

A framework for glass-box physics rule learner and its application to nano-scale phenomena

In Ho Cho^{1✉}, Qiang Li², Rana Biswas^{2,3,4} & Jaeyoun Kim²

Attempts to use machine learning to discover hidden physical rules are in their infancy, and such attempts confront more challenges when experiments involve multifaceted measurements over three-dimensional objects. Here we propose a framework that can infuse scientists' basic knowledge into a glass-box rule learner to extract hidden physical rules behind complex physics phenomena. A "convolved information index" is proposed to handle physical measurements over three-dimensional nano-scale specimens, and the multi-layered convolutions are "externalized" over multiple depths at the information level, not in the opaque networks. A transparent, flexible link function is proposed as a mathematical expression generator, thereby pursuing "glass-box" prediction. Consistent evolution is realized by integrating a Bayesian update and evolutionary algorithms. The framework is applied to nano-scale contact electrification phenomena, and results show promising performances in unraveling transparent expressions of a hidden physical rule. The proposed approach will catalyze a synergistic machine learning-physics partnership.

¹CCEE Department, 326 Town, Iowa State University, Ames, IA 50011, USA. ²ECpE Department, 1212 Coover Hall, Iowa State University, Ames, IA 50011, USA. ³Ames Laboratory, 311, Ames, IA 50011, USA. ⁴Microelectronics Research Center, 311, Ames, IA 50011, USA. ✉email: icho@iastate.edu

We have seen many triumphs of deep learning architectures successfully working on two-dimensional (2D) data sets, e.g.^{1–3}. But, the extension of deep learning to three-dimensional (3D) data sets is an active research area facing many theoretical and computational challenges^{4,5}. Although advanced machine learning (ML) methods gradually master and replace the complex actions that require the human-level intelligence and control capabilities^{6,7}, there exists a deep chasm between advanced ML methods and mature areas of science and engineering. Often, successful ML methods are limited to opaque decision-making capability about simple tasks. A direct adoption of an advanced ML method can hardly guarantee successful learning and prediction of the real-world experimental data that often involve 3D objects and multifaceted physical measurements. Scientific discovery often follows a typical process — measuring complex physical quantities, investigating the observed data, and deriving a rule that best describes the target physical phenomenon. This fundamental process poses several obstacles to the advanced ML methods. Physical measurements are often scattered over complex 3D spaces or objects. A few apparent descriptors are insufficient for learning. The size and volume of physical experimental data are relatively small for direct adoption of advanced ML methods. Scientists seek to find “expressions” of physical rules, not merely a “black-box” prediction.

Recently, a new research paradigm emerges to address these hurdles, i.e., the so-called physics- or theory-guided ML paradigm⁸. Domain science is used to help ML to predict physically sound solutions, e.g. human neurology⁹, quantum mechanics¹⁰, and heterogeneous composite structures¹¹. Governing physical rules (often partial differential equations) are often fed into ML, e.g., geophysics by an hierarchical graph model¹² and wave and fluid flow by deep learning¹³. The present goal of searching hidden rules shares the similar notion with auto-encoder methods in pursuit of salient latent terms¹⁴. Still, this study differs from prior works in several aspects. First, this framework focuses on obtaining transparent expressions of hidden physical rules. Second, it seeks to leverage basic physics and scientists’ experience rather than relying upon the predefined global governing equations. Third, it is built upon two ingredients, the externalized multi-layer convolved information index and flexible link functions which reinterpret and inherit the deep learning’s successful philosophies. Last, the identified expressions can easily evolve by embracing other physics and more experience. In essence, this framework is purely data-driven, requiring no distributional assumptions about priors and posteriors compared to¹⁴. A comparable recent work to our framework would be¹⁵ which pursues both hidden coordinates and a parsimonious form of governing equation of dynamic systems although the key methods and procedures are different.

This research proposes a framework for a constantly evolving, “glass-box” (as opposed to the black-box) rule learner which can help extract hidden physical rules behind complex real-world phenomena by integrating experienced scientists’ knowledge and the central notions of deep learning. In particular, we propose a framework that can facily deal with multifaceted measurements of 3D specimens by incorporating basic physics and scientists’ experience. Also, the framework can extract hidden rules in terms of transparent expressions which will evolve with increasing data through a Bayesian update scheme.

For feasibility tests, this research applied the framework to nano-scale contact electrification (CE), a process which is difficult to elucidate due to its complex geometry, intractable small length scale, and interaction of several physical mechanisms such as friction, fracture, demolding, and friction-induced charging (called tribocharging). The feasibility tests show a promising

performance of the proposed framework in reproducing complex 3D distributions of CE-induced electric potentials over the 3D point cloud along with transparent expressions.

Results

Quantified basic physics and experiences of scientists. The first question is about how to quantify and transform basic physics and scientists’ experiences into ML-friendly quantities. Depending upon the underlying physics and domain-specific knowledge, the resultant quantities may differ in forms and ranges.

This study presents an example of how to quantify and infuse the scientists’ simple experience. The target experiments adopted for the present feasibility tests is about CE during the demolding process of the hardened polydimethylsiloxane (PDMS) specimen from the polyethylene terephthalate (PET) base mold which holds unique nano-scale patterns. The physical measurements are distributed over the 3D point cloud and each point is on a smooth surface (Fig. 1a, b). The demolding occurs in a specific direction, e.g., left-to-right (Fig. 1c, d), and the separation of PDMS from the patterned PET mold (Fig. 1e) has an important role in determining the direction of crack propagation and frictional stress development (Fig. 1f, g), which scientists know as prior knowledge. We quantify this basic experience by a directional vector, $\mathbf{u} = \{u_x, u_y, u_z\} \in \mathbb{R}^3$. It is true that well-known analyses such as principal component analysis or Isomap¹⁶ with raw data may identify some information about the directivity of demolding test or reduced features, but this study regards scientists’ basic experiences as a starting point and focuses on how to infuse them into the subsequent rule learning. In particular, $\mathbf{u} = \{0.001, 0, 1\}$ was used to mimic left-to-right demolding process along the global X-axis whereas $\mathbf{u} = \{0.001, -0.0014, 1\}$, for the tilted demolding process. We denote the unit normal vector to the friction surface as $\mathbf{n} = \{n_x, n_y, n_z\} \in \mathbb{R}^3$, and each point has unique normal vector (Fig. 1b). Assuming scientist’s decent control of demolding, \mathbf{u} is constant for all points of an individual specimen while \mathbf{n} is varying at every point due to nano-scale patterns (Fig. 1e). Another important physical principle stemming from the scientists’ prior knowledge is that the friction-induced electric potential is proportional to the surface friction. Physically, the more surface friction, the more friction-related fracture energy can develop. Thus, to systematically quantify and infuse the degree of how much the demolding direction is aligned with the tangential direction of the surface friction, we leverage simple terms of the projection of demolding direction onto the friction surface, which can be mathematically described by

$$\mathbf{P}_u = \mathbf{u} - (\mathbf{u} \cdot \mathbf{n})\mathbf{n}, \quad (1)$$

