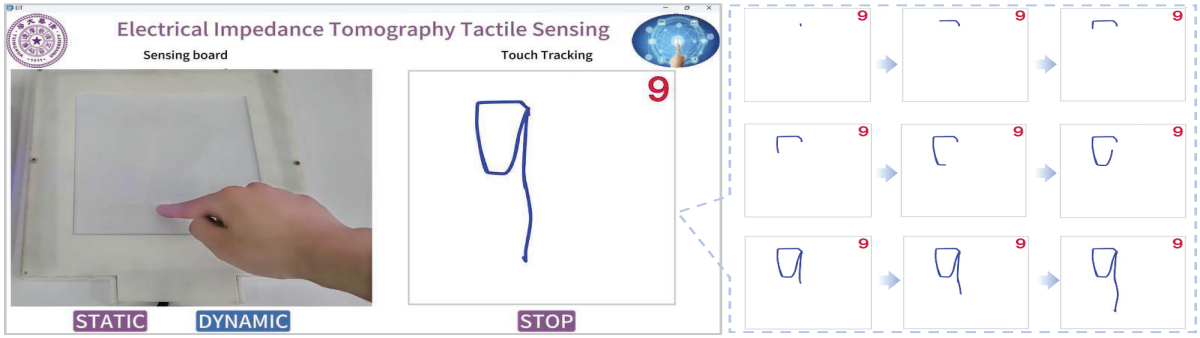

Figure D5 Schematic diagram of touch tracking based on the proposed sensor.

the robot can dynamically adjust its movement strategy in real-time to avoid any potential damage to itself and the surrounding environment. In this experiment, the initial grasping trajectory has been pre-programmed.

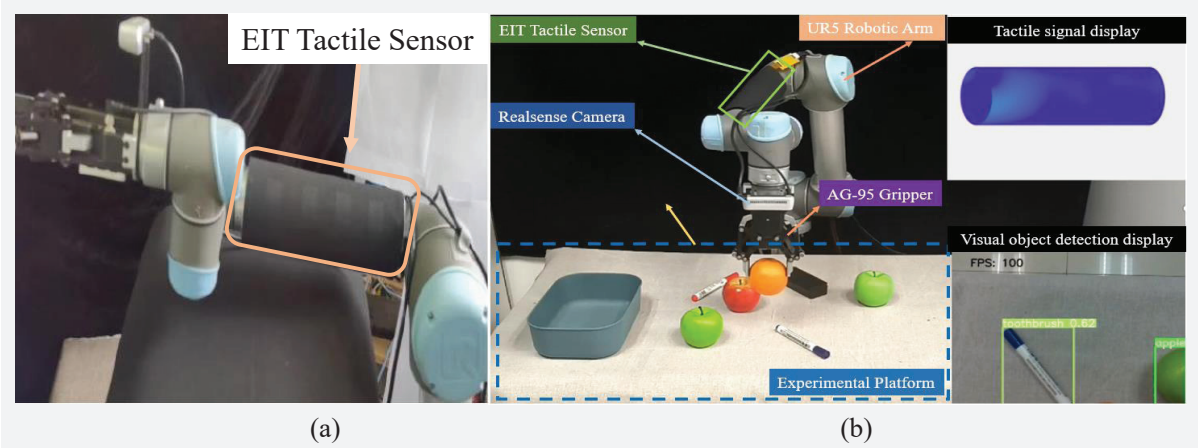

Figure D6 Robotic arm obstacle avoidance experiment. (a) The sensor is installed on the robotic arm; (b) An obstacle is randomly placed on the trajectory of the grasping task.

To detect collisions accurately, we propose a method to extract contact signals using the tactile image. In this approach, we formulate the collision detection problem as a binary classification task. The magnitude of the contact force is directly related to the conductivity amplitude observed in the tactile image. Therefore, we extract the maximum value from the conductivity image to determine whether contact exists. The collision detection signal (CD) is defined as follows:

$$CD = \begin{cases} 1, \sigma_{\max} \geqslant \varepsilon_{\sigma} \\ 0, otherwise \end{cases} \tag{D1}$$

where $\sigma_{\max}$ is maximum value of the conductivity image and ε_{σ} represents contact detection threshold. Since the sensing performance of ERT-based sensors is position-dependent, we set different contact thresholds for six different contact areas to achieve the best detection performance. When the conductivity peak exceeds the set threshold, it indicates that the robot has encountered an obstacle and needs to avoid it.

By integrating tactile sensors into robotic systems and exploiting the obtained tactile feedback, we demonstrate the sensor's ability to enhance the robot's perception and decision-making processes, allowing it to effectively avoid obstacles in real-world scenarios.

References

1 D. Silvera-Tawil, D. Rye, M. Soleimani, and M. Velonaki, "Electrical impedance tomography for artificial sensitive robotic skin: A review," *IEEE Sensors Journal*, vol. 15, no. 4, pp. 2001–2016, 2014.
2 W. Zheng, H. Liu, D. Guo, and W. Yang, "Adaptive optimal electrical resistance tomography for large-area tactile sensing," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 10 338–10 344.
3 O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*. Springer, 2015, pp. 234–241.
4 P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1125–1134.
5 V. Nguyen, T. F. Yago Vicente, M. Zhao, M. Hoai, and D. Samaras, "Shadow detection with conditional generative adversarial networks," in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 4510–4518.
6 H. Park, K. Park, S. Mo, and J. Kim, "Deep neural network based electrical impedance tomographic sensing methodology for large-area robotic tactile sensing," *IEEE Transactions on Robotics*, vol. 37, no. 5, pp. 1570–1583, 2021.
7 K. Park, H. Park, H. Lee, S. Park, and J. Kim, "An ert-based robotic skin with sparsely distributed electrodes: Structure, fabrication, and dnn-based signal processing," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 1617–1624.