

## ARTICLE OPEN



# Lossless multi-scale constitutive elastic relations with artificial intelligence

Jaber Rezaei Mianroodi<sup>1,2</sup>✉, Shahed Rezaei<sup>3</sup>, Nima H. Siboni<sup>1,2</sup>, Bai-Xiang Xu<sup>3</sup> and Dierk Raabe<sup>1</sup>

A seamless and lossless transition of the constitutive description of the elastic response of materials between atomic and continuum scales has been so far elusive. Here we show how this problem can be overcome by using artificial intelligence (AI). A convolutional neural network (CNN) model is trained, by taking the structure image of a nanoporous material as input and the corresponding elasticity tensor, calculated from molecular statics (MS), as output. Trained with the atomistic data, the CNN model captures the size- and pore-dependency of the material's elastic properties which, on the physics side, derive from its intrinsic stiffness as well as from surface relaxation and non-local effects. To demonstrate the accuracy and the efficiency of the trained CNN model, a finite element method (FEM)-based result of an elastically deformed nanoporous beam equipped with the CNN as constitutive law is compared with that obtained by a full atomistic simulation. The trained CNN model predicts the elasticity tensor in the test dataset with a root-mean-square error of 2.4 GPa (3.0% of the bulk modulus) when compared to atomistic calculations. On the other hand, the CNN model is about 230 times faster than the MS calculation and does not require changing simulation methods between different scales. The efficiency of the CNN evaluation together with the preservation of important atomistic effects makes the trained model an effective atomistically informed constitutive model for macroscopic simulations of nanoporous materials, optimization of nanostructures, and the solution of inverse problems.

*npj Computational Materials* (2022)8:67; <https://doi.org/10.1038/s41524-022-00753-3>

## INTRODUCTION

The laws of mechanics are rather well established at the macroscopic continuum scale. This applies particularly to the linear relationship between stress and strain, such as cast in Hooke's linear elasticity law nearly 350 years ago. The material-specific properties are represented through the elastic stiffness modulus, which in three dimensions is a fourth-order tensor. However, when leaving the idealized continuum description of bodies and zooming into their underlying micro-cosmos, also referred to as microstructure, a very complex landscape emerges which is characterized by a wide range of structural features and defects. Some of these microstructure features do not alter the macroscopic stiffness substantially or only in a rather modest and linear fashion, but others lead to more drastic changes in the elastic response. This applies particularly to the porosity of the material, a property that refers to the volume fraction, dispersion, and connectivity of the open volume. In such cases, the elastic modulus of the representative volume element of a material can be altered quite substantially. The relationship between porosity and elastic stiffness as well as its size-dependence across several scales is of importance for several material classes. One example is bone, where diseases such as osteoporosis lead to a loss in bone mass through the highly dispersed micro-architectural deterioration of bone tissue, an effect that entails decay in elastic stiffness and bone fragility<sup>1</sup>. A direct relationship between porosity across all size scales and the resulting stiffness and strength also applies to wood, which is currently gaining rapid momentum as a sustainable construction material<sup>2</sup>. Another example is the development of nano- and micro-porosity during the direct reduction of iron ores<sup>3</sup>. This effect leads to a gradual loss in stiffness which has severe consequences for the design of

corresponding direct reduction and fluidized bed reactors, a field of utmost relevance for the carbon-free reduction of metal oxides.

Also, many functional materials have a nano- or micro-porous structure, such as aerogels, many catalysts or supercapacitors<sup>4</sup>. Most of these materials do not only fulfill their respective functional role but must at the same time bear mechanical loads, where the elastic stiffness assumes a central role in their property portfolio. Another field where porosity and stiffness are closely connected is the domain of additive manufacturing which plays an increasing role in digital manufacturing. Most parts cannot be manufactured to 100% mass density so that better understanding of corresponding stiffness effects due to the inherited porosity is an important aspect<sup>5</sup>.

A size dependency of the elastic properties is particularly expected when it comes to nanoporous materials, which are attributed to the competition between surface and bulk energies, especially at smaller scales<sup>6–8</sup>. Classical models of elasticity should be extended correspondingly through the incorporation of surface elasticity models<sup>9–11</sup>. The reader is referred to refs. 12–15 and the references therein for an overview of surface effects in the case of elasticity and plasticity. One effective approach is to utilize interface finite elements on the surface of models that represent nanostructured materials. The latter method requires explicit input of the surface tension constants and surface elastic parameters in the case of surface elasticity. Also, such surface elements induce implementation and computational complications<sup>13,16</sup>. A size effect is also predicted in the torsion and bending of open-cell foams or lattices, where slender specimens appear stiffer than expected. Such size effects can be predicted by Cosserat elasticity (also known as micropolar elasticity). Such models are sensitive to strain gradients and introduce a characteristic length into the constitutive formulations, which, however, require complex

<sup>1</sup>Microstructure Physics and Alloy Design, Max-Planck-Institut für Eisenforschung, Düsseldorf, Germany. <sup>2</sup>Ergodic Labs, Lohmühlenstraße 65, 12435 Berlin, Germany. <sup>3</sup>Mechanics of Functional Materials Division, Institute of Materials Science, Technische Universität Darmstadt, Darmstadt, Germany. ✉email: [j.mianroodi@mpie.de](mailto:j.mianroodi@mpie.de)

experimental characterization to identify the required constitutive parameters (see also ref. <sup>17</sup>). When it comes to the mechanical and topological homogenization of such materials, further extensions are required to calibrate an effective generalized continuum model that is applicable across several scales<sup>18–20</sup>.

A straightforward way to calculate the true elastic stiffness of nanoporous materials is conducting atomistic calculations which naturally consider all relevant porosity and surface tension features arising from it<sup>21,22</sup>. However, this is not feasible when targeting the mechanical description of larger parts, revealing the typical scale dilemma often encountered in computational materials science. This problem is also often referred to as the multi-scale and multi-physics challenge in materials mechanics<sup>23</sup>.

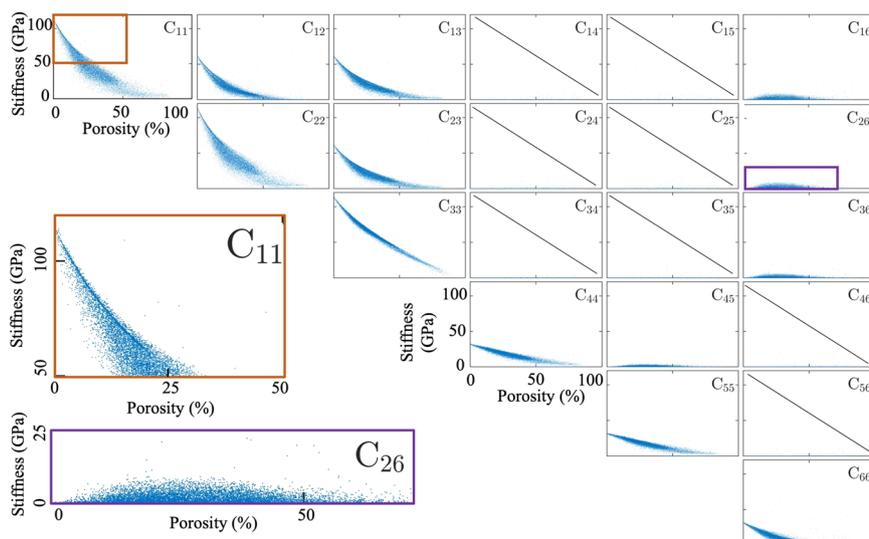
The multi-scale modeling strategy requires, on one hand, accurate and predictive simulation at lower scales taking various physical phenomena into account, and on the other hand, efficient methodologies to transfer the information between the scales. In the available scale-bridging techniques, there is usually a trade-off between the amount of information preserved in the up-scaling and the associated computational costs. One way to establish this micro-macro coupling is to make use of artificial intelligence (AI). Machine learning (ML) seems to provide a promising approach for efficient scale-bridging. As mentioned, the properties of a material depend to a large extent on a wide cosmos of defects, which is also referred to as nano- and microstructure, and on the mutual interaction among all these features. The high-dimensional and tensorial interaction among these multiple nanostructure and microstructure features makes the required dimensionality reduction of the scale transfer problem much more complex than just transferring the chemical composition and some overall geometrical factors from one scale to another. Empirical descriptors that could be predefined to capture and reflect some of these features and their change upon scale transitions are usually unknown and difficult to determine. Instead, in most cases, there are more complex and sometimes even weakly understood interactions hidden inside the material's nanostructure and microstructure. Therefore, image-based AI methods can be valuable tools for the study of microstructure-dependent structure–property relations<sup>24,25</sup> in general (due to the many effects and phenomena involved) and specifically for scale-bridging structure–property calculations (where even some of the physics for adequate coarse-graining have not been resolved yet). In ref. <sup>26</sup>, a summary is provided over recent advances in the application of AI techniques for numerical modeling of various types of materials, such as metals, polymers, ceramics, and composites. Owing to their high efficiency, allowing very fast calculations, AI techniques open up new efficient strategies to optimize and drastically accelerate structure–property calculations of future advanced engineering materials and structures<sup>27</sup>. These approaches could also help to discover scale bridging phenomena that had so far remained elusive, hidden behind the enormous chemical and structural complexity of modern engineering materials. Peng et al.<sup>28</sup> discussed the state of the art of combining ML and multi-scale modeling in various applications. Bock et al.<sup>29</sup> reviewed successful applications of ML and statistical learning methodologies in the field of continuum materials mechanics. They concluded that simulation-based data mining in combination with ML tools provides exceptional opportunities for the identification of fundamental interrelations within materials. Readers are also encouraged to consider reference<sup>30</sup> where stochastic multi-scale methods for modeling complex systems are introduced and discussed. In this work, it has been for example shown that data-driven methods can help to identify the probability density function from Monte Carlo simulation data.