where \mathbf{u} is the global demolding direction vector and \mathbf{n} is the surface normal vector. Since the current experimental data set consists of the observed electric potential over the 3D point cloud, the creation of the point-wise \mathbf{n} requires a specific algorithm, which is described in Supplementary Note 1. Aiming at a simple range term that accommodates all these experiences and physical knowledge, we define a point-wise information index (II). We denote the local II (i.e., the degree of projection of the demolding on the friction surface) as $II_{\text{local}} \in \mathbb{R}[-1, 1]$:

$$II_{\text{local}} = \text{sgn}(\mathbf{u}, \mathbf{n}) |\mathbf{P}_u|^2 / |\mathbf{u}|^2, \quad (2)$$

where $\text{sgn}(\mathbf{u}, \mathbf{n}) = +1$ if $\cos^{-1}(\frac{\mathbf{u}_{XY} \cdot \mathbf{n}_{XY}}{|\mathbf{u}_{XY}| |\mathbf{n}_{XY}|}) < \pi/2$; -1 otherwise; $\mathbf{u}_{XY} = \{u_x, u_y\}$, $\mathbf{n}_{XY} = \{n_x, n_y\}$. One simple physical reason behind this $\text{sgn}(\cdot)$ definition is that the specimens are assumed to lie on the XY plane and demolding occurs mainly toward the Z direction. Thus, $\text{sgn}(\mathbf{u}, \mathbf{n})$ gives $+1$ when the directions of demolding and surface normal are positively aligned, and -1 otherwise. It should be noted that the range of $II_{\text{local}} \in [-1, 1]$ has several physical

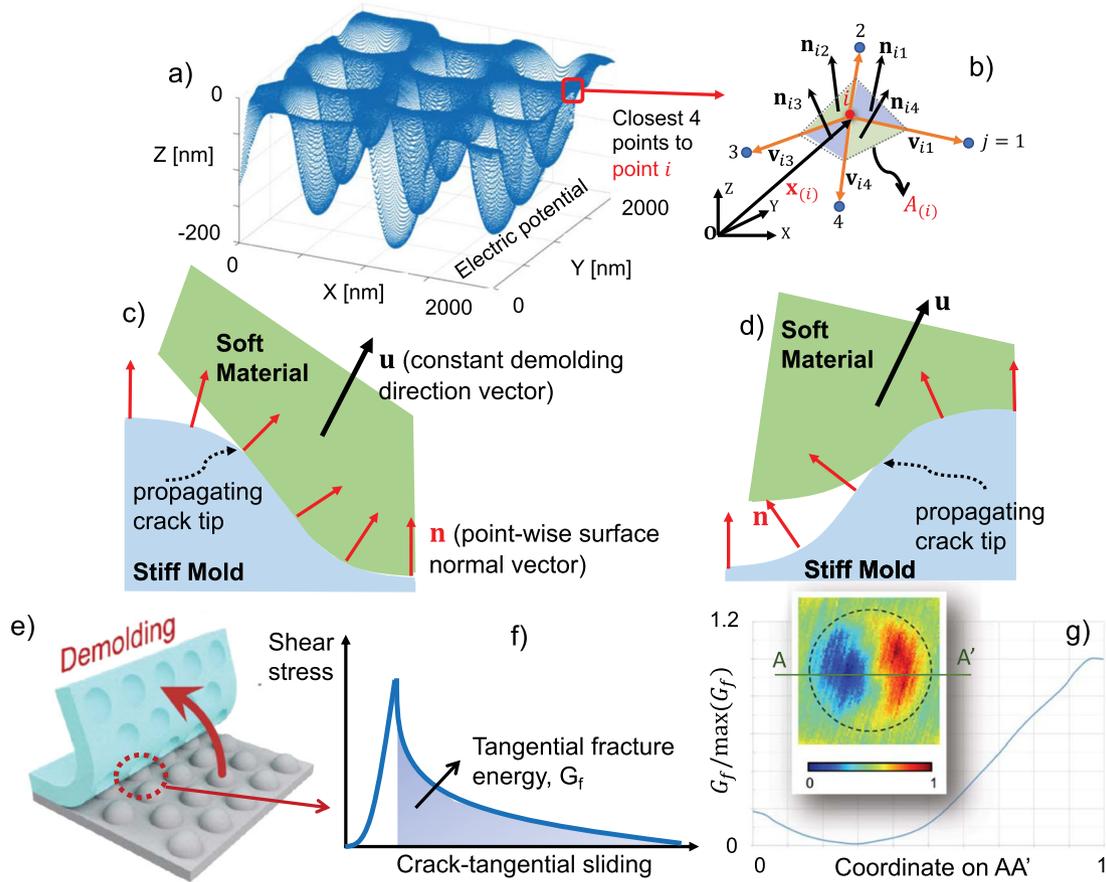


Fig. 1 Basic physics and scientist's experience to be quantified. **a** Example plot of the 3D point cloud of a nanocup array. **b** Point-wise normal vector \mathbf{n} and area A of a friction surface (see algorithms in Supplementary Note 1). **c-d** Illustration of the alignment of the demolding direction \mathbf{u} and normal vector \mathbf{n} ; **e** Demolding of nano-patterns in the lab. **f-g** Definition of tangential fracture energy G_f related to the tribocharging, and its calculated values over the circumference (AA') of a nanocup. Inset shows the normalized potential measured by Kelvin probe force microscopy (adapted from refs. 18, 19).

meanings. It will help confine the convolved \bar{II} in the simple range $[0, 1]$, which is compatible with the support range of the adopted cubic spline-based link functions. The simple monotonic range from -1 to 1 retains the similarity to activation values of typical ML methods. Clear physical interpretations exist: $+1$ indicating the positively aligned tangential friction whereas -1 meaning the negatively aligned condition. As long as these characteristics are satisfied, there exists ample room for different expressions of information index depending upon domain-specific knowledge. In computational mechanics, for instance, the spatial proximity between disparate materials can be represented by another form of information index by using virtual stress excitation¹¹. It should be noted, however, that the local index by itself may be insufficient to capture the interactions among nearby physical quantities and may suffer from irregularity and local spikes. These drawbacks are to be resolved by the convolved information index, similar to multi-layered convolution in deep learning. Example plots of the local information index calculated with practical nano-scale measurements are presented in Supplementary Note 2.

Generation of convolved information index. One of the key enablers of deep learning is the convolution process that allows information integration. If convolution is done over a spatial domain, ML can better understand the interaction of spatially distributed information and hidden patterns while applied to the temporal domain, the interactions between past and present information can be elucidated^{5,17}. Inheriting the philosophy of

the deep learning's convolution, the proposed framework seeks to spatially integrate the local II over the 3D point cloud. The key difference of this study's convolution is that we "externalize" the deep learning's multi-layered convolutions by conducting multiple convolutions at the input information level, not in the opaque deep network layers. Rather than a uniform integration, we adopt a weighted integration using Gaussian weight function (denoted ω) to realize the proximity-proportionate importance of information. This process creates "convolved" information index denoted as \bar{II} , which is calculated by

$$\bar{II}(\mathbf{x}_{(i)}) = \frac{1}{2} \left(1 + \int_V \omega(\mathbf{x}_{(i)}, \boldsymbol{\xi}) II_{\text{local}}(\boldsymbol{\xi}) d\boldsymbol{\xi} \right), \quad (3)$$

where the factor $1/2$ and addition of 1 in the parentheses are for normalization to $[0, 1]$. For the integration over a discrete 3D point cloud, with the uniformity assumption over a small patch we have an approximation of

$$\bar{II}(\mathbf{x}_{(i)}) \cong \frac{1}{2} \left(1 + \sum_{\boldsymbol{\xi}_{(j)} \in V} \omega(\mathbf{x}_{(i)}, \boldsymbol{\xi}_{(j)}) II_{\text{local}}(\boldsymbol{\xi}_{(j)}) A_{(i)} \right), \quad (4)$$

where $A_{(i)}$ is a patch area of the friction surface centered at the point i and calculated over the 3D point cloud (see the algorithm in Supplementary Note 1); $\boldsymbol{\xi}_{(j)}$ is the position vector of the point j . As a primary spatial weighting function, we used the Gaussian function which has been widely used in other fields with different

