2026 IEEE International Conference on Robotics and Automation (ICRA 2026)
June 1-5, 2026. Vienna, Austria

Touch with Insight: Physics-Aware Data-Driven Learning for EIT-Based Tactile Sensing

Kiyanoush Nazari[1], Yunqi Huang[1], David Hardman[2], Fumiya Iida[2], and Thomas George Thuruthel[1]

***Abstract*— Tactile sensing is essential for enabling dexterous robotic manipulation, yet estimating contact states such as location and force from high-dimensional sensor measurements remains challenging due to noise and complex nonlinear mappings between raw signals and physical interaction states. In this work, we propose a physics-informed contact modeling framework that combines the flexibility of deep models with inductive biases from physical modeling. Focusing on electrical impedance tomography (EIT) tactile skins, our approach incorporates knowledge of the EIT forward model by regularizing neural estimators with a latent-space consistency constraint, stabilizing the ill-posed inverse mapping from voltages to contact states. To support robust training and evaluation, we also develop a high-fidelity simulation pipeline that incorporates key hardware imperfections to better bridge the sim-to-real gap. We benchmark multiple architectures—including multilayer perceptrons, convolutional networks, Transformer-based models, and autoencoder regressors—on both real and synthetic datasets. Results show that the proposed hybrid approach consistently improves estimation accuracy, particularly for force prediction, and generalizes across domains. These findings highlight the value of embedding physical priors into learning pipelines for reliable tactile state estimation in robotic manipulation. (video, code)**

I. Introduction

Tactile sensing is key for enabling dexterous robotic interaction, complementing vision and proprioception by providing direct information about contact events such as location, force, and slip [1], [2]. A central challenge in tactile perception is to infer low-dimensional contact states from high-dimensional, often noisy sensor measurements. This challenge arises across a variety of tactile modalities, including vision-based, magnetic, acoustic, and electrical methods, where complex non-linear mappings are required to translate raw sensor signals into interpretable physical quantities.

Among these modalities, Electrical Impedance Tomography (EIT) has recently emerged as a promising tactile sensing approach [3], [4], [5], [6]. Originally developed as a medical imaging technique for monitoring tissue conductivity [7], EIT-based tactile sensors exploit the principle that applying small alternating currents through boundary electrodes into a conductive medium results in voltage measurements that encode changes in internal conductivity due to contact. These boundary voltages can, in turn, be used to infer tactile information such as contact location and applied pressure [8].

*This work was supported by The Advanced Research + Invention Agency (ARIA) as part of Robot Dexterity Programme.

[1] Department of Computer Science, University College London, London, UK.
[2] Department of Engineering, University of Cambridge, Cambridge, UK.

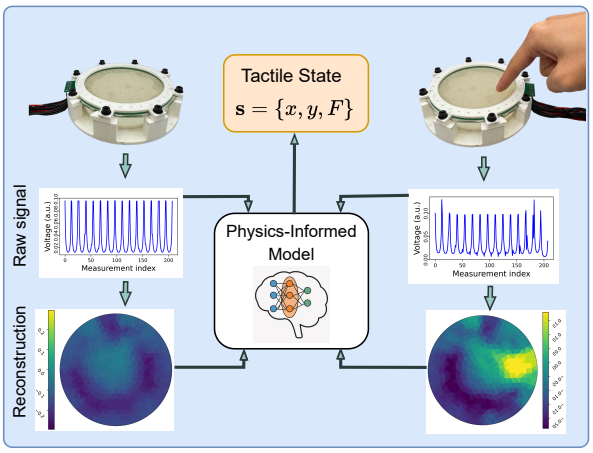


Fig. 1: A hydrogel-based EIT tactile skin with 16 boundary electrodes is indented by a conductive touch, producing boundary voltage measurements (raw signals). These signals are mapped through a physics-informed model to estimate the tactile state $s = \{x, y, F\}$, comprising contact location and normal force. For illustration, reconstructed conductivity maps highlight localized changes under contact.

Compared with vision- or magnet-based tactile sensors, EIT offers unique advantages as it requires minimal embedded electronics, supports large-area and flexible geometries, and enables low-cost fabrication of complex-shaped tactile skins [9], [10], [6]. Despite these benefits, accurate tactile state estimation from EIT data remains highly challenging. The measurements are sensitive to electrode variability, material properties, and noise in current injection and voltage readout [11], [12]. Furthermore, the inverse problem of EIT is fundamentally ill-posed and non-linear, where small variations at the boundary may lead to disproportionately large and unstable differences in reconstruction [13].

Classical model-based EIT reconstruction algorithms [14], [15] can, in principle, recover conductivity distributions, but their spatial resolution is limited, restricting their ability to provide precise tactile information. To overcome this, recent research has explored learning-based methods that enhance conductivity image reconstruction with data-driven regularization [16], [17], [18], or bypass reconstruction entirely by directly mapping boundary voltage measurements to contact states such as location and force [19].

While previous methods have been proposed to improve EIT-based tactile reconstruction, most existing approaches emphasize estimating the contact region within a coarse force

range rather than achieving accurate force mapping [20], [21], [22]. This limitation reflects a broader challenge in tactile sensing that is, the mapping from raw high-dimensional sensor signals to precise force states is highly nonlinear and often non-unique. In the case of EIT, this manifests as the complex relationship between applied forces, induced conductivity changes, and boundary voltage measurements. Whereas contact location can often be inferred from spatial patterns in measured signals, force estimation requires capturing subtle magnitude variations that are sensitive to material properties, contact geometry, and sensor–electrode configuration [23], [24]. Purely data-driven regression models, which treat this mapping as a black-box, typically struggle to generalize when confronted with variations in noise, sensor properties, or boundary conditions [25], [26]. To address this, it is advantageous to combine model-based and learning-based perspectives, where data-driven models provide flexibility and model-based structures offer inductive biases grounded in physics [27].