Different ML strategies have been investigated for modeling at different length scales. Wang and Sun<sup>31</sup> replaced the up-scaling procedure through an offline homogenization procedure by utilizing sub-scale simulations to generate a database to train

material models for geological materials. Xue et al.<sup>32</sup> proposed a data-driven multi-scale computational scheme to capture the non-linear mechanical behavior of cellular metamaterials. See also ref. <sup>33</sup> for similar studies. Kumar et al.<sup>34</sup> introduced an ML technique for the inverse design of metamaterials which can generate functionally graded cellular structures with tailored anisotropic stiffness. Wang et al.<sup>35</sup> developed a data-driven method for efficient multi-scale topology optimization. The application of ML has also been extended to more complicated material behavior including plasticity<sup>36,37</sup>. Readers are referred to refs. <sup>25,38</sup> for similar studies in different application fields. It is worth mentioning that the ML approaches can be applied to construct a direct solver for the usually well-known partial differential equations (PDEs) out of massive datasets. Samaniego et al.<sup>39</sup> explored deep neural networks as an option to approximate the solution of the underlying PDEs. Raissi et al.<sup>40</sup> introduced a physics-informed neural network that takes into account any given laws described by general non-linear PDEs. The authors demonstrated the effectiveness of the proposed framework through application to classical problems in fluids, quantum mechanics, and reaction-diffusion systems. Yang et al.<sup>41</sup> employed a conditional generative adversarial neural network to predict stress and strain fields directly from the material microstructure geometry. Wang et al.<sup>42</sup> introduced a genomic flow network and a mosaic flow predictor to estimate the solution of Laplace and Navier–Stokes equations in domains of untrained shapes and boundary conditions. Pandey et al.<sup>43</sup> proposed an ML-based surrogate model for predicting spatially resolved crystal orientation evolution under uniaxial tensile loading. It is also important to make sure that the trained network prediction satisfies the fundamental physical and thermodynamics laws<sup>44</sup> or fulfills the objectivity and possible material symmetry conditions<sup>45</sup>.

In this scientific context, convolutional neural networks (CNNs) are employed extensively to extract material property–structure relationships based on microstructure images<sup>46,47</sup>. Cecen et al.<sup>48</sup> employed a CNN to link a three-dimensional microstructure to its effective (homogenized) properties. The authors showed that the trained CNN can learn physically interpretable microstructure features and accurately predict desired properties. Rao and Liu<sup>49</sup> proposed a three-dimensional deep CNN to predict the anisotropic effective material properties for representative volume elements with random inclusions. The dataset generated by a computational homogenization approach was used for training the network. They concluded that the trained networks can predict unseen data, indicating that the network is capable of capturing the microstructural features of the system and could produce an accurate prediction of the effective anisotropic stiffness tensor. The CNN approach has been also applied in several works to study microstructure descriptor relations to material properties<sup>50–53</sup>. These descriptors are usually not readily interpretable in terms of crisp physics mechanisms. However, when carefully applied, such data-driven methods can provide useful material property prediction capabilities. Recently, Mianroodi et al.<sup>24</sup> applied a U-Net approach to predict stress fields in geometrically complex and heterogeneous non-linear elastoplastic material systems. It was shown that the U-Net-based prediction of the stress fields in such a highly non-linear and heterogeneous mechanical system was about 8000 times faster than that obtained by a typical spectral solver (e.g., ref. <sup>54</sup>). Interestingly, the U-Net was also capable to reproduce the stress distribution in geometries topologically far from those that had been used for training.

Although the body of literature on multiple possible and promising applications of ML in material science is growing fast, a general AI-assisted concept for transferring constitutive relations directly from atomistic to continuum scale is missing. There have been recent works on the application of ML for the prediction of elastic stiffness for multi-scale modeling<sup>55,56</sup>. However, these are



**Fig. 1** Components of the elasticity tensor calculated with molecular statics. The elasticity components for 18,172 atomistic boxes with randomly generated pore structure as a function of the box porosity volume fraction, defined in Eq. (4). All the plots (except the zoomed versions of  $C_{11}$  and  $C_{26}$  at the bottom left) have the same range in horizontal (porosity) and vertical (stiffness) axes.

typically designed to homogenize the continuum scale elasticity calculations that do not include atomistic effects. A direct calculation of the elasticity tensor from atomistic simulations as data source for ML has not been studied before to the best knowledge of the authors. Furthermore, most conventional multi-scale approaches involve loss of information since in each step of the up-scaling, some level of approximation or averaging (also referred to as homogenization) has to be usually introduced to reduce model complexity. This is in particular the case when discrete atomistic data is used at mesoscopic or macroscopic scales. Here, we show that the trained network is capable of capturing surface and size effects directly from the atomistic scale, practically replacing the formerly required computationally expensive atomistic simulation for stiffness calculations. The ML approach seems to transfer the information (elasticity tensor in this case) from microscopic scale to macroscopic scale without any loss (i.e., without losing the atomistic effects). Note that in data transmission, such methods are called “lossless,” as opposed to “lossy” methods of data transfer or compression. Analogously, we call the current approach a lossless multi-scale method. In the current work, we use Mishin’s interatomic potential for Al<sup>57</sup> to calculate the elasticity tensor for large sets of randomly generated nanoporous structures. A CNN is then trained using the topology data (stored as images of the atomic structure) and the calculated elasticity tensor components as a reference dataset. Once the network is trained, it is evaluated in terms of several test cases, showing its ability to capture full atomistic details (such as pore surface effects) while being orders of magnitude faster than the atomistic calculations. The efficiency of the AI-based method coupled with an accurate description of the atomistic scale effects on the material’s elastic constitutive response results in a lossless scale-bridging approach. The training data from the atomistic calculations as well as the performance and prediction capability of the trained CNN are presented in the Results section, followed by the Discussion section. In the “Methods” section, details about preparing the atomistic data, network architecture, as well as training are discussed.

## RESULTS

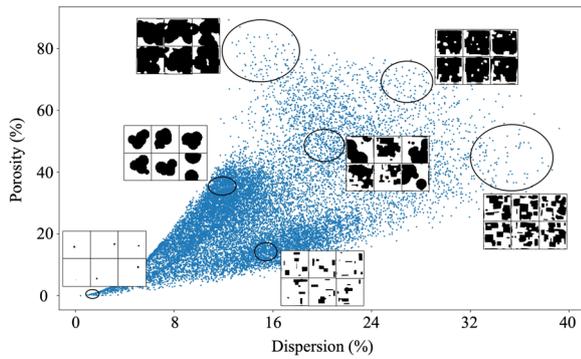
### Atomistic results

From the randomly generated structures (as explained in the “Methods” section) and the resultant atomistically simulated

elasticity tensors, only those with positive definite values are selected. Since the pore structures are random, in some cases the geometry is not stable, leading to incorrect elasticity calculations. After filtering these unacceptable cases based on the approach explained in the “Methods” section, 18,172 samples with physically meaningful input and output data are used for training the CNN. All of this data is visualized in Fig. 1. Note that the randomly shuffled dataset is divided into training and validation subsets with sizes of 17,172 and 1000, respectively. For testing the prediction capability of the trained network, an additional set of 2572 randomly generated cases are employed. This test dataset is not used in the training or validation of the network.

As observed in Fig. 1, the dispersion of the data is increasing as the porosity is increased. This is expected as there are more random ways to create the same porosity volume fraction when it is increased. Also, note that more of the cases in high porosity regions are filtered out due to the instability of the generated structures. Therefore, the number of data for structure regimes with higher porosity is lower compared to those with low porosity. As it will be explained in the method section, randomly generated structures with higher porosity fractions are often unstable and cause issues for the automated elasticity tensor calculations. Therefore, there is an unbalanced distribution of data at different porosity levels. As will be shown later, this results in reduced prediction accuracy for the higher porosity levels with a lower number of training data. One possible remedy is to increase the weight of the higher porosity data in the training. We have not implemented this in the current study since, although the filtering scheme removes most of the unphysical atomistic simulations, we expect lower reliability of the data obtained for higher porosity levels.

As it is shown in Fig. 1, only 13 components are non-zero (within a tolerance of  $10^{-3}$  GPa), instead of 21 for a general elasticity tensor. This is due to the assumed two-dimensional (2D) variation of the pore structures and translates to the introduction of a symmetry-breaking effect, rendering the structure response monoclinic (rather than cubic) with a plane of symmetry<sup>58</sup>. Note that the Voigt notation, i.e.,  $(11, 22, 33, 23, 13, 12) \rightarrow (1, 2, 3, 4, 5, 6)$  is employed. Therefore the CNN has 13 scalar outputs for one input image as summarized above. The magnified views of the  $C_{11}$  and  $C_{26}$  plots are also shown in Fig. 1. As apparent for the  $C_{11}$  plot, there is a high density of points along a curve from 0% porosity up to about 20%. This high point density portion of the curve is



**Fig. 2** Distribution of the input data in the feature space. Example clusters of the 18,172 randomly generated nanoporous structures used in the training and the distribution of the data in porosity fraction (Eq. (4)) vs. dispersion (Eq. (5)) space. The black regions are the empty space and the white regions are the fcc lattice.

visible as a line shadow in this region, which seemingly disappears for porosity values above 20%. One possible explanation is the percolation threshold, i.e., as the porosity volume fraction increases, at a certain threshold (which highly depends on the structure) most of the pores will be connected, increasing the dispersion in the data.

For the  $C_{26}$  elastic constant, the maximum dispersion of the data is around 20–25% porosity while the data points are closer to each other for porosity values around 0% and above 50%. A similar trend is visible for other components such as  $C_{16}$  and  $C_{36}$ . This could be also related to the percolation threshold discussed above.

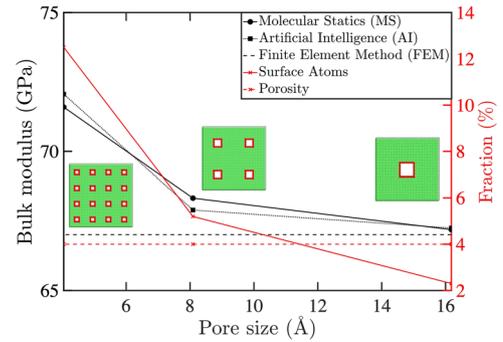
We further study the distribution of all training data in the space of porosity volume fraction (Eq. (4)) vs. the pore dispersion (Eq. (5)) as shown in Fig. 2. As expected, the pore structures based on the random circular shapes have lower dispersion (lower surface to bulk atoms) compared to the rectangular-shaped pores at the same pore volume fraction levels. High dispersion parts of the data are almost entirely made of interconnected rectangular pores (dispersion of about 35%), while the highest pore volume fraction (about 80%) in the data consists of interconnected circular pores at a relatively moderate dispersion of about 16%. In the intermediate regime (i.e., the porosity of about 50% and dispersion of about 20%), the data consists of a mixed combination of circular and rectangular-shaped structure features as seen in Fig. 2.