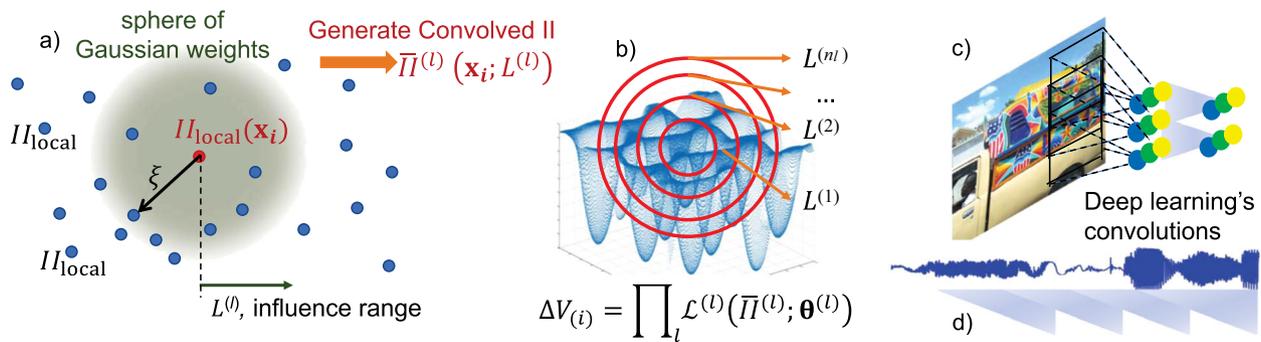


Fig. 2 Comparison of the proposed framework's convolved information index over the 3D point cloud and deep learning's typical convolutions. **a** Each point has its own I_{local} , and the weighted integration of Eq. (3), a spatial convolution, generates the “convolved information index” (denoted as \bar{I}) at all points. **b** Multiplicative interaction among several link functions (LFs) of multiple influence ranges similar to the deep learning's multiple convolutions over several layers. **c** Deep learning's typical 2D convolution of image and **d** a temporal convolution for continuous aggregation of past information (adapted from ref. 17).

names (other weighting functions may be used):

$$\omega(\mathbf{x}_{(i)}, \xi) = \left(L(2\pi)^{1/2}\right)^{-N} \exp\left(-\frac{|\mathbf{x}_{(i)} - \xi|^2}{2L^2}\right) = \mathcal{N}(\mathbf{x}_{(i)}, L^2), \quad (5)$$

where position vectors $\mathbf{x}_{(i)}$, $\xi \in V$; L is the influence range parameter; N is the dimension parameter (herein $N = 3$ for the 3D point space). Of particular importance is the radius of influence range L . With a larger value of L , the information over a large space can be incorporated, but over-smoothing effects may occur; with a smaller L , adjacent information to the current point becomes more important, but the near information may be peaked which may lead to an over-fitting issue (visual comparisons of varying L are presented in Supplementary Note 2). When compared to the convolution of image data, the larger L may correspond to a large image patch used for the convolution. Figure 2 illustrates how the spatial information integration is conducted over the 3D point cloud, and also compares its similarity to the spatial or temporal convolutions used in typical deep learning methods. In the multi-layered deep learning, convolution can take place multiple times over several layers. Analogous to this philosophy, we generate multiple convolved IIs with various values of L and consider their interactions via multiple link functions.

Feasibility tests with nano-scale experimental data. To confirm the feasibility of the proposed framework, we applied it to real experimental data sets of nano-scale electric potentials. The charge values were measured by Kelvin probe force microscopy over patterned nano-cups represented by the 3D point cloud (Figs. 3–5; raw test data are adapted from^{18,19}). Raw data in Fig. 3a, b are from $4000 \text{ nm} \times 4000 \text{ nm}$ square area with interval of 7.8125 nm , and Fig. 3c, d show the reproduced potentials using our framework. Raw data in Fig. 4a, b are measured over $3000 \text{ nm} \times 3000 \text{ nm}$ square area with interval of 5.86 nm , and Fig. 4c, d show the corresponding predictions. Figure 5a, b has the same setup as Fig. 3, and Fig. 5c, d are the prediction plots. The demolding direction is different in all experiments including rightward, diagonal, and downward as marked in figure. Another complexity arises from the diverse heights of nano-cups: 154 nm for Fig. 3a, 93.5 nm for Fig. 4a, 117 nm for Fig. 5a, and 50.2 nm for Fig. 5c, respectively. Nano-scale patterns are also different, i.e. nano-cup arrays and parallel ridges. Since the goal of this study is to learn underlying physical rules, the proposed framework should

overcome all the aforementioned apparent diversities and complexities.

To confirm general learning capability, the Bayesian-evolution training is conducted on two seemingly disparate experimental data sets. In particular, the proposed glass-box learning begins from raw data of Fig. 3a, and then the identified rules are inherited to the next learning of Fig. 4a. After each training finishes, the best-so-far rules are used to reproduce the distribution of electric potentials of training data sets as shown in Fig. 3c, d and Fig. 4c, d. Out of many possible combinations of multiple link functions (LFs) with different influence ranges, the best-so-far rule is found to be the combination using 3 LFs of $L = 8, 50, 100 \text{ nm}$ (here, L is the spatial convolution influence range). The mathematical expression is attained by the proposed framework as

$$\Delta V_{(i)} = \prod_{l=1}^3 \mathcal{L}^{(l)}(\bar{I}_{(i)}^{(l)}; \theta^{(l)}), \quad (6)$$

where $\theta^{(l)} = \{\mathbf{a}^{(l)}, \mathbf{x}^{*(l)}\}$ are the free parameters of the l th LF $\mathcal{L}^{(l)}$ and their values are summarized in Supplementary Table 1. Then, the identified physical rules are used to predict substantially different experimental data sets, Fig. 5a, b. Visual illustrations of the identified rules are presented in Supplementary Note 3. The identified rule of Eq. (6) appears to be able to predict the spatial distributions of potentials which are substantially irregular and complex. It should be noted that the electric potential is a relative quantity, and thus we focus on overall shapes of the potential rather than specific values of certain locations. The best-so-far rule appears to reasonably reproduce the peaks and patterns over the 3D space. As anticipated, the identified rule is defined at the material point level, and thus the prediction can be done regardless of substantial differences in demolding directions, nano-patterns, and different geometries. Since the inheritance takes place for underlying rule expressions, the learned physical rules can be used for general cases (Fig. 5) regardless of substantial differences of nano-scale experiments.