In this work, we propose a hybrid strategy that conditions the regression process on a secondary consistency module. Instead of relying solely on the ill-posed reverse problem of mapping voltages to contact states, we also learn a complementary mapping from contact states to a low-dimensional latent representation of the observation space. For EIT, this latent space is learned from voltage measurements via an autoencoder, and the secondary module acts as a forward-inspired regularizer that enforces consistency between predicted tactile states and the manifold of physically plausible voltage patterns. By coupling the regression with this latent-space constraint, the model respects the underlying physics while retaining the expressive power of learning-based methods. This hybrid, physics-informed regularization provides a principled way to stabilize complex tactile state estimation tasks, particularly in settings where raw sensor-to-state mappings are non-unique and highly nonlinear.

However, implementing such physics-informed approaches to learn the complex mappings between physical parameters and sensor responses requires substantial training data. While real-world experiments are essential for collecting data, simulation remains an important tool for scaling data generation and validating new approaches [28], [29]. Existing simulation tools like EIDORS [30] and pyEIT [31] provide forward and inverse problem solvers based on finite element methods (FEM) to model conductivity changes and distributions within the sensing medium. These simulations enable rapid data generation, facilitating comprehensive training and evaluation of learning-based methods. However, simulated environments often diverge from real-world measurements due to electrode contact impedance variability, measurement noise, crosstalk, and geometric mismatches, making direct transfer from simulation to hardware difficult [32]. To address this, we develop a new data generation pipeline that explicitly incorporates these sim-to-real gap factors, producing synthetic EIT tactile data that more closely mirrors the behavior of physical sensors.

In this work, (i) we present a physics-informed learning framework for EIT-based tactile sensing, where neural estimators are regularized through consistency with a latent representation informed by the EIT forward model. This hybrid approach leverages physical structure to improve force estimation while retaining the flexibility of deep learning. (ii) We conduct a comprehensive benchmark of learning models—including MLPs, CNNs, CNN–Transformer hybrids, and Transformer-based architectures—for both contact location and force estimation, providing a clear comparison across model families. (iii) To support robust evaluation, we also develop a high-fidelity simulation pipeline that incorporates key sim-to-real effects, generating synthetic data that closely reflects real sensor behavior. Across both real and simulated datasets, we show that the proposed physics-informed pipeline consistently enhances force estimation accuracy, particularly in data-driven baselines such as MLPs. Together, these results highlight the value of integrating physical priors into neural tactile estimators and establish a methodology applicable to diverse architectures and sensing contexts.

II. Methodology

A. Electrical Impedance Tomography (EIT) Formulation

In an EIT-based tactile sensor, a set of boundary electrodes is placed around a conductive medium (e.g., a circular hydrogel skin shown in Fig.1). By injecting small alternating currents through pairs of electrodes and measuring the resulting voltages on the remaining electrodes, one can infer conductivity changes inside the medium caused by contact. In this work, we consider the common configuration of 16 equally spaced electrodes arranged on the boundary of a circular domain. A sinusoidal current

$$I(t) = I_0 \sin(\omega t), \tag{1}$$

with amplitude I_0 and angular frequency ω is sequentially applied between chosen pairs of electrodes according to a measurement protocol (e.g., adjacent or opposite current injection). For each current injection, the measured electrode voltages are collected into a vector

$$V = \{V_1, V_2, \ldots, V_M\} \in \mathbb{R}^M, \tag{2}$$

where M denotes the number of independent voltage measurements. Under the commonly used *adjacent current injection protocol*, each of the n_{el} electrodes is used once as a source-sink pair, while voltages are measured on the remaining $n_{\text{el}} - 3$ electrodes. Thus, the total number of independent measurements is $M = n_{\text{el}} \times (n_{\text{el}} - 3)$. For $n_{\text{el}} = 16$, this gives $M = 208$ independent voltage measurements. Fig.,1 illustrates a representative no-touch voltage measurement from the sensor, comprising 208 boundary voltage values acquired across all electrode pairs.

The governing equation of EIT in the conductive domain $\Omega \subset \mathbb{R}^2$ is the elliptic partial differential equation

$$\nabla \cdot (\sigma(\mathbf{r}) \nabla u(\mathbf{r})) = 0, \quad \mathbf{r} \in \Omega, \tag{3}$$

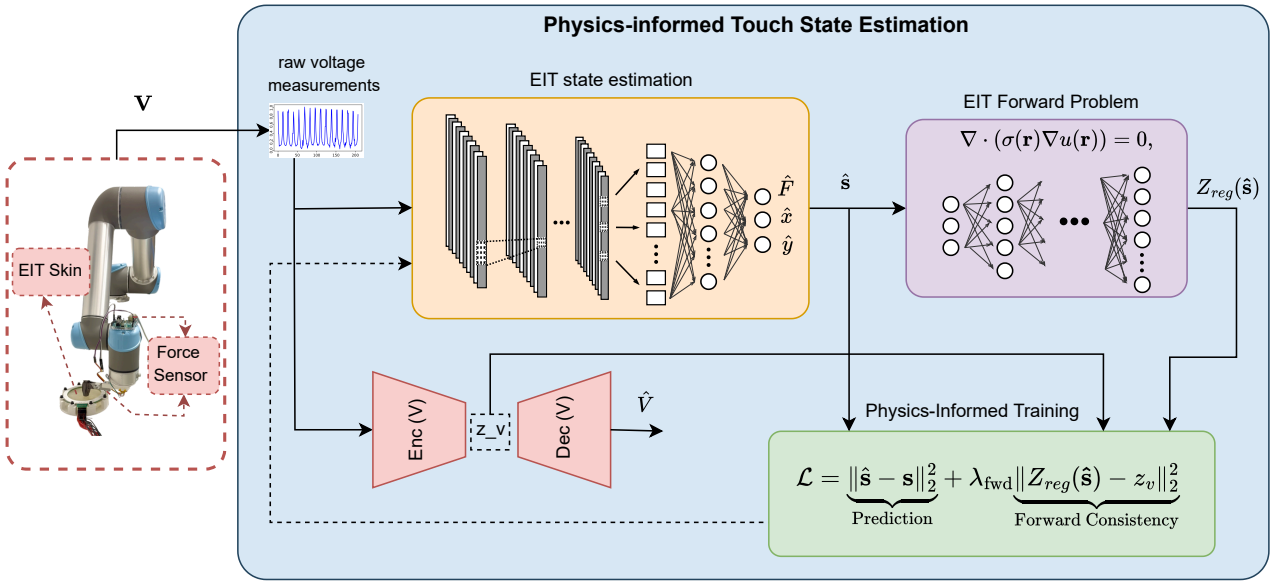


Fig. 2: Overview of the proposed physics-informed touch state estimation framework. Raw voltage measurements from the EIT skin are used to predict contact states $(\hat{x}, \hat{y}, \hat{F})$ through a data-driven network. In parallel, a voltage autoencoder learns a latent representation of the measurement manifold shaped by EIT physics. During training, predicted states are passed through the EIT forward model, and a consistency loss aligns the resulting latent codes with those from the voltage encoder. This joint objective combines prediction accuracy with physics-based regularization, improving robustness and accuracy of tactile state estimation.