### Benchmarking AI predictions

The trained network is benchmarked against 2572 randomly generated cases of the test dataset. These data were not included in the training or the validation of the training. The root-mean-square error (RMSE) of the trained network on the test dataset is about 2.4 GPa, which is comparable to the RMSE of 2.8 GPa on the validation dataset. These correspond to about 3.0 and 3.5%, respectively, of the equivalent bulk modulus (79.0 GPa) of single-crystalline Al. However, besides the 2572 cases of the atomistic data reserved for the test dataset as explained above, we also look into a few special cases here to evaluate and interpret the performance of the network. Note that these cases were neither part of the training nor the validation datasets.

### Size effects

One major difference between continuum mechanics and atomic systems is the non-local nature of the interactions in the latter. Due to the quantum-mechanical nature of their electronic bonding, atoms (particularly in metals) interact over a wide distance (where in simulations a cutoff radius is usually defined) as



**Fig. 3** Size dependency of the bulk modulus. Effective bulk modulus (black) as a function of pore dispersion in three simulation boxes is shown. Classical linear elastic continuum mechanics (solved by FEM, dashed black) shows no size effect, while atomistic calculations (solid black) and AI predictions (dotted black) show a significant size effect. Note that the pore volume fraction (dashed red) is constant for all three cases while the fraction of atoms that are located at the inner surfaces of the pores (solid red) increases as the pores get smaller, an effect that leads to a stiffer response. In this figure and the others to follow, green and red correspond to the atoms with fcc lattice and unknown (surface) structures, respectively, identified using common neighbor analysis as implemented in the freeware visualization and analysis package Ovito<sup>70</sup>.

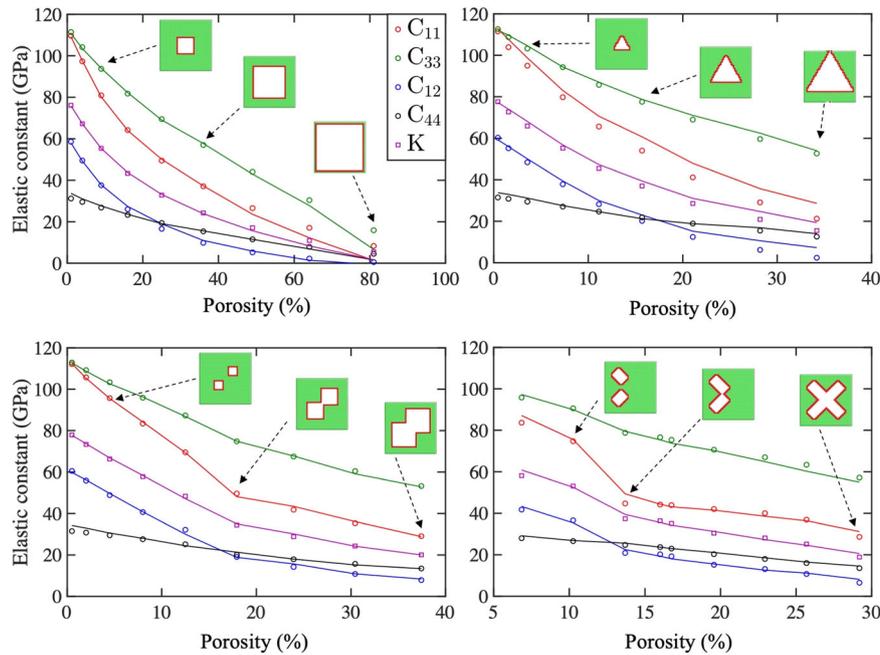
opposed to local material models used in classical solid mechanics. This non-local interaction leads to the distinction between atoms at a free surface with broken neighbor bonds and atoms in the perfect bulk face-centered cubic (fcc) lattice where all bonds are saturated. As systems become smaller, the ratio of atoms at surfaces to the atoms in the bulk increases (see red solid line in Fig. 3). To demonstrate the performance of the trained CNN, we investigate the size dependency of the effective bulk modulus of a system with porosity and compare the results with the molecular statics (MS) simulations, as well as classical continuum mechanics solved using the finite element method (FEM). A linear elastic constitutive law, without any phenomenological size effect, is assumed in the FEM calculations. Note that the bulk modulus in this work is the equivalent bulk modulus of single-crystalline Al with the pore structure.

To this end, three boxes with square-shaped pores are considered as shown in Fig. 3. In all cases, the simulation size and orientation are the same as explained in the “Methods” section. The pore in the simulation box on the right-hand side of the figure has an edge length of  $L_x/5$ , resulting in a porosity volume fraction of  $p = 4\%$ . In all other steps, the edge length of the pore is divided by two (area reduced by a factor of four) while the number of pores is quadrupled. Therefore, the number of removed atoms ( $N - n$ ) and the porosity volume fraction, see Eq. (4), of the three boxes remains unchanged (dashed red line in the figure) while dispersion parameter  $d$ , see Eq. (5), is increased from 56.3 to 125.0% and 300.0% in the three cases, as the pore sizes become smaller.

The same input structure is fed to the trained CNN and the predicted bulk modulus is plotted in Fig. 3 as solid black squares. As observed, the CNN prediction captures the pore size dependency of the bulk modulus with good agreement compared to the MS calculations. Note that the trained CNN is predicting the 13 components of the elasticity tensor as discussed above. The effective bulk modulus is calculated using

$$K = -V \frac{dP}{dV} \approx -V \frac{\Delta P}{\Delta V}, \quad (1)$$

where  $V$ ,  $\Delta P = (S_{11} + S_{22} + S_{33})/3$ , and  $\Delta V = V(1 + E_{11})(1 + E_{22})(1 + E_{33}) - V$  are the volume and the differential of the hydrostatic pressure and volume, respectively. Stress  $\mathbf{S}$  and strain  $\mathbf{E}$  are related through the elastic relation  $\mathbf{S} = \mathbf{CE}$ . Under the small deformation



**Fig. 4** Benchmarking the AI prediction of the elasticity tensor. Atomistic calculations (markers) and AI predictions (solid lines) for selected components of elasticity tensor and effective single-crystal bulk modulus for cases with simple geometries of one square pore (top left), triangular pore (top right), two square pores (bottom left) and two rectangular pores (bottom right).

assumption and with initially undeformed/stress-free conditions, the differentials can be approximated as

$$\Delta V = 3\epsilon V, \quad \Delta P = (S_{11} + S_{22} + S_{33})/3, \quad (2)$$

where  $E_{11} = E_{22} = E_{33} = \epsilon$  is assumed. Combining Eqs. (1) and (2), we get

$$K \approx -\frac{1}{9\epsilon} (S_{11} + S_{22} + S_{33}). \quad (3)$$

The bulk modulus is then obtained by employing Eq. (3) and applying an equal triaxial strain state on the system. Such displacement boundary conditions were applied for both, the MS and the linear elastic FEM calculations. In the FEM case, the solid phase is equipped with the anisotropic elastic property of single-crystalline Al while the pore has zero stiffness. Note that in all the simulations and results reported in this paper, we work with single-crystalline material and a full anisotropic elasticity model. Homogenized polycrystalline properties could be extracted from these for any given grain structure<sup>59</sup>.

The classical continuum prediction for the single-crystal bulk elastic modulus of this system is size-independent (dashed black line in Fig. 3) with the value of 67.0 GPa. When comparing these three structures using classical continuum mechanical computations, the overall bulk modulus remains the same. Also, note that in all these cases the pore volume fraction is kept constant and the box is periodic. If we change the pore volume fraction, we will see a difference in the responses even by employing classical FEM computations.

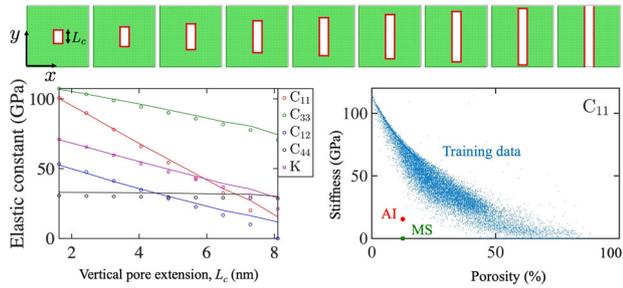
It is worth mentioning that by utilizing enhanced non-local continuum methodologies where surface effects are also considered (e.g., refs. <sup>10,13,16,60</sup>) one can also use FEM calculation for data generation. For a proper comparison, these models should be first calibrated based on atomistic calculations<sup>16</sup>. Moreover, the atomistic origin and the related description of size dependency or more generally nanostructure dependency of elasticity can go beyond the capability of the existing phenomenological constitutive laws. On the other hand, the MS calculations show a size effect in the bulk modulus with an increase from 67.2 to 71.6 GPa (6.5% increase) as the pores get

smaller from 16.1 to 4.0 Å (solid black line in the figure). This is correctly captured by the trained CNN with a relative error of about 0.6% compared to the MS-based values. Note that this error is for the prediction of the stiffness tensor in the simulation boxes with the special simple case of the square-shaped pore discussed above. The RMSE of the trained network in the test dataset is 2.4 GPa (or 3.0% of the bulk modulus).

### Porosity and connectivity

Here, we systematically evaluate the CNN-based prediction of various components of the single-crystal elasticity tensor against the MS results for special cases shown in Fig. 4.