Discussion

One of the strengths of the proposed glass-box learner is its transparency and clear interpretability. In view of Eq. (3), each convolved information index may be regarded as a marginal likelihood with the Gaussian conditional probability (dimension

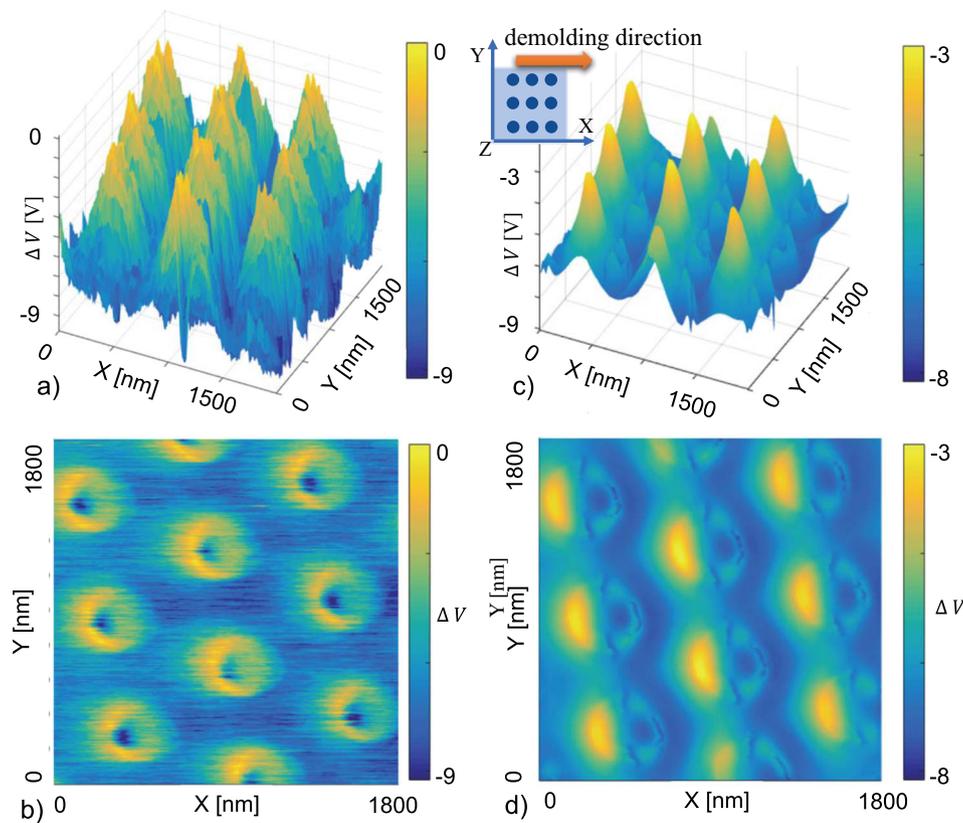


Fig. 3 Initial training with raw data and the first prediction of electric potentials. **a–b** Bird's eye and top view of the real experimental data of the 3D point cloud showing a specimen with the rightward demolding direction parallel to the X-axis as shown in inset (experimental data values are adapted from Fig. 2a, d of ref. ¹⁹). **c–d** Reproduced results using the identified physical rule with the best-so-far 3 link functions of convolved IIs in Eq. (6).

$N = 3$; influence range L):

$$\mathbb{E}_{\mathcal{N}(\mathbf{x}_{(i)}, L^2)}(II_{\text{local}}(\xi)) := \frac{1}{2} \left(1 + \int_V (L(2\pi)^{1/2})^{-N} \exp\left(-\frac{|\mathbf{x}_{(i)} - \xi|^2}{2L^2}\right) II_{\text{local}}(\xi) d\xi \right) \quad (7)$$

Here, the factor 1/2 and addition of 1 used for normalization (Eq. 3) do not change the intended meaning of the likelihood. The best-so-far LFs are regarded as a weighted summation of constant, linear, quadratic, and cubic polynomials, respectively). If we decompose the best-so-far LF's cubic spline basis, it will help elucidate the probable relationships; each polynomial basis informs the dominant relationship (e.g., linear, parabolic or high nonlinearity) between the target physics and II (compare Supplementary Figures 5 and 6 in Supplementary Note 4). This clear interpretability may help explicitly reveal a dominant relationship and approximation in subsequent investigations of domain scientists.

From the statistical angle, all the bases take the marginal likelihood $\mathbb{E}_{\mathcal{N}(\mathbf{x}_{(i)}, L^2)}(II_{\text{local}}(\xi); \theta)$. In particular, the observed physical rule behind the electric potential and our information index is written in an explicit form as

$$\Delta V_{(i)} = \prod_{l=1}^3 \left[a_1^{(l)} + a_2^{(l)} \times \mathbb{E}_{\mathcal{N}(\mathbf{x}_{(i)}, L^{(l)2})}(II_{\text{local}}) + \sum_{j=3}^5 \left(a_j^{(l)} \times b_j^{(l)} \left(\mathbb{E}_{\mathcal{N}(\mathbf{x}_{(i)}, L^{(l)2})}(II_{\text{local}}) \right) \right) \right], \quad (8)$$

where $b_j^{(l)}(x)$ is basis of the adopted cubic spline for LFs

(details about this basis are presented in the section of Flexible and transparent link functions in Methods); target ΔV and local index II_{local} are defined at each material point, i.e. each data point of the 3D point cloud; $\theta^{(l)} = \{a^{(l)}, x^{*(l)}\}$ are summarized in Supplementary Table 1. By all means, this identified physical rule about ΔV is not a fixed, unique form. Rather, it suggests a best-so-far probable and physically explainable expression regarding the target phenomena. There are myriad of ways to use this identified physical rule based on scientists' knowledge. For instance²⁰, suggested a plausible causal pathway between the frictional charge and contact-surface temperature difference, $\Delta V \propto \Delta T$, at nano-scale surfaces. Since the local information index always preserves physical meaning, e.g., direction alignment of demolding and tangential friction in Eq. (2), researchers may derive a new physical governing equation by simply linking the ΔT to Eq. (8) via a new II. For instance, atomic-scale investigations^{21–23} into thermally activated Prandtl and Tomlinson (PT) model suggests an interesting relationship among friction F_1 , temperature T , and velocity v : $F_1(v, T) = F_c - (\beta k_B T \ln(v_c/v))^{2/3}$ where v_c is the critical velocity, k_B is the Boltzmann constant, and F_c means the maximum slip-inducing force at zero temperature. It would be an interesting endeavor to link this PT relation to the II (e.g., $F_1 = f(II_{\text{local}})$ with an invertible expression f) and finally to ΔV by plugging f^{-1} into Eq. (8). Such a new merge across multiple physics will be transparent by virtue of the glass-box learning and prediction of the proposed framework.