where $\sigma(\mathbf{r})$ is the spatial conductivity distribution and $u(\mathbf{r})$ is the electric potential field. On the boundary $\partial\Omega$, Neumann boundary conditions enforce the applied current through electrodes, and Dirichlet conditions ensure voltage reference constraints.

The **forward problem** consists of solving for the potential $u(\mathbf{r})$ and electrode voltages V given a known conductivity distribution $\sigma(\mathbf{r})$. Conversely, the **inverse problem** seeks to reconstruct $\sigma(\mathbf{r})$ (or its low-dimensional representation) as shown in Fig.1, from the measured voltages V, which is ill-posed and highly sensitive to noise.

B. EIT Tactile State Estimation

In tactile sensing with EIT, our goal is to infer the *contact state*

$$\mathbf{s} \;=\; [\,x,\; y,\; F\,]^{\top} \in \mathbb{R}^3,$$

consisting of the contact location (x, y) on the sensor and the normal contact force F, from a boundary voltage measurement vector $\mathbf{V}$ introduced in Eq.(2).

We assume a supervised dataset $\mathcal{D} \;=\; \big\{\,(\mathbf{V}_i,\; \mathbf{s}_i)\,\big\}_{i=1}^{N}$, where each $\mathbf{V}_i$ is a stacked vector of electrode voltages from one EIT frame and $\mathbf{s}_i$ is the corresponding contact state. To reduce static offsets and improve conditioning, we subtract a no-touch (baseline) voltage

$$\tilde{\mathbf{V}}_i \;=\; \mathbf{V}_i - \mathbf{V}_{\text{ref}}, \qquad \mathbf{V}_{\text{ref}} \;=\; \frac{1}{|\mathcal{R}|}\sum_{j\in\mathcal{B}} \mathbf{V}_j,$$

with $\mathcal{R}$ the set of reference (zero-force) frames. The learning objective is to find a regression map

$$f_\theta :\; \mathbb{R}^M \to \mathbb{R}^3, \qquad \hat{\mathbf{s}} \;=\; f_\theta(\hat{\mathbf{V}}),$$

parameterized by θ. The baseline supervised loss on a mini-batch $\mathcal{B}$ is the (optionally weighted) MSE

$$\mathcal{L}_{\text{sup}}(\theta) \;=\; \frac{1}{|\mathcal{B}|}\sum_{i\in\mathcal{B}} \Big(\|\hat{x}_i - x_i\|_2^2 + \|\hat{y}_i - y_i\|_2^2 + w_F \,\|\hat{F}_i - F_i\|_2^2 \Big), \tag{4}$$

where (x_i, y_i, F_i) are standardized ground-truth labels and $w_F \geq 1$ upweights force.

If ground-truth locations are provided in a robot frame (u, v), we align them to the EIT unit-disk coordinates via a similarity transform

$$\begin{bmatrix} x \\ y \end{bmatrix} \;=\; s\,\mathbf{R}\begin{bmatrix} u \\ v \end{bmatrix} + \mathbf{t}, \quad \mathbf{R} \in \mathrm{SO}(2),\; s \in \mathbb{R}_+,\; \mathbf{t} \in \mathbb{R}^2,$$

estimated once from calibration samples (e.g., using contact "hotspot" angles). The aligned (x, y) are then used in Eq.(4). The tactile state estimation problem is stated as, given $\mathcal{D}$, learn $\theta^\star = \arg\min_\theta \mathcal{L}_{\text{sup}}(\theta)$, yielding a model that maps voltages to contact location and force with high accuracy and generalization.

C. Physics-Informed Learning via a Latent Voltage Prior

Directly penalizing mismatch in the EIT measurement space (e.g., through a differentiable forward solver) can be brittle and computationally heavy. Instead, we encode the *physics of EIT* implicitly by learning a compact latent manifold of boundary voltages from data. Constraining the

predictor to produce contact states whose implied latent voltages lie on this manifold, steers learning toward physically plausible solutions without requiring a PDE-in-the-loop during training.

Our approach, shown in Fig.2, augments a supervised predictor with a physics-informed latent regularizer. The pipeline consists of three modules trained sequentially:

1) **Voltage autoencoder** (E_v, D_v)**.** A low-dimensional latent representation $\mathbf{z} = E_v(\hat{\mathbf{V}}) \in \mathbb{R}^K$ is learned from standardized, baseline-subtracted voltages $\hat{\mathbf{V}} \in \mathbb{R}^M$ by minimizing the reconstruction loss

$$\mathcal{L}_{\mathrm{AE}} = \frac{1}{N}\sum_{i=1}^{N} \left\| D_v(E_v(\hat{\mathbf{V}}_i)) - \hat{\mathbf{V}}_i \right\|_2^2, \qquad K \ll M. \quad (5)$$

This defines a manifold of physically plausible voltage patterns.

2) **State-to-latent regressor** g_ϕ**.** Given contact states $\mathbf{s} = [x, y, F]^\top \in \mathbb{R}^3$, we learn a mapping into the same latent space by aligning $g_\phi(\mathbf{s})$ with the teacher embedding $E_v(\hat{\mathbf{V}})$:

$$\mathcal{L}_{\mathrm{ZReg}} = \frac{1}{N}\sum_{i=1}^{N} \left\| g_\phi(\mathbf{s}_i) - E_v(\hat{\mathbf{V}}_i) \right\|_2^2. \quad (6)$$

The regressor function g_ϕ mimics the EIT forward problem to learn a mapping between contact states to corresponding latent voltage representation. This function helps the Predictor function f_θ to predict more plausible tactile states from raw voltage measurements.