As seen in Fig. 4 top left for the case of a square-shaped pore, the agreement between MS (marked points) and CNN prediction (solid lines) is excellent. For example, considering the  $C_{33}$  results, the relative error starts from 0.6% for the smallest pore size and increases to 8.7% for the second-largest pore (before the last one). The CNN prediction deviates about 58% from the atomistic calculation results in the last point of the comparison at a very large porosity of 81%. The deviation is attributed mainly to the reduced number of training data provided at such high porosity levels as discussed. The CNN prediction also shows a relatively good agreement with the MS results in the system with a triangular pore as shown in the right top panel of Fig. 4. Although the error of 1.1% for the initial microstructure is increased to 34% for the largest pore, the network still captures the important trend in the results. The higher error compared to square-shaped pores in the previous case is due to the fact that no triangular or similar-shaped pores were included in the randomly generated training dataset, as opposed to random circular and rectangular-shaped pores. In the case of two growing squares in the bottom left panel of this figure, an interesting change in the slope of the plot (in particular for  $C_{11}$ ) is visible. This change in the slope is due to the two squares reaching each other and overlapping. The change in slope is correctly captured by the CNN prediction signifying the effectiveness of the trained network for including effects of topology, particularly of pore percolation. Since the maximum porosity in this case is up to 37%, the CNN prediction agrees much



**Fig. 5 Elasticity tensor components of a box with growing pore.** Top row: simulation boxes with the vertically growing rectangular pore of length  $L_c$ . Bottom left: MS calculation (markers) and AI prediction (solid lines) for single crystal elasticity tensor components in a vertically growing rectangular pore as a function of  $L_c$ . Bottom right: the  $C_{11}$  training data as well as the MS calculation and AI prediction for the case of  $L_c = L_y = 8.0$  nm.

better (compared to the previous cases) with the MS results with a maximum relative error of 2.2%. A similar test for two rectangular pores of increasing size has been performed with the results shown at the bottom right panel of Fig. 4. Similar to the previous case, the overlapping of the two rectangles leads to changes in the slope of the plots. The agreement between CNN predictions and MS results is quite good with a maximum error of about 5.0%. The changes in the slope of the plots indicate the effect of connectivity of the pores on the data, which seems to be captured by the CNN correctly.

Next, we study a case of a vertically growing pore with both, CNN and MS calculations. As shown in Fig. 5, this case is analogous to a crack growth scenario, where depending on the crack orientation, a strong asymmetry for the elasticity components is expected. Indeed we find that some components of the elasticity tensor drastically decrease (e.g.,  $C_{11}$ ) while others remain almost constant (e.g.,  $C_{44}$ ) or decrease with a lower rate (e.g.,  $C_{33}$ ), as the vertical pore length increases.

According to Fig. 5 bottom left, the agreement between MS and CNN-based prediction is apparent for smaller pore lengths. The trained model is also sensitive to the pore orientation and captures the stiffness degradation correctly. As the pore becomes longer in the vertical direction and therefore gets closer to the box boundaries, the error of the CNN prediction increases. In the limit of a fully separated box ( $L_c = 8.0$  nm), as expected, the MS prediction for  $C_{11}$  provides a complete loss in stiffness, whereas the AI prediction for this case is  $C_{11} = 15.4$  GPa. This case, as shown in Fig. 5 at the bottom right, is completely out of the range of the training data. However, the AI prediction, although associated with the high error, is also out of the training data range and qualitatively points in the correct direction. For better evaluation of fully grown pores (which could for instance mimic cracks or delamination features) and separated boxes, the training data should include such cases. However, in this work, we only focus on cases with non-zero stiffness and non-fractured boxes.

### Lossless multi-scale modeling

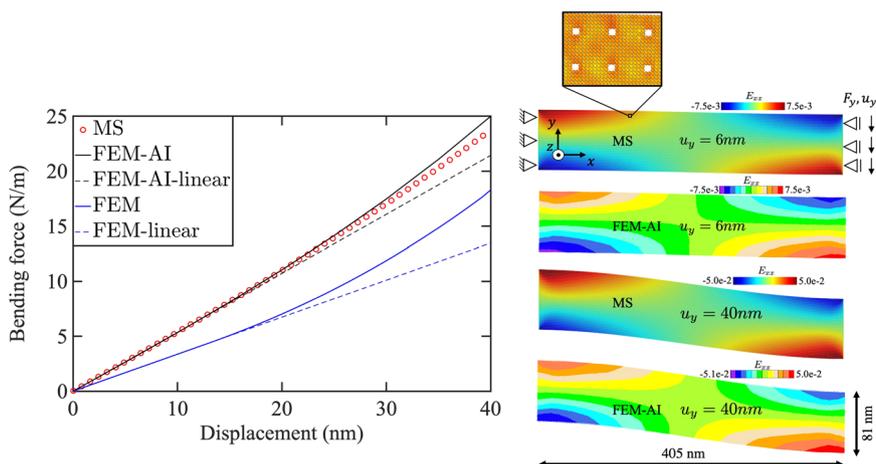
The CNN prediction of the elasticity tensor has been incorporated in a FEM as a continuum scale constitutive law to simulate the bending of a nanoporous beam. The beam dimensions are  $(l_x, l_y, l_z) = (405.0, 81.0, 1.2)$  nm with free surfaces in  $x$  and  $y$  and periodic boundary condition along the  $z$  axis. The left side of the beam is fixed in all directions while the right side is displaced step-wise in the  $-y$  direction up to  $|u_y| = 40$  nm. The atomistically modeled beam has square-shaped pores with side length and spacing of 0.40 and 1.46 nm, respectively, and consists of 2,304,480 atoms. Note that the porosity of the beam corresponds to the smallest pore size studied in Fig. 3.

The same beam geometry is simulated by FEM in conjunction with a full anisotropic elasticity tensor predicted by AI for this porosity. Here, the standard finite element model for elasticity is utilized where the anisotropic elastic stiffness tensor is simply an input for the FEM. For the case of small deformation, the elastic stiffness tensor in Voigt notation is denoted by  $C$ . According to Hook's law, we have  $\tilde{\sigma} = C\tilde{\epsilon}$ , where the stress and strain tensors in Voigt notation are denoted by  $\tilde{\sigma}$  and  $\tilde{\epsilon}$ , respectively. The component of the elastic stiffness matrix is directly obtained from the output of the trained CNN. The results of such a model are represented by "FEM-AI-linear." Using the Green–Lagrange strain tensor, we end up with "FEM-AI" where the underlying atomistic geometrical non-linearity of the strain tensor is considered. The beam is then loaded in the simulation up to  $|u_y| = 40$  nm. The resulting force-displacement curve from FEM-AI agrees well with the one obtained from the full atomistic simulations (MS). Note that, as shown in Fig. 3, this pore size introduces a significant surface/size effect compared to a continuum pore. As seen from the results, using the AI-based constitutive model, seems to correctly transfer the mechanical behavior of the system from the atomistic scale into a continuum model.

The results of the FEM simulations obtained by using Hooke's law for the linear elastic constitutive response, i.e., without the AI-based size-dependent elastic model, are shown in Fig. 6 left with the blue curves, for the same pore distribution. As it is seen here, ignoring the size dependency of the elastic stiffness introduces a significant error in the predicted bending behavior of the beam. For example, at  $|u_y| = 20$  nm, MS and FEM-AI both predict a bending force of 11.0 N/m while the size-independent FEM calculation predicts a bending force of 7.0 N/m. Note that the surface effects increase the stiffness of the system as shown in Fig. 3. Therefore, as expected, the atomistic simulation and the FEM simulation in conjunction with the AI-based elastic constitutive model predict higher forces for the same beam displacement compared to the FEM with no surface effects. The deviation (about 5.2% at  $|u_y| = 40$  nm) between the MS and the FEM-AI simulations for larger displacements is partly due to the increased pore deformation and its effect on the elasticity tensor. The FEM-AI simulation in its current version does not update the pore shape to reevaluate the elasticity tensor during the loading. This will be straightforward to implement and expected to resolve the deviation at larger deformation values.

For a better comparison and to emphasize the potential of the AI-based multi-scaling constitutive modeling capability, we compared the computational time required for different approaches as well as different steps of the training in Table 1. In the first row, the required computational times for training the network (as explained in the "Methods" section) are listed. Next, the computation time for one elasticity tensor calculation with MS and AI is compared. The bottom two rows of the table list the computation times required for the nanoporous beam bending simulations as discussed above.

As listed in Table 1, although the training time of the CNN is about 2 days, the evaluation of the trained network for the prediction of the full elasticity tensor is about 230 times faster than the MS calculations. This significant speed-up in elasticity tensor calculation will benefit FEM-MS-based multi-scale modeling, where each integration point of the elements in the continuum scale FEM model is represented by an atomistic simulation box. Since in the current simple example of a uniform pore structure in the beam exposed to a bending load, an on-the-fly evaluation of the AI-based elasticity tensor was not necessary, the FEM method with the AI-calculated constitutive relation (FEM-AI) is significantly (about five orders of magnitude in this particular example) faster than the full atomistic (MS) calculation, while still capturing important size effects as shown in Fig. 6. Note that the speed-up in this case highly depends on the problem complexity



**Fig. 6 Nanoporous beam under bending load modeled with different methods.** Comparison of force per unit length in  $z$  of the beam as a function of displacement in  $y$  (left) based on fully atomistic simulations (red markers) and a large deformation finite element calculation equipped with the AI-based constitutive elastic relation (black line), as well as the FEM model with size-independent constitutive relation (blue line). The dashed lines are the system response ignoring the large deformation effects. Strain fields calculated at two different beam deflections of 6 and 40 nm based on MS and FEM-AI are shown on the right side.

**Table 1.** Computational time for training the network and a single elasticity tensor calculation with MS and AI, as well as the beam bending test (bottom two rows) with different approaches.

	Time (sec)	System
AI training	$2.0 \times 10^5$ (~2 days)	NVIDIA Tesla P4 GPU
Elasticity tensor calculation		
MS	101	Intel i7-860 2.80 GHz (single core)
AI	0.43	Intel i7-860 2.80 GHz (single core)
Nanoporous beam bending test		
MS	$2.03 \times 10^6$ (~23 days)	Intel Xeon 6150 2.70 GHz (x2, 36 cores)
FEM-AI	35.5	1 CPU (averaged over 10 simulations)

and the numbers reported in the lower part of Table 1 are only representative of the current case and presented here as the first point of reference. The main speed-up reported is the calculation of the fully anisotropic elasticity tensor with AI as compared to the time-consuming MS methods. It should be also considered in this context that the elastic response is the most simple, fast, and straightforward mechanical response feature to predict in such atomistic calculations. When addressing more complex size-dependent nanostructural features, such as the inelastic response of materials under load, atomistic calculations are computationally much more costly, so that corresponding AI approximations of the size dependence (or other types of) underlying and homogenized constitutive response might profit even more from the use of such methods.