This study proposed a framework to seek a constantly evolving, glass-box physical rule learner by addressing how to infuse scientists' knowledge and experience into quantifiable, ML-friendly terms. The framework's two primary ingredients, the convolved II and LF, are inspired by deep learning's multi-layered convolutions. The framework's constantly evolving capability arises from

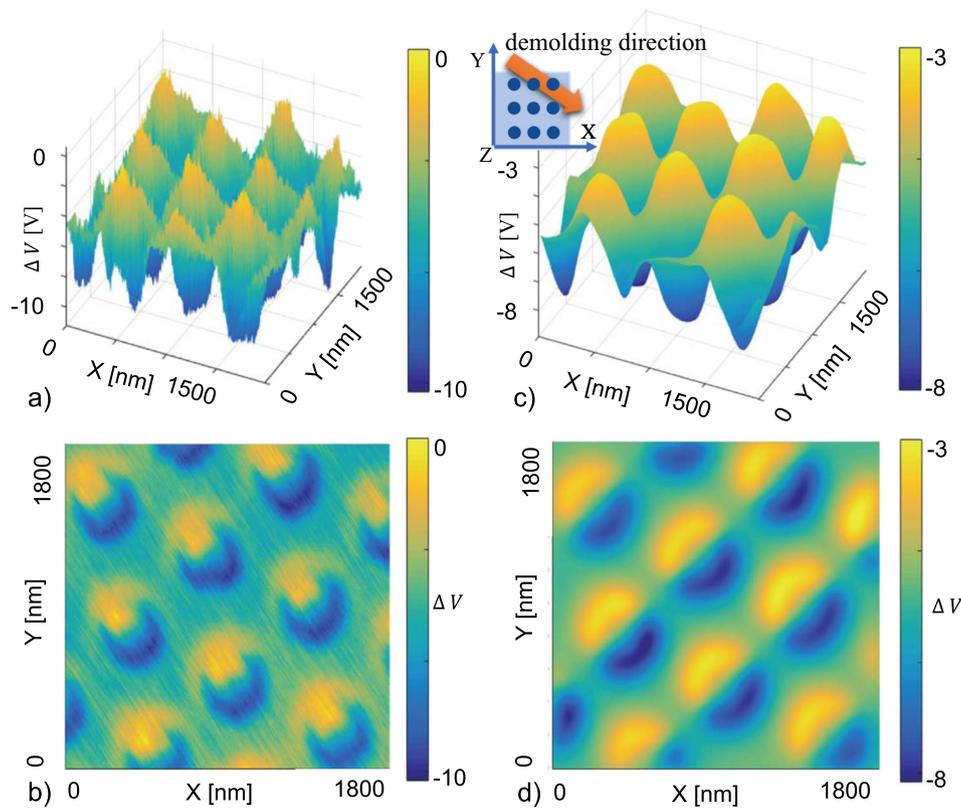


Fig. 4 Bayesian training with raw data and reproduction of electric potentials. **a–b** Bird-eye and top view of the real experimental data showing a specimen with the diagonal demolding direction as shown in inset (experimental data values are adapted from Fig. 2b, e of ref. ¹⁹). **c–d** Reproduced results using the identified physical rule with the best-so-far 3 link functions of convolved IIs in Eq. (6). A Bayesian update is used to inherit trained physical rule from Fig. 3.

the combination of a Bayesian update and evolutionary algorithm, which may translate into a maximization of log-likelihood, requiring no distributional assumptions about the priors and posteriors. The framework seeks to identify explicit mathematical expressions about the target physical phenomenon and IIs via LFs. Practical feasibility tests with complex nano-scale CE phenomena showed a promising capability of the framework in identifying reasonable expressions about intractable electric potential distribution across the 3D point cloud measured from nano-patterned specimens with varying geometry and demolding directions. By virtue of the transparency of LFs, the revealed physical rule will serve as a gateway for numerous possible rules, and with such a fertile partnership with ML, the new discovery will return to the hands of scientists in diverse disciplines. In general, this framework can be applied to other disciplines where multi-dimensional, multi-faceted physical data sets and the limited access to internal states pose challenges: e.g., new meta-material design, geophysics, and complex heterogeneous bodies¹¹. Overall, this study demonstrates how advanced ML methods inspire domain scientists and how they can combine to tackle hitherto intractable scientific questions, promoting more cross-disciplinary collaborations. This framework will spark imagination of scientists to develop their domain-specific IIs and also invigorate ML community to embrace IIs into their successful platforms.

Still, there are ample rooms for further sophistication of this initial framework: to name a few, more flexible and versatile basis for link functions²⁴ or use of an extensive library of possible mathematical expressions as done in¹⁵; more advanced ML methods and evolutionary algorithms for efficient searching of latent feature and high-dimensional parameters^{14,25}; automated

approaches to finding optimal layers of LFs; integration of diverse multiphysics rules such as nano-scale heat transfer, nano electrification and mechanical friction. Toward any extensions and applications, ideas in this study will serve as a fertile ground for departure.

Methods

Flexible and transparent link functions. In deep learning methods, the hidden layers' weights embed important relations and interactions of variables and neurons in terms of numerical values. The meaning of weights gradually becomes opaque as the number of hidden layers of convolution processes increases. To emphasize higher transparency and interpretability, this study suggests using a LF that describes the impact of convolved information index \bar{I} on the hidden physical rules in terms of clear mathematical expressions. LF is denoted as $\mathcal{L}(\bar{I}; \theta)$ where θ is a set of free parameters prescribing the LF. Since the true form of a hidden physical rule remains unknown, this study suggests to borrow the power of an evolutionary algorithm to enable LF to continue to learn, train, and evolve. The framework focuses on evolving θ of LF, rather than finding a single set of parameters. Although there is no restriction on the selection of specific LF, this study chose the cubic spline basis, which is highly smooth and flexible. The cubic spline curves consist of a few cubic polynomials connected (at the so-called knots) such that the curves are continuous up to the second derivatives²⁴. For example, when a practical cubic spline basis²⁶ (denoted as b_i) is adopted, LFs are given as

$$\mathcal{L}(\bar{I}; \mathbf{a}, \mathbf{x}^*) = \sum_i^p a_i b_i(\bar{I}) \quad (9)$$

where $b_1(x) = 1$, $b_2(x) = x$, and

$$b_{i+2}(x) = \frac{\left[(x_i^* - \frac{1}{2})^2 - \frac{1}{12} \right] \left[(x - \frac{1}{2})^2 - \frac{1}{12} \right]}{4} - \frac{\left[(|x - x_i^*| - \frac{1}{2})^4 - \frac{1}{2} (|x - x_i^*| - \frac{1}{2})^2 + \frac{7}{240} \right]}{24}, \quad (10)$$

for $i = 1 \dots p - 2$. Here, x_i^* is i th knot location. Therefore, to completely describe one LF, we need to identify $p + (p - 2)$ unknowns, i.e. $\mathbf{a} = \{a_1, \dots, a_p\}$ and $\mathbf{x}^* = \{x_1^*, \dots, x_{(p-2)}^*\}$. For brevity, we denote the total unknown parameters as $\theta = \{\mathbf{a},$