3) **Predictor** f_θ**.** The main predictor maps voltages to contact states, $f_\theta : \mathbb{R}^M \to \mathbb{R}^3$. It is trained with a supervised loss defined in Eq.(4) and a latent physics regularizer that enforces consistency between the predicted state's latent code and the one extracted from measured voltages:

$$\mathcal{L}_{\mathrm{latent}} = \frac{1}{|\mathcal{B}|}\sum_{i\in\mathcal{B}} \left\| g_\phi(f_\theta(\hat{\mathbf{V}}_i)) - E_v(\hat{\mathbf{V}}_i) \right\|_2^2. \quad (7)$$

The final training objective is

$$\mathcal{L}_{\mathrm{total}} = \mathcal{L}_{\mathrm{sup}} + \lambda_{\mathrm{lat}}\, \mathcal{L}_{\mathrm{latent}}. \quad (8)$$

The design rationale for this pipeline consists of *(i)* The voltage autoencoder learns the *measurement manifold* shaped by EIT physics (electrode array, drive/measure protocol, forward map, and typical conductivity perturbations). Matching the predicted state's latent to the voltage-derived latent (Eq. 7) penalizes predictions that would imply voltages off this manifold, providing a physics-shaped prior at low computational cost. *(ii)* In many EIT layouts, multiple (x, y, F) combinations can produce similar boundary voltages. The latent prior encourages *joint* consistency of (x, y, F) with the entire voltage pattern, which we find particularly helpful for force. *(iii)* Because E_v is trained across diverse operating points, $\mathcal{L}_{\mathrm{latent}}$ acts as a stabilizer against sensor noise and slow baseline drift.

Our training protocol forms as training (E_v, D_v) with Eq. (5) first, then train g_ϕ with Eq. (6) using teacher targets $\mathbf{z}^\star$. Finally, we train f_θ with Eq. (8). At test time we use f_θ alone; (E_v, g_ϕ) are not needed for inference.

D. Dataset Generation

1) Real-world tactile dataset: To evaluate our approach on physical measurements, we constructed a custom EIT-based tactile sensing setup. The tactile sensor consists of a hydrogel-based conductive skin fabricated by combining gelatin, glycerol, water, citric acid, and salt (NaCl) in a precise ratio of 1:1.5:2.5:0.2:0.1 by weight, following the method described by Hardman et al. [33]. The skin, with a radius of 8 cm and thickness of 3 mm, was mounted on a rigid supporting structure to ensure mechanical stability during indentation experiments. Sixteen electrodes were implemented on a custom PCB-based circuit that provided precise electrode positioning. The electrodes were equally distributed along the circumference of the circular skin, ensuring uniform current injection and voltage measurement points. This layout is illustrated in Fig. 2.

Data collection was performed using a UR5 robotic manipulator. A cylindrical probe was attached to the end-effector, with a calibrated force sensor (Sparkfun TAL220 Series Parallel Beam Load Cell) mounted in series to record the ground-truth contact force. To study the effect of contact geometry, three probes of different diameters were used: 7 mm, 10 mm, and 15 mm. The probing locations followed a predefined spatial pattern covering the sensor surface consisting of 192 contact points. At each contact location, the UR5 positioned the probe above the target point and performed a vertical indentation in 1 mm increments. The indentation continued until the measured contact force reached a target value of 8–10 N, ensuring a range of indentation depths and forces were captured. For every indentation step, the system recorded: (i) the EIT voltage measurements across all electrode pairs, (ii) the ground-truth contact location from the robot kinematics stated in UR5 base frame, and (iii) the ground-truth contact force from the force sensor. This procedure yielded a dataset of paired samples $\{(\mathbf{V}, \mathbf{s})\}$, where $\mathbf{V}$ are baseline-subtracted voltages and $\mathbf{s} = [x, y, F]^\top$ are the corresponding ground-truth states. This dataset therefore provides diverse coverage across contact locations, indentation depths, and probe geometries, enabling robust evaluation of learning-based and physics-informed state estimation methods.

2) PyEIT simulated dataset with sim-to-real gap injection: To complement the real dataset, we generated a synthetic EIT tactile dataset using the PyEIT forward solver on a 16–electrode circular domain. Touch interactions were modeled as circular anomalies of radius $r_{\text{touch}} = 0.1$ placed on a polar grid of (r, θ) locations, with contact force $F \in [0, 1]$ mapped to inclusion permittivity via $\varepsilon(F) = 1 + 9F$. For each sample we record baseline–subtracted voltages $\Delta\mathbf{v} = \mathbf{v}_{\text{touch}} - \mathbf{v}_0 \in \mathbb{R}^M$, where M is the number of channels under adjacent stimulation (e.g., $M = 208$ for 16 electrodes with the standard parser).

To reduce the gap between ideal simulations and real measurements, we injected five dominant sim–to–real effects in EIT sensing. We model the measured voltages as

$$\tilde{\mathbf{v}} \;=\; \mathbf{C}\left(\mathbf{g}(F) \odot \mathbf{v}\right) \;+\; \boldsymbol{\eta}, \quad (9)$$

where $\mathbf{g}$ are per–channel gains (contact impedance variability), $\mathbf{C}$ is a crosstalk mixing matrix, $\odot$ denotes element–wise product, and $\boldsymbol{\eta}$ is additive noise.

Electrode variability and force drift. Gains are drawn as $g_i = g_{M-1-i} = 1 + \delta_i$, $\delta_i \sim \mathcal{N}(0, \sigma_g^2)$, with an optional force–dependent modulation

$$g_i(F) = g_i\big(1 + \alpha(F - 0.5)\big).$$

This linear form provides a first–order approximation of contact–impedance changes with applied force, centered at the mid–range ($F = 0.5$) for stability. We set $\sigma_g = 0.08$, yielding $\sim$8% random channel deviations consistent with real baselines, and $\alpha = 0.1$, allowing up to 10% gain drift across the normalized force range. Values were tuned to reflect realistic sensor behavior without destabilizing the forward solver.