## DISCUSSION

In this work, we show a lossless and efficient approach to using AI techniques to derive constitutive laws for multi-scale modeling of materials with complex nanostructures. Two ideas are presented, namely, (1) using computations at the atomistic scale and transferring the information to the continuum level to arrive at physics-motivated model parameters without any ad hoc assumptions or empirical approximations. (2) A CNN is used to directly relate the image of an arbitrary complicated material structure to its homogenized properties. More specifically, we focused on

obtaining the anisotropic elastic properties of a nanoporous aluminum single crystal, as a reference model substance, which can be well simulated by atomistic methods, owing to the good quality of the existing potentials and the absence of magnetic effects. Random porous structures are generated and their homogenized anisotropic elasticity tensor is calculated using MS. We show that all the non-local effects arising from the physical nature of the material's response at the atomistic scale (i.e., surface and size effects) are captured in this step. The complex pore structure and the details of the atomistic reconstructions at the pore surfaces necessitate a neural network design that can detect the local properties and aggregate that information for use at a coarse-grained scale. The CNNs are one of the few appropriate candidates to achieve this flow of information. The trained CNN confirms this by adequately predicting the elasticity tensor of completely new porous structures which were not included in the training set. Finally, the proposed methodology is applied to a simple structural problem of a beam under bending load. We show that by using a FEM simulation in conjunction with the AI-based constitutive relations, one can efficiently and accurately predict the material's structural behavior (in the elastic regime, with full size dependence) with the accuracy of atomistic simulations. We also observed that the trained AI can predict anisotropic elastic properties about 230 times faster compared to performing corresponding MS calculations. Based on the latter observation, one can also conclude that such materials-related accelerated and lossless AI-based multi-scale modeling will have at least the same amount of speed-up compared to a conventional hierarchical multi-scale simulation combining FEM and MS.

Although the performance of the proposed approach using the investigated dataset is significant, there is room for further improvements. As an example, it would be interesting to enrich the current training dataset by including more cases with much higher porosity, including zero stiffness cases as well as open-cell topologies. Finally, instead of training the symmetric part of the stiffness tensor, one can try to predict the Cholesky factor of a tangent stiffness matrix to impose a weak convexity on the strain energy function<sup>61</sup>. The latter point is specifically interesting when it comes to non-linear material behavior. It should be noted that in the current manuscript, as a starting work, only simplified cases of randomly generated microstructure patterns are employed for training the neural network. The current approach could be extended to include three-dimensional pore structures with closer

geometric shapes to the experimentally observed porosity in specific materials and applications.

The present work opens up many opportunities for efficient and yet accurate inverse design strategies. For future work, the current methodology can be extended to study similar engineering materials with complex inherent porosity features such as metamaterials, biological matter, additively manufactured materials, batteries, or foams<sup>62–64</sup>. Investigations should be done to compare the introduced AI-based model with more advanced non-local elasticity models such as introduced by the micropolar<sup>60</sup> or peridynamics<sup>65</sup> theory frameworks. Furthermore, instead of finding the symmetric part of the stiffness tensor as an output, one can focus on obtaining the total scalar elastic energy as a function of the given deformation gradient. In this manner, one can also derive physics-based material non-linearities for the hyperelastic response. The proposed approach may also be extended to perform studies on arbitrary (nano-) composite volume elements with various phases. As a result, one can obtain not only the mechanical properties (e.g., stiffness tensor) but also other physical properties such as effective thermal conductivity or the mobility of chemical species in such complex types of matter. The reported collection of data can also help to understand how the porosity with an arbitrary shape influences the overall stiffness. Such information could be further used for developing damage models at coarse-grained scales. In other words, the structure-dependent change in the stiffness tensor, for instance, due to creep porosity or micro-fracture patterns, can be used to represent the level of damage or degradation due to the change in porosity within each material point. Finally, other fundamental and essential properties such as direction-dependent fracture energy<sup>66</sup> may be extracted from atomistic simulations. The anisotropic fracture energy and elastic properties can be directly passed to macroscopic phase-field damage models at larger scales for conducting efficient calculations based on atomistic data.

## METHODS

In this section, the general workflow to create the atomistic data, the network architecture, and the training procedure are explained. The inputs to the atomistic part are the initial box size, the desired ranges of the porosity volume fraction and shapes as well as the number of random

configurations. The outputs are the images of the atomic structure of the porous simulation box and the numerical value of the elasticity stiffness tensor. The network input is the images, and the output is the components of the elasticity tensor.

## Workflow and preparation of atomistic data

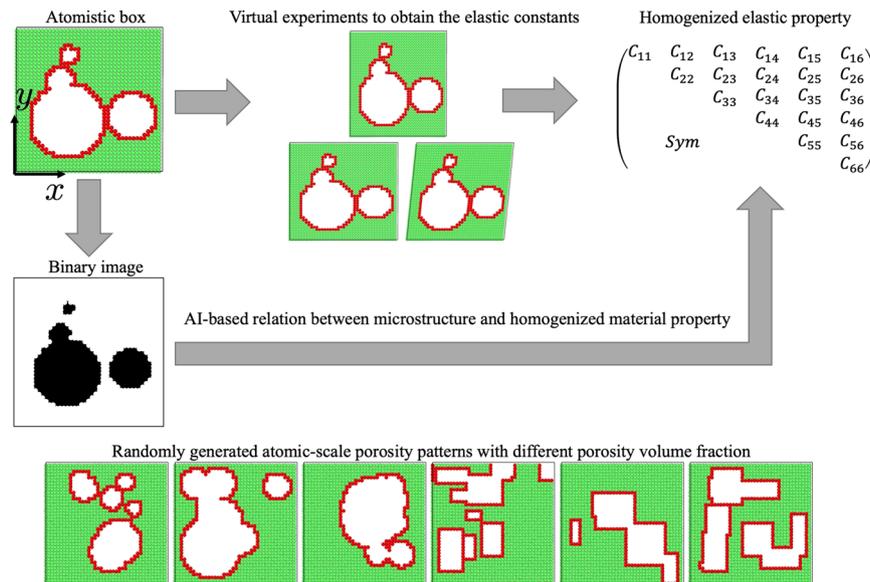
An important step in training a neural network consists in obtaining an accurate and diverse set of training data. In this work, we use MS to calculate the anisotropic elastic constants of fcc aluminum with randomly distributed porosity. The Mishin embedded atom potential<sup>57</sup> in conjunction with LAMMPS<sup>57</sup> is employed to calculate the elastic constants. Note that this interatomic potential is very accurate in the prediction of elastic properties of Al. The bulk modulus predicted by the potential is within a 3% error of density functional theory (DFT) calculations<sup>68</sup>. The surface energies generally agree with the values reported in the literature, however, with larger errors of about 20%<sup>57</sup>. The fully periodic simulation box size is set to  $(L_x, L_y, L_z) = (20, 20, 3)a_0$  where  $a_0 = 4.05 \text{ \AA}$  is the lattice constant. The schematics of the method along with exemplary nanoporous structures are shown in Fig. 7.

The box is periodic in all directions, representing an infinite array of the pore structure. Initially, the box is filled with  $N = 4800$  Al atoms. Then a random target porosity volume fraction in the range of (0, 90) percent is selected. The porosity volume fraction is defined as

$$p = \frac{N - n}{N} \times 100, \quad (4)$$

where  $n$  and  $N$  are the numbers of atoms in the porous and non-porous (filled) boxes, respectively. Note that  $p$  is 0% and 100% in the case of a full simulation box ( $n = N$ ) and empty space ( $n = 0$ ), respectively.

The procedure to generate the atomistic data is as follows: first a target porosity volume fraction, for brevity, referred to as porosity in this work, is randomly selected. Then a pore with either a rectangular or a circular shape is created by removing the atoms from the box using the LAMMPS "region" and "delete\_atoms" commands. The pore geometry and position are selected randomly with the edge lengths in the range of  $(L/64, L/3)$  where  $L = L_x = L_y$  for the rectangular and the radius in the range of  $(L/32, L/4)$  for the circular pore. Then the porosity ( $p$ ) is calculated and compared to the initially selected porosity. If the  $p$  is above or equal to the desired value, the geometry is accepted and the calculation for relaxation and subsequent elastic stiffness follows. Otherwise, another random pore is created as explained above. This procedure is repeated in a loop until  $p$  is above or equal to the desired porosity value. After the desired porosity volume fraction is reached, the atomistic sample is relaxed to a zero stress state using a combination of the conjugate gradient and FIRE algorithms,



**Fig. 7 Schematic of the workflow.** Depicting the elasticity tensor calculation from molecular statics and corresponding image-based artificial intelligence prediction. In the binary image generated from the corresponding atomistic structure, the pore area is indicated in black while the solid material region is in white. Selected examples of the randomly generated circular and rectangular pore structures are shown in the bottom row.

which is a damped dynamics method described in<sup>69</sup>. Once all stress components (including shear components) are zero, the elastic stiffness calculation is performed using the script in LAMMPS which works by distorting the system infinitesimally in different directions and calculating the change of stress, thereby, measuring the components of the elasticity tensor. The details of this script are available in the example folder of the LAMMPS source code.

In addition to the porosity volume fraction introduced in Eq. (4), we introduce an atomistic measure for quantifying the pore dispersion

$$d = \frac{n_s}{N - n} \times 100, \quad (5)$$

where  $n_s$  is the total number of surface atoms, identified using common neighbor analysis<sup>70</sup>. Note that for a given porosity volume fraction  $p$ , i.e., constant  $N - n$ , the dispersion  $d$  is maximized when the pore surface area, i.e., number of atoms at the surface  $n_s$ , is maximized. This typically translates into dispersing the same pore volume fraction into smaller average pore sizes. Note that other quantification measures of porous media such as percolation are also available. In the current case, since the structures are randomly generated, at each porosity volume fraction level, we generate many structures with different percolation and dispersion.

Without loss of generality, we consider 2D pore structures in the  $x$ - $y$  plane, while along the  $z$  direction, no structure variation is assumed, i.e., mimicking an extruded nanostructure. The scheme presented should also work for general 3D nanoporous structures, except that the training of the CNN model becomes computationally more expensive. Thus, only variations in the  $x$ - $y$  plane are important here, resulting in the possibility of representing the atomic positions and topology of the pores by a single image. The image is created after full stress relaxation and serves as the input for the CNN. Note that the images are created using the built-in dump image functionality of the LAMMPS package with the correspondingly adjusted view and zoom settings. The atom diameter in the image render is set to 2.9 Å to make sure the bulk fcc parts of the box are rendered as full white regions. The corresponding outputs are the components of the elasticity tensor.