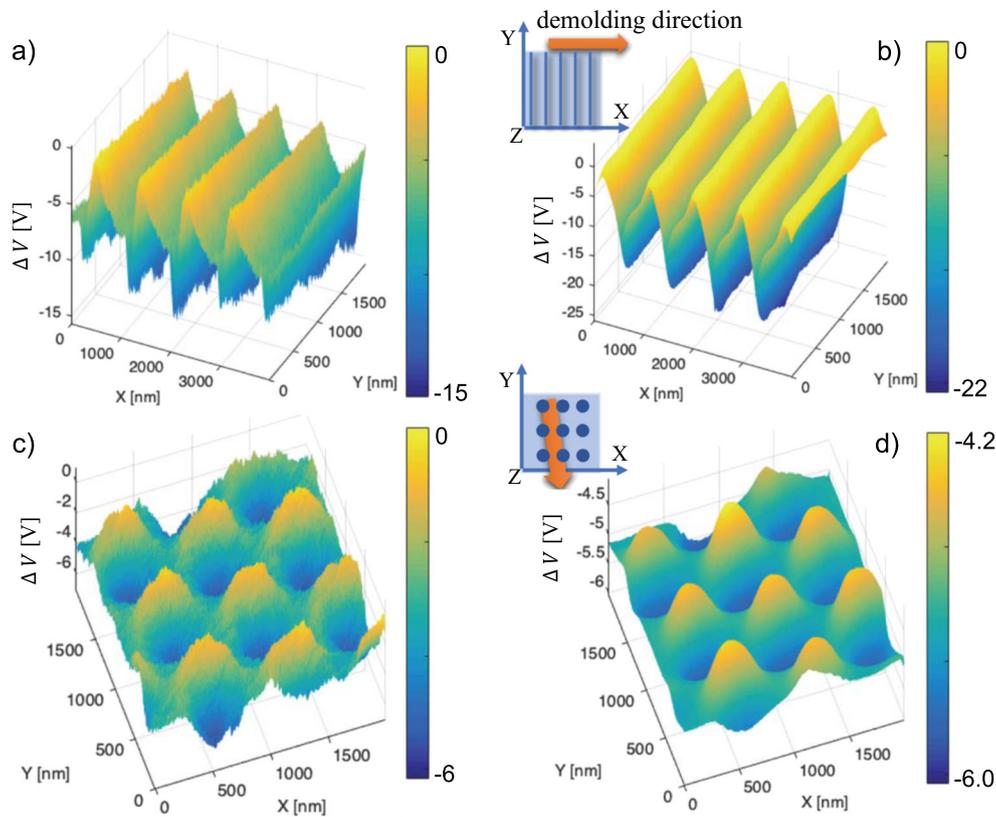


Fig. 5 Prediction tests with substantially different experimental data sets. **a–b** Rightward demolding direction parallel to the X -axis as shown in inset. **a** Experimental data (raw data values are adapted from Fig. S2a, c in ref.¹⁹) and **b** predicted electric potentials using the best-so-far physical rule with three link functions of convolved IIs in Eq. (8). **c–d** Downward demolding direction parallel to the Y -axis as shown in inset. **c** Experimental data (raw data values are adapted from Fig. 2c, f in ref.¹⁹) and **d** predicted electric potentials using the best-so-far physical rule with three link functions of convolved IIs in Eq. (6). The experimental data for these prediction tests were not used for the Bayesian training, and the identified physical rules are solely based on training on the raw data of Fig. 3 and Fig. 4.

\mathbf{x}^* . Regarding the LF's flexibility, when a physical relationship is monotonic, a smooth shape LF may suffice, but when the hidden relations are in complex relations, the more flexible shape may be better. The adopted cubic spline basis can accommodate all of these characteristics. It should be emphasized that the adopted cubic spline basis is not used for direct regression. Rather than aiming at direct regression, this framework seeks to leverage the flexibility and transparency of cubic spline basis for finding “expressions” of LFs. Comparable to the multiple convolution over layers of deep learning, we allow an interaction of multiple LFs with different influence ranges. Thus, a target physical response ΔV is, in general, obtained with $\theta^{(l)} = \{\mathbf{a}^{(l)}, \mathbf{x}^{*(l)}\}$ by

$$\Delta V_{(i)} = \prod_l \mathcal{L}^{(l)}(\bar{\Pi}_{(i)}^{(l)}; \theta^{(l)}), \quad (11)$$

where $\Delta V_{(i)}$ is the predicted electric potential at point (i) , $\bar{\Pi}_{(i)}^{(l)}$ is the convolved II associated with l th influence range $L^{(l)}$ at the point (i) , and the number of total LFs is denoted by n_l which is also to be determined through learning. After we considered various numbers of LFs and different combinations such as additive or multiplicative, we found that $n_l=3$ with influence ranges (e.g., $L^{(l)}=8, 50,$ and 100 nm) and multiplicative combinations as in Eq. (11) lead to reasonable learning. As in the multiple convolutions over several layers of deep learning, this framework compares possible combinations of multiple LFs allowing interaction among different ranges' information and finds the best performance case as illustrated in Fig. 2. Unlike hyperparameters of other ML methods, the proposed LFs seek to offer “expressions” which will be inserted or interwoven with other physical phenomena.

Bayesian update ingrained into an evolutionary algorithm. Aiming at no distributional assumptions about the priors/posteriors as well as pursuing smooth evolution, this study adopts the combination of a Bayesian update and a modified genetic algorithm¹¹. The key evolutionary algorithm involves the preparation of initial generation, organism-wise evaluation of fitness score, and fitness-based spawning of the next generation. The prior best physical rules can be smoothly inherited by the Bayesian update-based fitness proportionate probability (FPP) rule. To accelerate the evolution speed of the modified genetic algorithm, an

individual variable-wise gene cross-over scheme²⁷ has been used, and the changing search range scheme²⁸ is used in an iterative manner for better performance as successfully done in¹¹. Since an individual s realizes a candidate of $\theta = (\mathbf{a}, \mathbf{x}^*)$ in current generation S , the raw cost of an individual s , termed as $\mathcal{J}(s)$, is calculated by

$$\mathcal{J}(s) = \frac{1}{n} \sum_i^n [\Delta V_{(i)} - \Delta V_{(i)}(\theta)]^2, \quad (12)$$

where $\Delta V_{(i)}$ is the true (measured) physical response at point (i) and $\Delta V_{(i)}$ is the predicted response by Eq. (11) with θ . This fitness score simply means the mean squared errors of the observed and the predicted surface potentials. Then, following typical genetic algorithm procedures^{11,29,30}, the normalized fitness score \mathcal{F} of an individual is calculated by

$$\mathcal{F}(s) = \frac{(1 + \mathcal{J}(s))^{-1}}{\sum_{s \in S} [(1 + \mathcal{J}(s))^{-1}]}, \quad (13)$$

where s denotes an individual in the entire generation S . Learning a hidden physical rule is not a one-time task, rather a continuous activity. As diverse new experimental data become available, the physical rule learner must embrace all the previous knowledge and learn new information. To seamlessly realize this continuous learning, this study infused a Bayesian update scheme into the evolutionary algorithm's FPP rule. Suppose we have the best-so-far generation, denoted as S^* and its associated fitness scores, $\mathcal{F}^*(s)$, $s \in S^*$. According to the FPP rule, the probability of selecting an θ for next parent is given by $\text{Prob}(\theta) \propto \mathcal{F}(s)$, $s \in S^*$. Thus, $\mathcal{F}^*(s)$ is regarded as a prior probability density function of parameters $\theta = \{\mathbf{a}, \mathbf{x}^*\}$, i.e. $\pi_{\text{prior}}(\theta)$ in the typical Bayesian formalism. For initialization of $\pi_{\text{prior}}(\theta)$, there are several choices: fully random initialization, expert knowledge-based³¹ or the principle of maximum entropy-based initialization³². This study intentionally departs from fully random initialization to investigate positive evolution trends without special initialization assumption. Thus, this framework is purely data-driven, requiring no distributional assumptions about priors and posteriors compared to some successful auto-encoder methods¹⁴. For the posterior distribution, we proposed the following two-stage procedure.