Crosstalk. Neighboring channel leakage is modeled by a tridiagonal wrap–around matrix

$$\mathbf{C} = (1 - 2\varepsilon)\mathbf{I} + \varepsilon(\mathbf{S}_\mathrm{L} + \mathbf{S}_\mathrm{R}),$$

where ε controls the coupling strength (set to 0.03 in our experiments). Here, $\mathbf{I}$ is the $M \times M$ identity matrix, $\mathbf{S}_\mathrm{L}$ is a cyclic left–shift operator (ones on the subdiagonal and in the top–right corner), and $\mathbf{S}_\mathrm{R}$ is its transpose, a cyclic right–shift operator. This structure ensures that each channel retains most of its own signal $(1-2\varepsilon)$, while a small fraction ε of its amplitude leaks symmetrically to the immediate neighbors, consistent with coupling patterns observed in real hardware.

Measurement noise. We add Gaussian noise $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$ with $\sigma = \sigma_\mathrm{rel}\,\mathrm{med}(|\mathbf{v}|)$, where $\sigma_\mathrm{rel} = 0.02$ specifies the noise strength relative to the median channel magnitude. This choice injects low–level variability comparable to that seen in repeated no–touch measurements on the real sensor.

Geometry mismatch and electrode jitter. To capture mismatch between the reconstruction model and the true plant, we scale the mesh boundary to an ellipse $(x, y) \mapsto (s_x x, s_y y)$ with $s_x, s_y \sim \mathcal{N}(1, \sigma_\mathrm{ell}^2)$, and perturb electrode angles $\phi_k' = \phi_k + \delta\phi_k$, $\delta\phi_k \sim \mathcal{N}(0, \sigma_\phi^2)$. We set $\sigma_\mathrm{ell} = 0.05$ (5% boundary scaling) and $\sigma_\phi = 0.5°$, reflecting small geometric distortions observed in practice.

This compact distortion model produces baseline–subtracted features

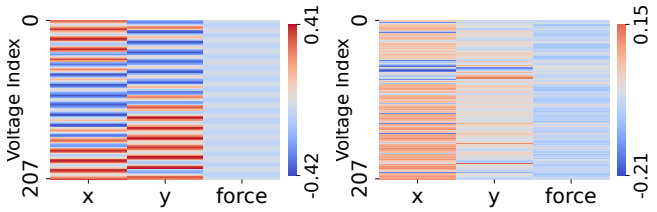
$$\Delta\tilde{\mathbf{v}} = \tilde{\mathbf{v}}_\mathrm{touch} - \tilde{\mathbf{v}}_0,$$

which more closely resemble real measurements. The resulting dataset covers a polar grid of 192 touch locations and a diverse distribution of forces, with distortions applied consistently within sessions to mimic hardware biases.

III. Results

A. Experimental Setup

For the real-world experiments with the hydrogel-based EIT skin, we collected approximately $N_\mathrm{real} \approx 3{,}900$ samples. At each contact location, the UR5 robot applied incremental indentations until a target force was reached, resulting in contact forces ranging from 0.5 N up to 10 N. We drove the skin with a single–tone sinusoidal excitation at $f_\mathrm{exc} = 10$ kHz (AD5930 clocked at 50 MHz), using adjacent drive. The drive and measurement gains (AD5270 digital rheostats) were auto–tuned to keep the differential voltage within the ADC's linear range (≈ 0.5–$2.1\,\mathrm{V_{p\text{–}p}}$).



(a) **simulated** dataset. (b) **real** dataset.

Fig. 3: Pearson correlation coefficients between boundary voltage features and target variables (contact location x, y, and force f). Stronger correlations are visible for x and y, while force exhibits weaker and more diffuse correlations, explaining the relative difficulty of force estimation.

For the simulated dataset, we reproduced a comparable contact point pattern on a circular 16-electrode sensor using the PyEIT library. Unlike the real case, the simulation is parameterized in terms of permittivity changes rather than direct force values, though the range was scaled to cover an equivalent dynamic range with finer increment steps. The final dataset contained approximately $N_\mathrm{sim} \approx 50{,}000$ samples, incorporating distortions that emulate key sources of sim-to-real gap.

All models were trained in PyTorch on a laptop with an Intel Core Ultra 7 CPU and NVIDIA GeForce RTX 4050 (6 GB), with 16 GB system RAM. Data were standardized channel-wise, split into 80% training and 20% testing partitions, with a validation set drawn from 20% of the training portion to monitor generalization performance during training and optimized using Adam unless otherwise stated.

B. Correlation Analysis of Voltage and Contact States

Before training predictive models, we performed an exploratory correlation analysis between the measured voltage features and the target outputs (contact location x, y, and contact force f). Such analysis is informative in this context because it highlights the degree of linear association between the high-dimensional EIT signals and the underlying contact states, thereby offering insight into the relative difficulty of each regression task. Fig. 3 shows the Pearson correlation coefficients computed between each voltage channel and the three target variables, for both the real and simulated datasets. Two consistent trends can be observed. First, the correlations between voltages and the spatial coordinates (x and y) are relatively strong, suggesting that contact location information is more directly encoded in the boundary voltage patterns. Second, the correlations with contact force are substantially weaker and more diffuse, reflecting the fact that force estimation from EIT is a more complex mapping problem. This observation is consistent with our later quantitative results, where all models achieve higher accuracy in

TABLE I: Prediction performance across baseline and physics-informed models on real and simulated datasets. Higher R^2 and lower MAE show better performance.

	Real Dataset						Simulated Dataset					
Model	R_x^2	MAE_x	R_y^2	MAE_y	R_f^2	MAE_f	R_x^2	MAE_x	R_y^2	MAE_y	R_f^2	MAE_f
SVR	0.99	0.7	0.99	0.7	0.76	0.88	0.99	1.92	0.99	1.97	0.82	0.62
MLP	0.99	0.8	0.99	0.7	0.82	0.75	0.99	1.73	0.99	1.71	0.89	0.48
CNN	0.99	1.0	0.99	1.1	0.79	0.84	0.99	2.0	0.99	1.97	0.90	0.43
Transformer	0.98	1.3	0.98	1.4	0.64	1.07	0.99	2.29	0.99	2.83	0.90	0.47
Autoencoder-MLP	0.98	1.3	0.98	1.2	0.65	1.05	0.99	1.98	0.99	1.91	0.91	0.45
CNN-Transformer	0.98	1.3	0.99	1.2	0.62	1.08	0.99	2.57	0.99	2.49	0.87	0.50
Physics-Informed (MLP)	0.99	1.1	0.99	1.1	**0.84**	**0.72**	0.99	1.3	0.99	1.3	**0.92**	**0.41**

predicting contact location than in predicting contact force.