Two datasets with circular and rectangular pore shapes are generated, respectively. Each dataset consists of about 10,000 cases, resulting in a total of 20,000 sets of an input image and output elastic constants. Exemplary pore structures selected from each of the datasets are shown in the bottom row of Fig. 7. An additional dataset with the size of 3000 (2000 circular and 1000 rectangular pores) is also created and reserved for testing the prediction capability of the trained network. All the atomistic data are filtered to remove unphysical cases resulting from the instability of the random structures or numerical issues in the calculation of the elasticity tensor. A simulation case is removed from the dataset if

1. the determinant of  $C$  is negative, or
2. any of the normal components of the elastic tensor ( $C_{ii}$  for  $i \in 1, 2, 3$ ) are above or below 130 or  $-0.01$  GPa, respectively.

The first condition enforces that the elasticity tensor is positive-definite and thus non-singular (can be inverted into compliance tensor). However, this condition alone does not exclude unphysically large values of the elasticity tensor due to numerical issues in unstable cases. Therefore the second condition is introduced. Note that the maximum value of the normal component of the elasticity tensor for the bulk defect-free material is about 114 GPa. A small negative value of  $-0.01$  GPa is still allowed to account for inherent fluctuations in the atomic calculations. After the application of the above-mentioned filtering scheme on all 23,000 simulation cases, the remaining data is divided into 17,172, 1000, and 2572 cases for training, validation, and testing, respectively. Note that, in addition to ensuring that only physically correct data are used for training the network, one may include the physical constraints directly into the network as well. In this context, the first condition above can be treated as an additional non-equality constraint in the network loss function. However, in the current work, the physical filtering of the training data will implicitly result in a network with physically correct predictions.

The simulation boxes are represented by images with a resolution of  $256 \times 256$  images. At this resolution, each atom is represented by about  $12 \times 12$  pixels, which is more than enough to represent an arbitrary pore structure in the simulation box. The focus in the present work is only on 2D structures. Furthermore, calculation of the elasticity tensor at finite temperature is an important extension that involves post-processing of the results in terms of ensemble averaging of the oscillating atomic positions. However, in the current work, we focus only on static results.

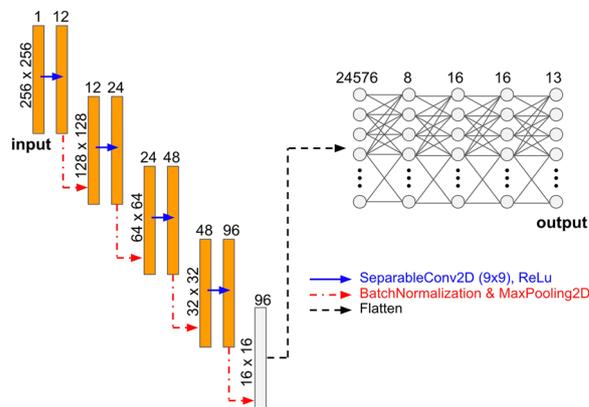
## Network architecture and training

Using ML has become ubiquitous in material science (see refs. 71–82 for a review). These applications include accelerated material discovery<sup>83–86</sup>, efficient interatomic potential development<sup>87–90</sup>, feature identification from complex patterns that have relevance for materials performance<sup>91–95</sup>, or facilitating predictive simulations which solve macroscopic (non-linear) partial differential equation systems<sup>24,40,96</sup>. This has the potential to revolutionize continuum-based simulations of materials, allowing a substantial enhancement in the modeling of systems and topologies with high complexity. Macroscopic simulations (with or without AI) require homogenized and averaged material property descriptors, e.g., elastic constants and yield strength, as their inputs. In a multi-scale approach, these macroscopic properties are estimated using microscopic simulations like molecular dynamics or ab initio calculations<sup>97</sup>. Interestingly, ML has been also applied to replace (at least partially) the expensive microscopic simulations. These efforts include replacing the computationally costly interatomic ab initio force calculations with a neural network<sup>98</sup> or replacing the trajectory prediction using graph neural networks<sup>99</sup>. A common feature of these ML solutions is that the atomistic coordinates are transferred to the network as its input.

Unlike these approaches, in our case, we use the image of the microstructure as the input and predict the aggregated effect of the interatomic forces and the arrangements of the atoms. In this respect, our work is similar to the approach introduced in<sup>100</sup> where 2D X-ray images of the microstructure are used as the input, and macroscopic material properties like porosity, specific surface area, and average pore size of each image are computed. CNN models of homogenized mechanical properties have been trained in relation to mesoscopic microstructure with multiple phases<sup>46–48</sup>, whereby the classical continuum theory was applied to each phase and for the evaluation of the effective properties. Our approach differs fundamentally from these CNN models in the sense that we apply atomistic calculations to obtain the property which automatically maps the surface effects, namely, the elastic stiffness changes resulting from the surface reconstructions of the many atoms surrounding the pores, and thus can recapitulate the resultant size dependency.

**Neural network architecture.** Neural networks vary in their basic neural units, the arrangement of these units in the layers and their connectivity, the character of the loss functions (e.g., in terms of the quantification of the deviations of the predicted values relative to the reference values in the training data), and the “reductionist” spirit of the network design (e.g., see refs. 101–103). We build our neural network under the assumption that the macroscopic properties can be obtained as a non-linear function of the coarse-grained local information. In other words, the network extracts the local information, coarse grains it, and finally passes this information to a (learnable) non-linear function. This is reflected in the architecture of our neural network, as shown in Fig. 8.

Given that (i) the input to our network is an image and (ii) the first step is to extract the local information, we use convolutional layers which are the most widely used neural network architecture for image processing<sup>104</sup>. We apply, successively, a number of convolutional layers where each layer is accompanied by a coarse-graining step which is implemented by a max



**Fig. 8** The artificial neural network architecture used in this work. Consists of a convolutional part (left) and dense layers (right). The input is a  $256 \times 256$  pixel binary image and the output is a set of 13 scalar values of the elasticity tensor.

pooling operation. The convolutional layer is commonly coded in the form of a simple matrix (of much lower rank compared to the image size). This matrix, which is commonly referred to as the kernel, is sequentially slid across the image and multiplied at each sequential position with a subset of the input array such that the output enhances certain topological pattern features such as edges, corners, gradients, etc. As a rule of thumb, including more kernels in the design of the network leads to a better performance of the network; simply put, the network then has a larger potential for detecting features. As mentioned before, candidate quantities for these features could be gradients, edges, etc. Note that these features are not hard-coded. Instead, the network learns by itself what the important features are. The process of learning the elements of these kernels is as follows. The elements of this matrix are initially set to be random numbers but during the training phase, the neural network learns an appropriate set of values for these elements such that (together with the rest of the network) the prediction error is minimized. To enable the network to obtain different types of (necessary) information, a number of kernels are used. After feature extraction and coarse-graining, the most coarse-grained information is mapped to the homogenized properties via a number of dense layers, which is essentially a non-linear transfer function. Note that the latent space in our network is relatively large, i.e., each microstructure image is mapped into 24,576 scalar descriptors which are learned during network training. Such a large array of descriptors is of course not readily interpretable by users but will provide the flexibility to our network to condense the microstructure information and transfer its important features.

The neural network that we utilized here is very similar to the *encoder* part of the U-Net architecture<sup>102</sup>, which has been shown to be well suited for capturing solutions to solid mechanics PDEs<sup>24</sup>. Generally speaking, the encoders map the information from one space to another, where the second space is smaller in its dimensions. Later, this condensed information is passed to the decoder whose task is to map the information from the low dimensional space back to a high dimensional space. In the case of the U-Net, the high-dimensional input is an image (which is for example a RGB image showing an object). This image is transferred to the encoder and condensed to a much smaller image. Then, the condensed data is sent to the decoder which creates an image with the size of the original image (which is for example a mask indicating the positions of that object in the original image). In the current approach, we replace the decoder with a trainable non-linear function (i.e., the dense layers shown in the top right part of Fig. 8) which maps the encoded representation to the scalar output values of the elasticity tensor.

**Neural network training.** In the current implementation, for the convolutional layers a kernel of size  $k \times k$  with size  $k = 20$  is used. The kernel size used here is considerably larger (i.e.,  $\approx 6$  times) than that used in the original CNN design. The rationale behind using a larger size is to capture more features of the non-local effects. The number of channels progressively increases from one channel input data to the embedded space which has 96 channels as shown in Fig. 8. After that, we use five dense layers in conjunction with a rectified linear unit activation function, which is accompanied by the last dense layer with no activation function (i.e., the identity activation function). The training is performed by minimizing the loss function which measures the mean squared difference between the predicted and atomistically calculated values across all the 13 outputs. Note that, as it was explained in the “Results” section, only 13 components of the full elasticity tensor are independent in the current case. This error minimization is done using the ADAM optimizer<sup>105</sup>, which is a stochastic first-order gradient-based method. The parameters of the ADAM optimizer are set to  $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$ ,  $\epsilon = 10^{-7}$ , and a learning rate within the range  $[1, 2] \times 10^{-4}$ . We use random samples in batches of size 128 for the gradient estimation and continue training for 1000 epochs. In each epoch, the complete training dataset is used. The data is divided into training and validation sets as explained in the “Results” section.

## DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## CODE AVAILABILITY

The code used for machine learning in this study is open source and accessible in [github.com/tensorflow](https://github.com/tensorflow). The LAMMPS Molecular Dynamics Simulator is also open source and accessible at [lammps.org](https://lammps.org). Other codes and scripts used in this work are available from the corresponding author upon reasonable request.