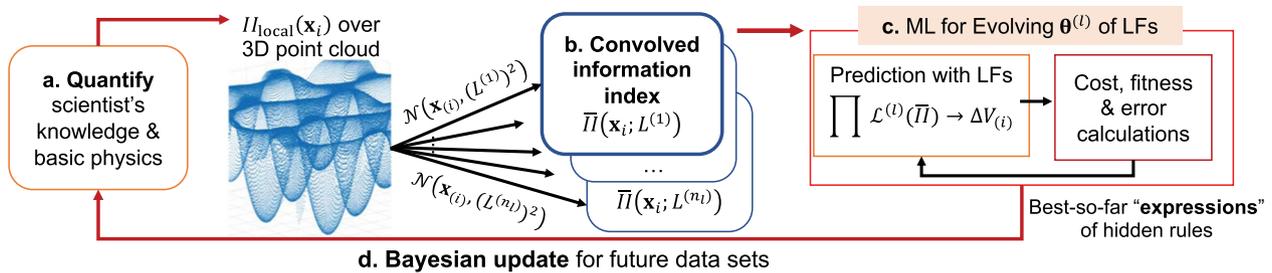


Fig. 6 Flowchart of the proposed framework. **a** Basic physics and experience is translated into a local I_{local} at all spatial point \mathbf{x}_i , **b** With various ranges $L^{(l)}$ and Gaussian weights \mathcal{N} , multiple convolved IIs (\bar{I}) are generated; **c** Using \bar{I} and prediction errors of target electric potentials (ΔV), evolutionary algorithm reveals LFs. **d** Bayesian update is incorporated for cross-specimen evolution.

Suppose that we have the prior best LFs and their S^* and that new experimental data become available. At the first learning generation with the new data, we can calculate the first fitness scores $\mathcal{F}(s; S^*)$ by applying the prior S^* and LFs to the new experiment. After the first generation, we can estimate the Bayesian fitness score (denoted as \mathcal{F}_B) as

$$\mathcal{F}_B(s) = \frac{1}{\kappa} \frac{\mathcal{F}(s; S^*) \mathcal{F}^*(s)}{\sum_{s \in S^*} \mathcal{F}^*(s)}, \quad (14)$$

where κ is needed for normalizing the Bayesian fitness to unity, which is simply given by

$$\kappa = \frac{\sum_{s \in S^*} \mathcal{F}(s; S^*) \mathcal{F}^*(s)}{\sum_{s \in S^*} \mathcal{F}^*(s)}. \quad (15)$$

Then, from the second learning generation of the new experiment, the probability of selecting two parents is proportional to the Bayesian fitness score as

$$\text{Prob}(\text{parent}_i | s) \propto \mathcal{F}_B(s), \quad (i = 1, 2). \quad (16)$$

Once again, an individual s realizes a candidate of $\theta = (\mathbf{a}, \mathbf{x}^*)$ in the new generation S , and thus the desired posterior distribution is obtained. In this way, the prior knowledge is smoothly inherited to the new experiment on the framework of evolutionary algorithm, thereby enabling constantly evolving physical rule learning. To allow for evolving with new data, the previous scores are inherited by the Bayesian score Eq. (14). Since the adopted evolutionary algorithm remembers prior generation's fitness scores, which offer the probability distribution of free parameters of LFs. As the Bayesian inheritance continues with new experimental data, the probability distribution of LFs will naturally evolve. Thus, the proposed framework can achieve evolving capability with increasing data. In the future, more dedicated investigations should focus on validation of the constantly evolving capability of LFs with sufficient, diverse test data. To some extent, the aforementioned combination of a Bayesian update and evolutionary algorithm can be viewed as a log-likelihood maximization as explained in Supplementary Note 4.

Overall flow of the evolving physical rule learner. Based on the aforementioned building blocks, the proposed framework has the overall architecture as shown in Fig. 6. For a target physical phenomenon, learning begins with quantifying the scientist's knowledge and experience into a simple, local information index (step **a** in Fig. 6). Spatial information integration is conducted with multiple influence ranges (i.e. $L^{(l)}$, $l = 1, \dots, n_l$) and Gaussian weights \mathcal{N} as marked between steps **a** and **b** in Fig. 6. Multiple convolved IIs and their interaction may be regarded as the counterparts to the deep learning's multi-layered convolutions. As mentioned earlier, the central novelty is that we "externalize" the deep learning's multi-layered convolutions by conducting multiple convolutions at the input information level (**b** in Fig. 6), being independent of learning engine. Then, a ML or optimization method (here, evolutionary algorithm) is used for learning and evolving internal parameters θ of the LFs, not for direct prediction of target responses (step **c** in Fig. 6). Thus, the ML methods' powerful strength of training and solution searching can focus on identifying mathematical expressions between the IIs and the target physical responses. The best-so-far "expressions" of the hidden rule will become a prior best generation for the next generation's LFs via the Bayesian update scheme (step **d** in Fig. 6) for future data sets.

Raw experimental data sets of electric potentials. The raw experimental data sets adapted for the feasibility tests are in a text-based matrix form of $\{x, y, z, \Delta V\}_{(i)}$, $i = 1, \dots, n$: (1) the horizontal demolding case (Fig. 3) has a 65536×4 matrix, (2) the diagonal demolding case (Fig. 4) has a 115600×4 matrix, and (3) the horizontal demolding case of parallel strip patterns (Fig. 5) has a 65534×4 matrix. Here, the unit of coordinates $\{x, y, z\}$ is [nm] while the electric potential charge ΔV is in [Voltage] or [V]. All of these raw training data sets are made publicly available (see Data Availability).

General settings for the evolutionary algorithm. The initial search ranges for all link functions' parameters $\mathbf{a}_i \in \mathbf{a}$ are set as $a_i \in [-3000, 3000]$ while $x_1^* \in [0, 1/3]$, $x_2^* \in [1/3, 2/3]$, and $x_3^* \in [2/3, 1]$ for the three knots of cubic spline regression basis. For typical settings for the genetic algorithm, 4 alleles are used per gene; the mutation rate is 0.01–0.02; the total organisms per generation is 10,000–20,000. The maximum generation number is 1000 which gave reasonably converged results.

Computational implementation of proposed algorithms. The spatial convolution of the 3D point cloud to generate the convolved information index may be computationally expensive depending upon the size of data points. This framework developed a scalable parallel program with C++ and *OpenMPI*. All other learning, evolutionary algorithm and a Bayesian update scheme are implemented on the parallel program. The developed program will be publicly shared for academic purposes upon request to the corresponding author. Iowa State University's high-performance computing facility, *Condo* cluster is used for this study.

Fabrication of tribocharged nano-cup array. The tribocharged PDMS nanocup array is fabricated by the sequence of (1) prepare a PET mold with a 750 nm-pitch triangular array of nanocones (Micro-continuum Inc.), (2) pour the liquid phase PDMS, Sylgard 184, Dow Corning, mixed with the curing agent, (3) solidify the specimen, and (4) peel off the PDMS from the PET mold (for more details see ref. 19).