C. Baseline Models

To benchmark EIT state estimation, we evaluated a range of machine learning architectures spanning classical regressors, spatially structured networks, and attention-based models. Support vector regression (SVR) with an RBF kernel was included as a classical non-linear baseline, providing a contrast to deep learning approaches. MLPs serve as a standard baseline, with three hidden layers (256, 128, 64 neurons, ReLU activations) and a linear output for (x, y, F). CNNs were implemented as 1D convolutional models with three blocks (kernel size 3; channels 64, 128, 256) followed by max-pooling, batch normalization, and two dense layers (128, 64). Transformers employed sinusoidal positional encoding, two encoder layers with four attention heads, a feed-forward block (128 hidden units), and a regression head. For hybrid designs, we tested an autoencoder–MLP, where a bottleneck autoencoder ($256 \rightarrow 64 \rightarrow$ latent $z = 16 \rightarrow 64 \rightarrow 256$) compressed voltages before MLP regression, and a CNN–Transformer, where convolutional layers extracted local features that were then processed by a transformer encoder and dense output.

This suite of baselines allows comparison between purely data-driven, spatially aware, and globally attentive architectures for EIT regression.

D. Tactile State Estimation Results

To evaluate the performance of the baseline models and the proposed physics-informed pipeline, we report results using two complementary metrics: the coefficient of determination (R^2) and the mean absolute error (MAE). The R^2 score indicates the proportion of variance in the ground-truth target explained by the model, with higher values denoting better predictive performance. MAE measures the average absolute deviation between the predicted and true values, providing an interpretable scale in the same units as the output variables. Together, these metrics capture both the goodness of fit and the average prediction accuracy.

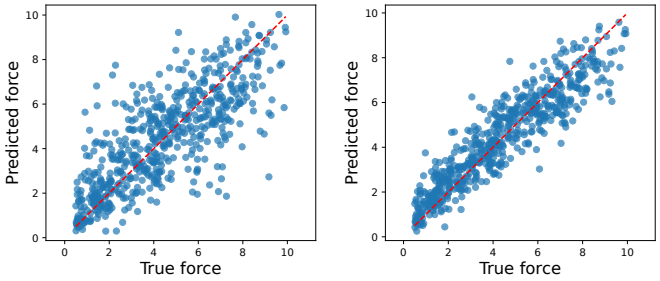
Table I summarizes the prediction results for all models across both real and simulated datasets. We report values for the three outputs of interest: contact position in the x- and y-axes, and the contact force.

Table I shows that contact location prediction (x and y) achieves consistently high performance across all baseline models, with R^2 values close to 0.99. This trend is expected, as our correlation analysis in Fig.3 revealed that boundary voltage patterns are strongly associated with the spatial location of contact. In fact, contact location forms a relatively straightforward regression mapping from the high-dimensional voltage space, which explains the uniformly strong performance across model families. For closer comparison, MAE is a more sensitive metric; interestingly, SVR and the MLP yield the lowest errors for location prediction, though the differences across models remain small.

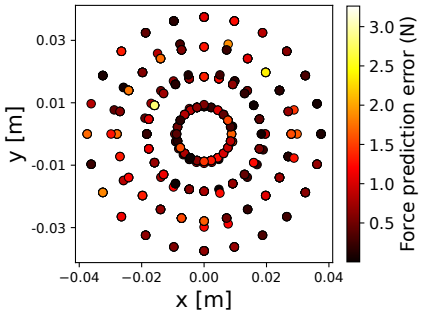
In contrast, contact force prediction presents a more challenging task. As anticipated, R^2 values drop considerably compared to location prediction, reflecting the weaker correlation between boundary voltages and applied force. Among the baseline models, the MLP achieves the strongest force estimation performance. This may be attributed to two factors: (i) the fully connected architecture of the MLP can flexibly approximate nonlinear mappings without imposing spatial locality biases (as in CNNs) or sequential dependencies (as in Transformers), which may not align with EIT data structure; and (ii) the relatively moderate dataset size favors MLPs, as larger architectures like Transformers require more samples to generalize effectively.

Most importantly, the proposed physics-informed MLP achieves the best overall force estimation performance. Compared to the baseline MLP, it yields a higher R_f^2 (0.84 vs. 0.82) and a lower MAE (0.72 vs. 0.75), demonstrating the benefit of embedding physics priors into the learning pipeline. These gains are consistent and highlight that physics-informed regularization helps the network capture latent structure in the EIT measurements that purely data-driven models overlook.

On the simulated dataset, all models achieve higher performance metrics for both contact location and force compared to the real dataset. This reflects the more structured and noise-free nature of synthetic data, even with the incorporated perturbation reflecting the real data structure. Among the baselines, Autoencoder-MLP performs best for force prediction, while SVR and CNN-Transformer models show the lowest accuracy. Importantly, the physics-informed MLP consistently outperforms all baselines, achieving the highest R_f^2 (0.92) and the lowest MAE (0.41). For simulated (x, y)



(a) CNN-Transformer (b) Physics-informed (MLP)



(c) Physics-informed MLP

Fig. 4: Qualitative force prediction on the test set (real dataset). In (a) and (b) each panel plots predicted vs. ground-truth force with the $y = x$ reference line. The physics-informed MLP exhibits tighter clustering around the diagonal and fewer large outliers compared to the worst-performing baseline. (c) Spatial distribution of force prediction errors for the physics-informed MLP on the real dataset.

prediction, all models achieve $R^2 \approx 0.99$, but MAE reveals clear differences: physics-informed MLP attains the lowest error ($\text{MAE}_x = \text{MAE}_y \approx 1.3$), while Transformer and CNN–Transformer baselines perform worst with substantially higher errors (MAE_y up to 2.8). Taken together with the real dataset results, this shows that the proposed physics-informed approach provides consistent improvements across both domains.