Received: 8 September 2021; Accepted: 15 March 2022;

Published online: 12 April 2022

## REFERENCES

1. Ström, O. et al. Osteoporosis: burden, health care provision and opportunities in the EU. *Arch. Osteoporos.* **6**, 59–155 (2011).
2. Wimmers, G. Wood: a construction material for tall buildings. *Nat. Rev. Mater.* **2**, 17051 (2017).
3. Kim, S.-H. et al. Influence of microstructure and atomic-scale chemistry on the direct reduction of iron ore with hydrogen at 700°C. *Acta Mater.* **212**, 116933 (2021).
4. Chowdhury, A. H., Salam, N., Debnath, R., Islam, S. M. & Saha, T. in *Nanomaterials Synthesis, Micro and Nano Technologies* (eds Beeran Pottathara, Y., Thomas, S., Kalarikkal, N., Grohens, Y. & Kokol, V.) Ch. 8 (Elsevier, 2019).
5. Zhang, B., Liu, S. & Shin, Y. C. In-process monitoring of porosity during laser additive manufacturing process. *Addit. Manuf.* **28**, 497–505 (2019).
6. Lucchetta, A., Brach, S. & Kondo, D. Effects of particles size on the overall strength of nanocomposites: molecular dynamics simulations and theoretical modeling. *Mech. Res. Commun.* **114**, 103669 (2021).
7. Wang, J., Duan, H., Huang, Z. & Karihaloo, B. A scaling law for properties of nanostructured materials. *Proc. Math. Phys. Eng. Sci.* **462**, 1355–1363 (2006).
8. Chen, S. & Yao, Y. Elastic theory of nanomaterials based on surface-energy density. *J. Appl. Mech.* **81**, 121002 (2014).
9. Shuttleworth, R. The surface tension of solids. *Proc. Phys. Soc. Sect. A* **63**, 444–457 (1950).
10. Gurtin, M. E. & Ian Murdoch, A. A continuum theory of elastic material surfaces. *Arch. Ration. Mech. Anal.* **57**, 291–323 (1975).
11. Javili, A., McBride, A., Steinmann, P. & Reddy, B. D. A unified computational framework for bulk and surface elasticity theory: a curvilinear-coordinate-based finite element methodology. *Comput. Mech.* **54**, 745–762 (2014).
12. Li, Z. & Steinmann, P. Rve-based studies on the coupled effects of void size and void shape on yield behavior and void growth at micron scales. *Int. J. Plast.* **22**, 1195–1216 (2006).
13. Firooz, S., Steinmann, P. & Javili, A. Homogenization of composites with extended general interfaces: comprehensive review and unified modeling. *Appl. Mech. Rev.* **73**, 040802 (2021).
14. Wei, G., Shouwen, Y. & Ganyun, H. Finite element characterization of the size-dependent mechanical behaviour in nanosystems. *Nanotechnology* **17**, 1118–1122 (2006).
15. Stein, P., Zhao, Y. & Xu, B.-X. Effects of surface tension and electrochemical reactions in li-ion battery electrode nanoparticles. *J. Power Sources* **332**, 154–169 (2016).
16. Rezaei, S., Mianroodi, J. R., Khaledi, K. & Reese, S. A nonlocal method for modeling interfaces: numerical simulation of decohesion and sliding at grain boundaries. *Comput. Methods Appl. Mech. Eng.* **362**, 112836 (2020).
17. Rueger, Z. & Lakes, R. S. Experimental study of elastic constants of a dense foam with weak cosserat coupling. *J. Elast.* **137**, 101–115 (2019).
18. Forest, S. & Trinh, D. Generalized continua and non-homogeneous boundary conditions in homogenisation methods. *J. Appl. Math. Mech.* **91**, 90–109 (2011).
19. Kouznetsova, V., Geers, M. G. D. & Brekelmans, W. A. M. Multi-scale constitutive modelling of heterogeneous materials with a gradient-enhanced computational homogenization scheme. *Int. J. Numer. Methods Eng.* **54**, 1235–1260 (2002).
20. Geers, M., Kouznetsova, V. G. & Brekelmans, W. A. M. Multiscale first-order and second-order computational homogenization of microstructures towards continua. *Int. J. Multiscale Comput. Eng.* **1**, 371–386 (2003).
21. Guillotte, M., Godet, J. & Pizzagalli, L. A fully molecular dynamics-based method for modeling nanoporous gold. *Comput. Mater. Sci.* **161**, 135–142 (2019).
22. Patil, S. P., Rege, A., Sagardas, I., Itskov, M. & Markert, B. Mechanics of nanostructured porous silica aerogel resulting from molecular dynamics simulations. *J. Phys. Chem. B* **121**, 5660–5668 (2017).
23. van der Giessen, E. et al. Roadmap on multiscale materials modeling. *Model. Simul. Mater. Sci. Eng.* **28**, 043001 (2020).
24. Mianroodi, J. R., H. Siboni, N. & Raabe, D. Teaching solid mechanics to artificial intelligence: a fast solver for heterogeneous materials. *Npj Comput. Mater.* **7**, 99 (2021).
25. Lin, B., Bai, Y. & Xu, B.-X. Data-driven microstructure sensitivity study of fibrous paper materials. *Mater. Des.* **197**, 109193 (2021).
26. Huang, J. S., Liew, J. X., Ademiloye, A. S. & Liew, K. M. Artificial intelligence in materials modeling and design. *Arch. Comput. Methods Eng.* **28**, 3399–3413 (2021).
27. Salehi, H. & Burgueo, R. Emerging artificial intelligence methods in structural engineering. *Eng. Struct.* **171**, 170–189 (2018).
28. Peng, G. C. Y. et al. Multiscale modeling meets machine learning: what can we learn? *Arch. Comput. Methods Eng.* **28**, 1017–1037 (2021).