Surface characterization with Kelvin probe force microscopy. Atomic force microscopy (AFM) (Multimode, Veeco) in the tapping mode is used to obtain the topography. AFM in the Kelvin probe force microscopy mode is used to measure the surface topography and potential. The settings for Pt/Ir coated tips (purchased from Bruker) are SCM-PIT-v2, the spring constant of 2.8 N.m^{-1} , and the resonance frequency of 75 kHz. The lift height and the typical scanning rate are 45 nm and 0.5 Hz, respectively (for more details see ref. 19).

Data availability

Raw training data of nano tribocharging experiments adapted from the authors' prior work¹⁹ are available at <https://sites.google.com/site/ichoddcse2017/home/Computational-Sci-Eng/nano-letters-2019>. All the computational programs are available upon proper request to the corresponding author.

Received: 9 December 2019; Accepted: 23 March 2020;

Published online: 08 May 2020

References

- Krizhevsky, A., Sutskever, I. & Hinton, G. E. Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, Edited by F. Pereira et al., Neural Information Processing Systems Foundation, Inc., 1097–1105 (2012).
- Long, J., Shelhamer, E. & Darrell, T. Fully convolutional networks Imagenet classification with deep convolutional networks for semantic segmentation. *Proc. International Conference on Computer Vision and Pattern Recognition, IEEE Computer Society, USA*, 3431–3440 (2015).
- He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. *Proc. International Conference on Computer Vision and Pattern Recognition, IEEE Computer Society, USA*, 770–778 (2016).
- Bronstein, M. M., Bruna, J., LeCun, Y., Szlam, A. & Vandergheynst, P. Geometric deep learning: going beyond euclidean data. *IEEE Signal Process. Mag.* **34**, 18–42 (2017).

5. Ahmed, E. et al. Deep learning advances on different 3D data representations: a survey. Preprint at <https://arxiv.org/abs/1808.01462v2> (2018).
6. Brown, N. & Sandholm, T. Superhuman AI for heads-up no-limit poker: Libratus beats top professionals. *Science* **359**, 418–424 (2018).
7. Mnih, V. et al. Human-level control through deep reinforcement learning. *Nature* **518**, 529–533 (2015).
8. Karpatne, A. et al. Theory-guided data science: a new paradigm for scientific discovery from data. *IEEE Trans. Knowl. Data Eng.* **29**, 2318–2331 (2017).
9. Leibo, J. Z. et al. View-tolerant face recognition and hebbian learning imply mirror-symmetric neural tuning to head orientation. *Curr. Biol.* **27**, 62–67 (2017).
10. Fischer, C. C., Tibbetts, K. J., Morgan, D. & Ceder, G. Predicting crystal structure by merging data mining with quantum mechanics. *Nat. Mater.* **5**, 641–646 (2006).
11. Cho, I. A framework for self-evolving computational material models inspired by deep learning. *Int. J. Numer. Methods Eng.* **120**, 1202–1226 (2019).
12. Janoos, F., Denli, H., & Subrahmanya, N. Multi-scale graphical models for spatio-temporal processes. *Advances in Neural Information Processing Systems*, Edited by Z. Ghahramani et al., Neural Information Processing Systems Foundation, Inc., 316–324 (2014).
13. Raissi, M., Perdikaris, P. & Karniakakis, G. E. Physics informed deep learning (Part II): Data-driven discovery of nonlinear partial differential equations. Preprint at <https://arxiv.org/abs/1711.10566v1> (2017).
14. Kingma, D. P. & Welling, M. Auto-encoding variational Bayes. Preprint at <https://arxiv.org/abs/1312.6114v10> (2013).
15. Champion, K., Lusch, B., Kutz, J. N. & Brunton, S. L. Data-driven discovery of coordinates and governing equations. *Proc. Natl. Acad. Sci.* **116**, 22445–22451 (2019).
16. Tenenbaum, J. B., de Silva, V. & Langford, J. C. A global geometric framework for nonlinear dimensionality reduction. *Science* **290**, 2319–2323 (2000).
17. van den Oord, A. et al. WaveNet: a generative model for raw audio. Preprint at <https://arxiv.org/abs/1609.03499v2> (2016).
18. Li, Q., Peer, A., Cho, I., Biswas, R. & Kim, J. Replica molding-based nanopatterning of tribocharge on elastomer with application to electrohydrodynamic nanolithography. *Nat. Commun.* **9**, 974 (2018).
19. Li, Q., Cho, I., Biswas, R. & Kim, J. Nanoscale modulation of friction and triboelectricity via surface nanotexturing. *Nano Letters* **19**, 850–856 (2019).
20. Lin, S. et al. Electron transfer in nanoscale contact electrification: effect of temperature in the metal-dielectric case. *Adv. Mater.* <https://doi.org/10.1002/adma.201808197> (2019).
21. Riedo, E. et al. Interaction potential and hopping dynamics governing sliding friction. *Phys. Rev. Lett.* **91**, 084502 (2003).
22. Persson, B. N. J. et al. On the nature of the static friction, kinetic friction and creep. *Wear* **254**, 835–851 (2003).
23. Jansen, L. et al. Temperature dependence of atomic-scale stick-slip friction. *Phys. Rev. Lett.* **104**, 256101 (2010).
24. Wood, S. *Generalized Additive Models: An Introduction with R* (CRC Press, Boca Raton, FL, 2006).
25. Duchi, J., Hazen, E. & Singer, Y. Adaptive subgradient methods for online learning and stochastic optimization. *J. Mach. Learn. Res.* **12**, 2121–2159 (2010).
26. Gu, C. *Smoothing Spline ANOVA Models* (Springer Science and Business Media, New York, 2013).
27. Tang, W. & Tong, Y. G. Improved genetic algorithm for design optimization of truss structures with sizing, shape and topology variables. *Int. J. Numer. Methods Eng.* **62**, 1737–1762 (2005).
28. Amirjanov, A. The dynamics of a changing range genetic algorithm. *Int. J. Numer. Methods Eng.* **81**, 892–909 (2010).
29. Koza, J. *Genetic Programming: On The Programming of Computers By Means of Natural Selection* (MIT Press, Cambridge, MA, 1992).
30. Cho, I., Song, I. & Teng, Y. Numerical moment matching stabilized by a genetic algorithm for engineering data squashing and fast uncertainty quantification. *Comput. Struct.* **204**, 31–47 (2018).
31. Beck, J. & Katafygiotis, L. S. Probabilistic system identification and health monitoring of structures In *Proc. 10th World Conf. on Earthquake Engineering* (1992).
32. Jaynes, E. T. Prior probabilities. *IEEE Trans. Sys. Sci. Cybernetics* **4**, 227–241 (1968).

Acknowledgements

This work was supported by the National Science Foundation under grants CMMI-170648 (R.B., J.K., Q.L., in part), CBET-1605275 (J.K., Q.L., I.C., in part), and OAC-1931380 (I.C., in part). The research reported in this paper is partially supported by the HPC@ISU equipment at Iowa State University, some of which has been purchased through funding provided by the NSF under MRI grant CNS-1229081 and CRI grant 1205413.

Author contributions

All the authors contributed to the writing of the manuscript; nano-scale experimental data collection and analyses were conducted by Q.L., R.B., and J.K.; I.C. is responsible for the computational algorithms and implementations.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s42005-020-0339-x>.

Correspondence and requests for materials should be addressed to I.H.C.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons

Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020