Fig. 4 contrasts test-time force prediction quality between the worst baseline (e.g. CNN-Transformer) and our physics-informed MLP. The latter (Fig.4b) produces a visibly tighter cloud around the identity line with fewer extreme deviations, consistent with the quantitative gains reported in Table I. This qualitative analysis gives an interpretable insight into the range of force prediction performance over the weakest and strongest models tested.

To further illustrate the performance of our best-performing model, we visualize the spatial distribution of force prediction errors across the sensor surface (Fig. 4c). The errors are generally uniform, without pronounced hotspots, indicating that the physics-informed MLP generalizes well across different contact locations.

E. Sim-to-Real Similarity Analysis

A major challenge in using EIT tactile sensors is the *sim-to-real gap*, where simulated voltage measurements do not perfectly reflect the statistics of real-world signals. Since our proposed framework leverages both simulated and real datasets, it is important to quantify this discrepancy and assess whether perturbation-based augmentation reduces the gap. To this end, we compare the baseline (PlainSim) and perturbed simulated data (PertSim) using E.(9) against the real no-touch voltage signals.

TABLE II: Sim-to-real similarity scores between plain and perturbed simulated data compared to real no-touch voltage signals.

Comparison	MMD↓	CORAL↓	NCC↑	SWD↓
PlainSim vs Real	1.39	163.7	0.713	77.5
PertSim vs Real	0.74	151.9	0.733	76.7

We adopt four complementary similarity metrics, chosen for their prevalence in domain adaptation and distribution alignment tasks. Maximum Mean Discrepancy (MMD) captures differences in distribution means via kernel embeddings. CORAL measures covariance alignment between feature spaces. Normalized Cross-Correlation (NCC) quantifies global signal similarity, and the Sliced Wasserstein Distance (SWD) evaluates distribution geometry in a robust manner. Lower values in MMD, CORAL, and SWD, and higher values in NCC, indicate better similarity.

As shown in Table II, introducing perturbations leads to a notable reduction in MMD (from 1.39 to 0.74), indicating a closer match in global distribution between simulated and real voltages. CORAL and NCC also show mild improvements, suggesting better alignment in covariance and overall signal similarity. SWD values remain similar across conditions, reflecting that both simulated datasets preserve the general distribution geometry. Overall, this analysis supports that perturbation improves realism of simulated EIT signals, thereby reducing the sim-to-real gap and enhancing their utility for training learning-based models.

IV. Conclusion and Future Work

In this paper, we addressed the challenging problem of tactile state estimation using Electrical Impedance Tomography (EIT), a modality characterized by high-dimensional, noisy boundary measurements and nonlinear voltage-to-state mappings. We introduced a novel physics-informed learning pipeline that integrates knowledge of the EIT forward model into the training of data-driven estimators. Our approach consistently improved prediction accuracy, particularly for the more challenging task of force estimation, across both real and simulated datasets. We further benchmarked a wide range of neural architectures, including CNNs, Transformers, and autoencoder-based models, and showed that our physics-informed variant of MLP achieves the best performance. To account for the sim-to-real gap, we generated a perturbed simulation dataset and quantified similarity against real measurements using multiple distributional metrics, demonstrating improved realism compared to plain simulations. Finally, we release both our real and simulated datasets to facilitate reproducibility and benchmarking in future studies.

Despite these advances, some limitations remain. First, the sim-to-real generalization of models is still constrained by residual discrepancies between real and simulated voltages, particularly under force variations. Second, our current experiments were restricted to single-point contact and limited probe geometries, whereas richer multi-contact interactions remain unexplored. Finally, while we investigated representative neural architectures, incorporating recent advances in physics-guided representation learning or domain adaptation could further enhance performance.

In future work, we aim to expand this study by: (i) exploring more advanced domain adaptation techniques to close the sim-to-real gap; (ii) extending to multi-contact and dynamic tactile interactions; and (iii) integrating our pipeline into closed-loop robotic manipulation tasks. We believe these directions will accelerate the deployment of EIT-based tactile skins in real-world robotic manipulation.

REFERENCES

[1] S. Luo, N. F. Lepora, W. Yuan, K. Althoefer, G. Cheng, and R. Dahiya, "Tactile robotics: An outlook," 2025. [Online]. Available: https://arxiv.org/abs/2508.11261

[2] K. Nazari, W. Mandil, M. Santello, S. Park, and A. Ghalamzan-E, "Bioinspired trajectory modulation for effective slip control in robot manipulation," *Nature Machine Intelligence*, vol. 7, no. 7, pp. 1119–1128, 2025.

[3] K. Park, H. Yuk, M. Yang, J. Cho, H. Lee, and J. Kim, "A biomimetic elastomeric robot skin using electrical impedance and acoustic tomography for tactile sensing," *Science Robotics*, vol. 7, no. 67, p. eabm7187, 2022.

[4] H. Chen, X. Yang, P. Wang, J. Geng, G. Ma, and X. Wang, "A large-area flexible tactile sensor for multi-touch and force detection using electrical impedance tomography," *IEEE Sensors Journal*, vol. 22, no. 7, pp. 7119–7129, 2022.

[5] W. Xin, F. Zhu, P. Wang, Z. Xie, Z. Tang, and C. Laschi, "Electrical impedance tomographic shape sensing for soft robots," *IEEE Robotics and Automation Letters*, vol. 8, no. 3, pp. 1555–1562, 2023.

[6] D. Hardman, T. G. Thuruthel, and F. Iida, "Multimodal information structuring with single-layer soft skins and high-density electrical impedance tomography," *Science Robotics*, vol. 10, no. 103, p. eadq2303, 2025.

[7] B. Brown, "Electrical impedance tomography (EIT): A review," *Journal of Medical Engineering & Technology*, vol. 27, no. 3, pp. 97–108, 2003.

[8] E. Smela, *Electrical Impedance Tomography for Tactile Imaging*, ser. 2053-2563. IOP Publishing, 2024.

[9] Y. Zhang, G. Laput, and C. Harrison, "Electrick: Low-cost touch sensing using electric field tomography," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, ser. CHI '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 1–14.