29. Bock, F. E. et al. A review of the application of machine learning and data mining approaches in continuum materials mechanics. *Front. Mater.* **6**, 110 (2019).
30. Bakirji, J. *Stochastic Multiscale Modeling of Complex Materials*. PhD thesis, Stanford Univ. (2020).
31. Wang, K. & Sun, W. A multiscale multi-permeability poroplasticity model linked by recursive homogenizations and deep learning. *Comput. Methods Appl. Mech. Eng.* **334**, 337–380 (2018).
32. Xue, T. et al. A data-driven computational scheme for the nonlinear mechanical properties of cellular mechanical metamaterials under large deformation. *Soft Matter* **16**, 7524–7534 (2020).
33. Le, B. A., Yvonnet, J. & He, Q.-C. Computational homogenization of nonlinear elastic materials using neural networks. *Int. J. Numer. Methods Eng.* **104**, 1061–1084 (2015).
34. Kumar, S., Tan, S., Zheng, L. & Kochmann, D. M. Inverse-designed spinoid metamaterials. *Npj Comput. Mater.* **6**, 110 (2019).
35. Wang, L., Tao, S., Zhu, P. & Chen, W. Data-driven topology optimization with multiclass microstructures using latent variable Gaussian process. *J. Mech. Des.* **143**, 031708 (2020).
36. Settgast, C., Hÿtter, G., Kuna, M. & Abendroth, M. A hybrid approach to simulate the homogenized irreversible elastic-plastic deformations and damage of foams by neural networks. *Int. J. Plast.* **126**, 102624 (2020).
37. Fuchs, A., Heider, Y., Wang, K., Sun, W. & Kaliske, M. Dnn2: a hyper-parameter reinforcement learning game for self-design of neural network based elastoplastic constitutive descriptions. *Comput. Struct.* **249**, 106505 (2021).
38. Karapiperis, K., Stainier, L., Ortiz, M. & Andrade, J. Data-driven multiscale modeling in mechanics. *J. Mech. Phys. Solids* **147**, 104239 (2021).
39. Samaniego, E. et al. An energy approach to the solution of partial differential equations in computational mechanics via machine learning: concepts, implementation and applications. *Comput. Methods Appl. Mech. Eng.* **362**, 112790 (2020).
40. Raissi, M., Perdikaris, P. & Karniadakis, G. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **378**, 686–707 (2019).
41. Yang, Z., Yu, C.-H. & Buehler, M. J. Deep learning model to predict complex stress and strain fields in hierarchical composites. *Sci. Adv.* **7**, eabd7416 (2021).
42. Wang, H., Planas, R., Chandramowlishwaran, A. & Bostanabad, R. Train once and use forever: solving boundary value problems in unseen domains with pre-trained deep learning models. Preprint at <https://arxiv.org/abs/2104.10873> (2021).
43. Pandey, A. & Pokharel, R. Machine learning based surrogate modeling approach for mapping crystal deformation in three dimensions. *Scr. Mater.* **193**, 1–5 (2021).
44. Masi, F., Stefanou, I., Vannucci, P. & Maffi-Berthier, V. Thermodynamics-based artificial neural networks for constitutive modeling. *J. Mech. Phys. Solids* **147**, 104277 (2021).
45. Fernández, M., Jamshidian, M., Böhle, T., Kersting, K. & Weeger, O. Anisotropic hyperelastic constitutive models for finite deformations combining material theory and data-driven approaches with application to cubic lattice metamaterials. *Comput. Mech.* **67**, 653–677 (2021).
46. Zhang, X. & Garikipati, K. Machine learning materials physics: multi-resolution neural networks learn the free energy and nonlinear elastic response of evolving microstructures. *Comput. Methods Appl. Mech. Eng.* **372**, 113362 (2020).
47. Li, X. et al. Predicting the effective mechanical property of heterogeneous materials by image based modeling and deep learning. *Comput. Methods Appl. Mech. Eng.* **347**, 735–753 (2019).
48. Cecen, A., Dai, H., Yabansu, Y. C., Kalidindi, S. R. & Song, L. Material structure-property linkages using three-dimensional convolutional neural networks. *Acta Mater.* **146**, 76–84 (2018).
49. Rao, C. & Liu, Y. Three-dimensional convolutional neural network (3d-cnn) for heterogeneous material homogenization. *Comput. Mater. Sci.* **184**, 109850 (2020).
50. Swetlana, S., Khatavkar, N. & Singh, A. K. Development of Vickers hardness prediction models via microstructural analysis and machine learning. *J. Mater. Sci.* **55**, 15845–15856 (2020).
51. Ford, E., Maneparambil, K., Rajan, S. & Neithalath, N. Machine learning-based accelerated property prediction of two-phase materials using microstructural descriptors and finite element analysis. *Comput. Mater. Sci.* **191**, 110328 (2021).
52. Liu, R. et al. Context aware machine learning approaches for modeling elastic localization in three-dimensional composite microstructures. *Integr. Mater. Manuf. Innov.* **6**, 160–171 (2017).
53. Xu, H., Liu, R., Choudhary, A. & Chen, W. A machine learning-based design representation method for designing heterogeneous microstructures. *J. Mech. Des.* **137**, 051403 (2015).
54. Roters, F. et al. DAMASK - The Düsseldorf Advanced Material Simulation Kit for modeling multi-physics crystal plasticity, thermal, and damage phenomena from the single crystal up to the component scale. *Comput. Mater. Sci.* **158**, 420–478 (2019).
55. Liu, Z. & Wu, C. Exploring the 3D architectures of deep material network in data-driven multiscale mechanics. *J. Mech. Phys. Solids* **127**, 20–46 (2019).
56. Balokas, G., Czichon, S. & Rolles, R. Neural network assisted multiscale analysis for the elastic properties prediction of 3D braided composites under uncertainty. *Compos. Struct.* **183**, 550–562 (2018).
57. Mishin, Y., Farkas, D., Mehl, M. J. & Papaconstantopoulos, D. A. Interatomic potentials for monoatomic metals from experimental data and ab initio calculations. *Phys. Rev. B* **59**, 3393–3407 (1999).
58. Sadd, M. H. In *Elasticity (Fourth Edition)* 4th edn (ed. Sadd, M. H.) Ch. 11 (Academic Press, 2021).
59. Roters, F. et al. Overview of constitutive laws, kinematics, homogenization and multiscale methods in crystal plasticity finite-element modeling: theory, experiments, applications. *Acta Mater.* **58**, 1152–1211 (2010).
60. Rueger, Z. & Lakes, R. S. Experimental cosserat elasticity in open-cell polymer foam. *Philos. Mag.* **96**, 93–111 (2016).
61. Xu, K., Huang, D. Z. & Darve, E. Learning constitutive relations using symmetric positive definite neural networks. *J. Comput. Phys.* **428**, 110072 (2021).
62. Deshpande, V., Fleck, N. & Ashby, M. Effective properties of the octet-truss lattice material. *J. Mech. Phys. Solids* **49**, 1747–1769 (2001).
63. Kulagin, R., Beygelzimer, Y., Estrin, Y., Schumilin, A. & Gumbsch, P. Architected lattice materials with tunable anisotropy: design and analysis of the material property space with the aid of machine learning. *Adv. Eng. Mater.* **22**, 2001069 (2020).
64. Liu, Y., St-Pierre, L., Fleck, N., Deshpande, V. & Srivastava, A. High fracture toughness micro-architected materials. *J. Mech. Phys. Solids* **143**, 104060 (2020).
65. Guo, X., Chen, Z., Chu, X. & Wan, J. A plane stress model of bond-based cosserat peridynamics and the effects of material parameters on crack patterns. *Eng. Anal. Bound. Elem.* **123**, 48–61 (2021).
66. Rezaei, S., Mianroodi, J. R., Brepols, T. & Reese, S. Direction-dependent fracture in solids: atomistically calibrated phase-field and cohesive zone model. *J. Mech. Phys. Solids* **147**, 104253 (2021).
67. Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **117**, 1–19 (1995).
68. Jacobs, P., Zhukovskii, Y., Mastrokov, Y. & Shunin, Y. Bulk and surface properties of metallic aluminium: DFT simulations. *Comput. Model. Simul. Eng.* **6**, 7–28 (2002).
69. Bitzek, E., Koskinen, P., Gähler, F., Moseler, M. & Gumbsch, P. Structural relaxation made simple. *Phys. Rev. Lett.* **97**, 170201 (2006).
70. Stukowski, A. Visualization and analysis of atomistic simulation data with ovito: the open visualization tool. *Model. Simul. Mater. Sci. Eng.* **18**, 015012 (2010).
71. Sha, W. et al. Artificial intelligence to power the future of materials science and engineering. *Adv. Intell. Syst.* **2**, 1900143 (2020).
72. Ye, W., Chen, C., Wang, Z., Chu, I. H. & Ong, S. P. Deep neural networks for accurate predictions of crystal stability. *Nat. Commun.* **9**, 1–6 (2018).
73. Butler, K. T., Davies, D. W., Cartwright, H., Isayev, O. & Walsh, A. Machine learning for molecular and materials science. *Nature* **559**, 547–555 (2018).
74. Schmidt, J., Marques, M. R. G., Botti, S. & Marques, M. A. L. Recent advances and applications of machine learning in solid-state materials science. *Npj Comput. Mater.* **5**, 83 (2019).
75. Ramprasad, R., Batra, R., Pailania, G., Mannodi-Kanakithodi, A. & Kim, C. Machine learning in materials informatics: recent applications and prospects. *Npj Comput. Mater.* **3**, 54 (2017).
76. Niezgod, S. R., Kanjarla, A. K. & Kalidindi, S. R. Novel microstructure quantification framework for databasing, visualization, and analysis of microstructure data. *Integr. Mater. Manuf. Innov.* **2**, 54–80 (2013).
77. McDowell, D. L. & Lesar, R. A. The need for microstructure informatics in process-structure-property relations. *MRS Bull.* **41**, 587–593 (2016).
78. Bereau, T., Andrienko, D. & Kremer, K. Research update: computational materials discovery in soft matter. *APL Mater.* **4**, 053101 (2016).
79. Wodo, O., Broderick, S. & Rajan, K. Microstructural informatics for accelerating the discovery of processing-microstructure-property relationships. *MRS Bull.* **41**, 603–609 (2016).
80. Agrawal, A. & Choudhary, A. Perspective: Materials informatics and big data: realization of the “fourth paradigm” of science in materials science. *APL Mater.* **4**, 053208 (2016).
81. Dimiduk, D. M., Holm, E. A. & Niezgod, S. R. Perspectives on the impact of machine learning, deep learning, and artificial intelligence on materials, processes, and structures engineering. *Integr. Mater. Manuf. Innov.* **7**, 157–172 (2018).
82. Masi, F., Stefanou, I., Vannucci, P. & Maffi-Berthier, V. Thermodynamics-based artificial neural networks for constitutive modeling. *J. Mech. Phys. Solids* **147**, 104277 (2021).
83. Raccuglia, P. et al. Machine-learning-assisted materials discovery using failed experiments. *Nature* **533**, 73–76 (2016).

84. Mefstahi, N. et al. Machine learning property prediction for organic photovoltaic devices. *Npj Comput. Mater.* **6**, 1–8 (2020).
85. Sun, W. et al. Machine learning-assisted molecular design and efficiency prediction for high-performance organic photovoltaic materials. *Sci. Adv.* **5**, 1–9 (2019).
86. Dai, D. et al. Using machine learning and feature engineering to characterize limited material datasets of high-entropy alloys. *Comput. Mater. Sci.* **175**, 109618 (2020).
87. Behler, J. & Parrinello, M. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.* **98**, 146401 (2007).
88. Kolb, B., Lentz, L. C. & Kolpak, A. M. Discovering charge density functionals and structure-property relationships with PROPhet: a general framework for coupling machine learning and first-principles methods. *Sci. Rep.* **7**, 1192 (2017).
89. Behler, J. Perspective: Machine learning potentials for atomistic simulations. *J. Chem. Phys.* **145**, 170901 (2016).
90. Dragoni, D., Daff, T. D., Csányi, G. & Marzari, N. Achieving dft accuracy with a machine-learning interatomic potential: thermomechanics and defects in bcc ferromagnetic iron. *Phys. Rev. Mater.* **2**, 013808 (2018).
91. Jiang, Y. et al. Topological representations of crystalline compounds for the machine-learning prediction of materials properties. *Npj Comput. Mater.* **7**, 1–8 (2021).
92. Adachi, Y., Taguchi, M. & Hirokawa, S. Microstructure recognition by deep learning. *Tetsu-to-Hagane* **102**, 722–729 (2016).
93. DeCost, B. L., Francis, T. & Holm, E. A. Exploring the microstructure manifold: Image texture representations applied to ultrahigh carbon steel microstructures. *Acta Mater.* **133**, 30–40 (2017).
94. DeCost, B. L. & Holm, E. A. A computer vision approach for automated analysis and classification of microstructural image data. *Comput. Mater. Sci.* **110**, 126–133 (2015).
95. DeCost, B. L. & Holm, E. A. Characterizing powder materials using keypoint-based computer vision methods. *Comput. Mater. Sci.* **126**, 438–445 (2017).
96. Rad, M. T., Viardin, A. & Apel, M. On theory-training neural networks to infer the solution of highly coupled differential equations. Preprint at <https://arxiv.org/abs/2102.04890> (2021).
97. Fernández, M., Rezaei, S., Rezaei Mianroodi, J., Fritzen, F. & Reese, S. Application of artificial neural networks for the prediction of interface mechanics: a study on grain boundary constitutive behavior. *Adv. Model. Simul. Eng. Sci.* **7**, 1 (2020).
98. Behler, J. & Parrinello, M. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.* **98**, 146401 (2007).
99. Sanchez-Gonzalez, A. et al. Learning to simulate complex physics with graph networks. Preprint at <https://arxiv.org/abs/2002.09405> (2020).
100. Alqahtani, N., Alzubaidi, F., Armstrong, R. T., Swietojanski, P. & Mostaghimi, P. Machine learning for predicting properties of porous media from 2d x-ray images. *J. Pet. Sci. Eng.* **184**, 106514 (2020).
101. Hochreiter, S. & Schmidhuber, J. Long short-term memory. *Neural Comput.* **9**, 1735–1780 (1997).
102. Ronneberger, O., Fischer, P. & Brox, T. U-net: convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-assisted Intervention* 234–241 (Springer, 2015).
103. Lechner, M. et al. Neural circuit policies enabling auditable autonomy. *Nat. Mach. Intell.* **2**, 642–652 (2020).
104. Chollet, F. et al. *Deep learning with Python*, Vol. 361 (Manning, 2018).
105. Kingma, D. P. & Ba, J. Adam: a method for stochastic optimization. Preprint at <https://arxiv.org/abs/1412.6980> (2014).

## ACKNOWLEDGEMENTS

We thank Professor Bob Svendsen for fruitful discussions during this work.

## AUTHOR CONTRIBUTIONS

J.R.M., S.R., and N.H.S. developed the initial concept and workflow. J.R.M. carried out the atomistic calculations and the neural network training. N.H.S. designed the neural network and carried out the network analysis. S.R. performed the finite element calculations. B.-X.X. and D.R. supervised the work and helped in the interpretation of the results. All authors discussed the paper and contributed to preparing the manuscript.

## FUNDING

Open Access funding enabled and organized by Projekt DEAL.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Correspondence** and requests for materials should be addressed to Jaber Rezaei Mianroodi.

**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) 2022