[10] H. Park, W. Kim, S. Jeon, Y. Na, and J. Kim, "Graph-Structured Super-Resolution for Geometry-Generalized Tomographic Tactile Sensing: Application to Humanoid Faces," *IEEE Transactions on Robotics*, vol. 41, pp. 558–572, 2025.

[11] D. S. Tawil, D. Rye, and M. Velonaki, "Improved image reconstruction for an eit-based sensitive skin with multiple internal electrodes," *IEEE Transactions on Robotics*, vol. 27, no. 3, pp. 425–435, 2011.

[12] A. Costa Cornellà, D. Hardman, L. Costi, J. Brancart, G. Van Assche, and F. Iida, "Variable sensitivity multimaterial robotic e-skin combining electronic and ionic conductivity using electrical impedance tomography," *Scientific Reports*, vol. 13, no. 1, p. 20004, 2023.

[13] 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, 2015.

[14] M. Cheney, D. Isaacson, J. C. Newell, S. Simske, and J. Goble, "NOSER: An algorithm for solving the inverse conductivity problem," *International Journal of Imaging Systems and Technology*, vol. 2, no. 2, pp. 66–75, 1990.

[15] A. Adler and R. Guardo, "Electrical impedance tomography: Regularized imaging and contrast detection," *IEEE Transactions on Medical Imaging*, vol. 15, no. 2, pp. 170–179, 1996.

[16] 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.

[17] H. Chen, X. Yang, G. Ma, Y. Wang, and X. Wang, "Enhancing Tactile Sensing in Robotics: Dual-Modal Force and Shape Perception with EIT-based Sensors and MM-CNN," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. Yokohama, Japan: IEEE, 2024, pp. 3311–3317.

[18] K. Kim, J.-H. Hong, K. Bae, K. Lee, D. J. Lee, J. Park, H. Zhang, M. Sang, J. E. Ju, Y. U. Cho *et al.*, "Extremely durable electrical impedance tomography–based soft and ultrathin wearable e-skin for three-dimensional tactile interfaces," *Science Advances*, vol. 10, no. 38, p. eadr1099, 2024.

[19] Y. Huang and T. G. Thuruthel, "Electrical Impedance Tomography Based Finger-Shaped Soft Artificial Skin," *IEEE Robotics and Automation Letters*, pp. 1–8, 2025.

[20] Y. Chen, M. Yu, H. A. Bruck, and E. Smela, "Compliant multi-layer tactile sensing for enhanced identification of human touch," *Smart Materials and Structures*, vol. 27, no. 12, p. 125009, nov 2018. [Online]. Available: https://dx.doi.org/10.1088/1361-665X/aaeae4

[21] G. Ma, H. Chen, P. Wang, and X. Wang, "Multi-Frame Constrained Block Sparse Bayesian Learning for Flexible Tactile Sensing Using Electrical Impedance Tomography," *IEEE Transactions on Computational Imaging*, vol. 8, pp. 438–448, 2022.

[22] H. Dong, R. Bingnan Liu, L. Micklem, S. Peisan E, F. Giorgio-Serchi, and Y. Yang, "Data-Efficient Tactile Sensing With Electrical Impedance Tomography," *IEEE Sensors Journal*, vol. 25, no. 11, pp. 19 724–19 733, Jun. 2025.

[23] A. Elsanadedy, Y. Mamatjan, M. Ahmadi, and A. Adler, "Characterisation of Conductive Polymer for EIT based Sensor," in *Proceedings of the International Conference on Electrical and Computer Systems*, Ottawa, Ontario, Canada, Aug. 2012, p. 117.

[24] K. Nazari, W. Mandill, M. Hanheide, and A. G. Esfahani, "Tactile dynamic behaviour prediction based on robot action," in *Towards Autonomous Robotic Systems: 22nd Annual Conference, TAROS 2021, Lincoln, UK, September 8–10, 2021, Proceedings 22*. Springer, 2021, pp. 284–293.

[25] Y. Huang, D. Hardman, C. Bascucci, F. Clemens, and T. G. Thuruthel, "Design and modelling of electrical impedance tomography-based 3d-printed patterned soft tactile skins," in *2025 IEEE 8th International Conference on Soft Robotics (RoboSoft)*, 2025, pp. 1–6.

[26] W. Mandil, K. Nazari *et al.*, "Action conditioned tactile prediction: case study on slip prediction," *arXiv preprint arXiv:2205.09430*, 2022.

[27] N. Elijah, K. Nazari, M. Esfandiari, and E. A. Ghalamzan, "Advancing slip classification in robotic manipulation through tactile data representation and model selection," *IEEE Sensors Journal*, 2025.

[28] J. Huang, Y. Guo, Y. Jiang, F. Wang, L. Pan, and Y. Shi, "Recent advances and future prospects in tactile sensors for normal and shear force detection, decoupling, and applications," *Journal of Semiconductors*, vol. 45, no. 12, p. 121601, 2024.

[29] A. Ghalamzan, K. Nazari, H. Hashempour, and F. Zhong, "Deep-lfd: deep robot learning from demonstrations," *Software Impacts*, vol. 9, p. 100087, 2021.

[30] A. Adler and W. R. B. Lionheart, "Uses and abuses of EIDORS: An extensible software base for EIT," *Physiological Measurement*, vol. 27, no. 5, p. S25, Apr. 2006.

[31] B. Liu, B. Yang, C. Xu, J. Xia, M. Dai, Z. Ji, F. You, X. Dong, X. Shi, and F. Fu, "pyeit: A python based framework for electrical impedance tomography," *SoftwareX*, vol. 7, pp. 304–308, 2018.

[32] H. Dong, R. B. Liu, S. Teng, D. Hu, Peisan, E, F. Giorgio-Serchi, and Y. Yang, "Efficient tactile perception with soft electrical impedance tomography and pre-trained transformer," 2025. [Online]. Available: https://arxiv.org/abs/2506.02824

[33] D. Hardman, T. George Thuruthel, and F. Iida, "Self-healing ionic gelatin/glycerol hydrogels for strain sensing applications," *NPG Asia Materials*, vol. 14, no. 1, 12 2022.